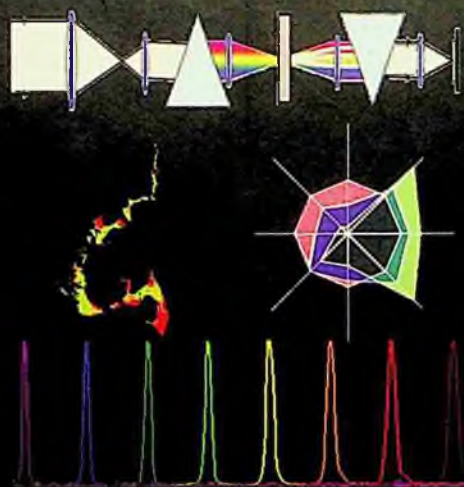


SPECTRAL SENSING RESEARCH FOR WATER MONITORING APPLICATIONS AND FRONTIER SCIENCE AND TECHNOLOGY FOR CHEMICAL, BIOLOGICAL AND RADIOLOGICAL DEFENSE

Editors

Dwight Woolard

Janet Jensen



**SPECTRAL SENSING RESEARCH
FOR WATER MONITORING
APPLICATIONS AND FRONTIER
SCIENCE AND TECHNOLOGY FOR
CHEMICAL, BIOLOGICAL AND
RADIOLOGICAL DEFENSE**

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

Selected Topics in Electronics and Systems — Vol. 48

**SPECTRAL SENSING RESEARCH FOR WATER MONITORING APPLICATIONS AND
FRONTIER SCIENCE AND TECHNOLOGY FOR CHEMICAL, BIOLOGICAL
AND RADIOLOGICAL DEFENSE**

Copyright © 2008 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN-13 978-981-283-323-5

ISBN-10 981-283-323-4

Editor: Tjan Kwang Wei



FOREWORD

This special journal issue contains select papers from the *2006 International Symposium on Spectral Sensing Research (2006 ISSSR)* that fall into the subject matter areas of: (1) Multispectral/Hyperspectral Techniques for Water Monitoring Applications; and (2) Frontier Spectroscopic Science and Technology for Chemical, Biological and Radiological (CB&R) Defense. The specific focus of the *2006 ISSSR* was on the creation of new technology-program oriented networks that will serve as a research and development foundation for the advancement of the state-of-the-art in spectroscopic-based early-warning sensor capabilities. In recent years, spectral sensing has experienced rapid technical advancement that has led to practical field sensors. Therefore spectral-based techniques exhibit the clear potential for providing more effective, economical and supportable (i.e., reagentless) solutions to military and homeland defense early-warning monitoring requirements for water, surface and air related sensing applications.

An increased emphasis on *reagentless* spectroscopy is motivated primarily by performance issues associated with traditional chemical and biological (CB) point and standoff techniques. In particular, sensors have been previously developed and fielded that rely heavily on reagents and/or burdensome support structures that are expensive and difficult to maintain and that have serious false alarm issues. Previously implemented technologies include biological assays, mass spectrometry and ion mobility. Other explored methodologies include novel materials (mips, smart ligands, amino acid sequences, aptamers, sol gel, aerogel, electro-conducting polymers, etc.) or bulk property interactions (electrochemistry, surface acoustic wave, surface plasmon resonance, thermal capacity) and combinations of the two.

At this time, extensive expertise exists in the multispectral/hyperspectral community for applications such as airborne and space-based sensing and imaging which has proved effective in monitoring weather, resource management (agriculture, forestry), oil/mineral deposits and CB detection in air releases. Hence, spectral-based techniques clearly have potential for providing near to mid-term solutions for many of the monitoring problems associated with CB&R contaminations of water, surfaces and air. However, the ultimate realization of such spectroscopic techniques will probably require the fusion of many types of spectral-sensing techniques and modality. Therefore, standoff and point interrogation sensors are now sought that can provide for extremely high confidence in CB&R detection and monitoring scenarios and the goal of the *2006 ISSSR* and this

companion special issue is to organize and focus the science and technology base towards these important challenges.

This special issue begins with a collection of research and development papers on the subject of spectroscopic techniques for “Water Sensing and Monitoring.” These papers address the Joint Services issued formal requirements for new capabilities in early warning monitoring of water supplies against a CB event. These requirements put a new type of problem on the table that is considered to be “non-traditional” in a number of ways. Specifically, new ways of thinking and innovative approaches and technology are needed to solve this problem. For example, the civilian and military drinking water providers have extensive history, expertise and a proven track record that is unsurpassed at delivering high quality, safe drinking waters to their customers. But these standards and monitoring methods were developed to address traditional public health requirements, to be able to protect the public health in a court of law and to provide a level of validated data that could be used to guide or modify treatment protocols over extended time periods. Detection is based on classical laboratory methods that require extensive sample collection, handling, transportation, labor, and logistics. Hence, it is accepted and expected to take days to conduct and weeks to report.

In contrast the Joint requirements call for “early warning” monitoring. “Early warning” must be “real-time”, “in-line”, low cost, low labor, easy to support, maintain and use. A mandate for “reagent less” technologies is given to reduce costs of disposables, logistics support, contamination and fouling in the field. An early warning monitor is not developed as a tool for “detect-to-treat” or “detect-to-sue”, but simply as a type of smoke-alarm to “detect-to-warn” to alert a population of an abnormal toxic or pathological event. It must provide enough time so that people can reduce or eliminate catastrophic exposure. The goal is to prevent human morbidity and mortality. It may also be used as a trigger to invoke more extensive and classical methods for forensic confirmation or treatment.

The Standoff Detection and Spectral Sensing communities are accustomed to rapid data acquisition and reporting, early warning monitoring and contamination avoidance. Standoff Detection was created and developed for the purpose of early warning and contamination avoidance. They have worked since the 1960s to develop the types of technologies that can provide this type of monitoring for air and more recently surfaces, but historically have not been asked to address water. Water for these communities is considered a “non-traditional” background and has not received consideration or dialogue. A deep well of technological capability exists in this community and it needs to be pumped for water answers. To this end, the goal of the ISSSR 2006 water monitoring session was to bring together these different expert communities to foster new thoughts and collaborations, identify new approaches, new ways of using existing

technologies, highlight emerging and enabling technologies with the focus on developing new and novel solutions to answer the Joint Service Requirements.

This special contains a very good representative sampling of the papers that were presented at the *2006 ISSSR* on the subject of spectroscopic techniques for “Water Sensing and Monitoring.” In particular, the special issue begins with papers that focus on the subject of real-time, on-line monitoring of drinking water distribution networks for both chemical and biological agent threats. The subject matter then shifts towards innovations such as chip-size wavelength detectors that allow for highly accurate and spectrally broad, compact spectrometers for optical identification of bio-agents in fluids, and multi-color native fluorescence spectroscopy that enables identification of multiple classes of bio-molecules. This special issue also contains papers from the diverse subject areas: Water quality monitoring using satellite imagery; Deep UV laser-induced fluorescence (LIF) for the detection of trace species and dissolved organic compounds in water; and, Materials and sampling protocols for effective coupling to miniaturized and portable FTIR for aqueous-based detection of toxic compounds. Papers are also included that focus on emerging techniques such as: Terahertz spectroscopy of proteins in water that utilize both direct absorption and circular dichroism; and, a novel Surface enhanced Raman Spectroscopy (SERS) active material with enhanced sensitivity, specificity and speed in detecting chemical agents and their hydrolysis products. The water monitoring section concludes with the two novel sample preparation methodologies: Front-end liquid flow cells which utilize electric field concentration methods for collecting spores onto the surface of an attenuated total reflection (ATR) IR crystal; and, Liquid-Flow concentration with a reversible filter that enables preconcentration of microorganisms into a tiny volume of liquid for enhanced spectral detection.

This special issue also contains a collection of technical papers on the subject of “Frontier and Emerging Spectroscopic Science and Technology” that have relevance to a general array of CB&R defense and security applications. Here, the initial focus of this issue is on new scientific techniques and advanced technology that offer the potential for enhancing and extending the effectiveness of existing hyperspectral/multispectral sensing methodologies. Examples of these research and development efforts include: Adaptive spectrometers and multispectral imagers that allow for real-time optimization of standoff CB sensor data; Physics-based algorithms for the more effective detection of gaseous effluents using airborne LWIR hyperspectral imagery of complex scenes; Application of combined infrared and Raman spectroscopy for collecting the vibrational signatures of chemical warfare agent simulants, their hydrolysis and degradation products and a number of toxic industrial compounds (TICs); Development of double-transducer noncollinear acousto-optic tunable filter (AOTF) cells that enable enhanced visible to midwave infrared spectral imaging; Demonstration of a remotely operable miniature spectrometer for the detection of TICs using near infrared spectroscopy; Utilization of multidimensional phase shaped femtosecond pulses to control the fragmentation and

ionization of mass spectrometry; New insights into surface-plasmon-resonance phenomenon for optical sensing applications; Demonstration of an optical characterization unit based upon fluorescence spectroscopy-on-a-chip that utilizes a compact fluidic platform to achieve point detection and analysis of pathogens; and, Enhanced UV light emitting diodes with higher efficiencies, increased power and extended lifetimes for reagentless bio-agent detection and water purification.

The later portion of this special issue is on novel technology and phenomenology that has potential for impacting spectroscopic-based sensing in the future. Some examples of these research and development efforts include: Identification of a novel semiconductor-based hetero-systems that allows for the design of an optically-triggered interband resonant-tunneling-diode (OT-I-RTD) oscillator capable of producing significant levels of terahertz (THz) frequency output power; Presentation of a novel THz detector concept that is based upon magnetoresistance oscillations in two-dimensional electron systems in semiconductors that allows for the simultaneous determination of the radiation frequency and intensity from a simple electrical measurement; Investigations of a novel quantum-mechanical phenomenon associated with the electromagnetic response of cold-atom systems that offers the potential for completely new sensing modalities from RF to the far-infrared (terahertz); Studies on dilute-magnetic-semiconductor (DMS) based multi-barrier structures offering a new type of very high speed transistors for integrated sensor applications at THz frequencies; Research on InN and In_n/InGa_n nanostructures that show their promise for THz sources pumped by low cost 1550 nm fs fiber and CW lasers; Introduction of a novel type of fluctuation-enhanced sensing that exploits the information contained in microfluctuations to increase both selectivity and sensitivity of conventional chemical sensors; Quantification of phenomenology of time-resolved visible, near- and mid-IR spectra from fast-transient high-explosive fireballs showing its utility for explosive classification; Development of a new broadband Chirped-Pulse Fourier Transform Microwave (CP-FTMW) spectrometer that allows for high resolution and high throughput analysis of chemical warfare agents even when present in mixtures; Studies that quantify the environmental effects on the THz spectra of DNA and on the THz spectra of nanostructures useful for the study of the vibrational, electronic and optical properties of DNA; Demonstration of a compact ultrasensitive wideband superconducting mm/submm integrated digital spectrometer for chemical and biological agent detection; Illustration of the effectiveness of laser-ionization mass spectroscopy for the detection of explosives and chemical warfare agents; Development of in-line optical fiber structures that allow for large interaction lengths and high sensitivity in surface plasmon and evanescent sensors for chemical and biological agents; Design of functional surfaces for application in point sensing of a single or a few bio-molecular targets to enable label-free ultra-low concentration detection; Development of hybrid films that possess fluorescence intensity dependency on exposure to methanol vapors; and, Introduction of a novel nanoscale imaging technology for application in THz-frequency transmission microscopy based sensing.

Finally, the organizing committee *2006 ISSSR* and the Editors of this special issue would like to recognize the following Best Paper Presentations that lead off this collection of technical papers:

Programmable Adaptive Spectral Imagers for Mission-Specific Application in Chemical/Biological Sensing

By Neil Goldstein, *Spectral Sciences, Inc*

Design & Optimization of an I-RTD Hybrid THz Oscillator Based Upon $\text{In}_{1-x}\text{Ga}_x\text{As} / \text{GaSb}_y\text{As}_{1-y}$ Heterostructure Systems

By Weidong Zhang, *North Carolina State University*

Study of Transport and Devices Based on the Photo-Excited Two-Dimensional Electronic System

By Ramesh Mani, *Harvard University*

Editors:

Janet Jensen, U.S. Army Edgewood Chemical Biological Center
Dwight Woolard, U.S. Army Research Office

CONTENTS

Foreword	v
Water Sensing and Monitoring Sessions	
Operational and Laboratory Verification Testing of a Heuristic On-Line Water Monitoring System for Security <i>D. Kroll and K. King</i>	3
Real-Time On-Line Monitoring of Drinking Water for Waterborne Pathogen Contamination Warning <i>J. A. Adams and D. McCarty</i>	15
Chip-Size Wavelength Detectors <i>O. Schmidt, P. Kiesel, M. Bassler and N. Johnson</i>	33
Class Identification of Bio-Molecules Based on Multi-Color Native Fluorescence Spectroscopy <i>M. Bassler, O. Schmidt, P. Kiesel and N. M. Johnson</i>	43
Chesapeake Bay Water Quality Monitoring Using Satellite Imagery <i>B. Ramakrishna, C-I Chang, B. Trou and J. Henqemihle</i>	53
LIF Detection of Trace Species in Water Using Different UV Laser Wavelengths <i>A. V. Sharikova and D. K. Killinger</i>	61
Development of Materials and Sampling Methods for IR-Based Detection of Toxic Compounds in Water <i>B. J. Ninness, L. D. Doucette, B. McCool and C. P. Tripp</i>	69
THz Spectroscopy of Proteins in Water: Direct Absorption and Circular Dichroism <i>J. Xu, K. W. Plaxco and S. J. Allen</i>	81

A SERS-Based Analyzer for Point and Continuous Water Monitoring of Chemical Agents and Their Hydrolysis Products <i>S. Farquharson and F. E. Inscore</i>	91
Reagentless Bio-Sampling Methods for IR Detection <i>L. D. Doucette, H. Li, B. J. Ninness and C. P. Tripp</i>	101
PreConcentration of Microorganisms into a Tiny Volume of Liquid for Enhanced Spectral Detection <i>S. Zaromb, D. Martell, N. Schattke and G. Hankins</i>	111
Frontier Session	
Programmable Adaptive Spectral Imagers for Mission-Specific Application in Chemical/Biological Sensing <i>N. Goldstein, P. Vujkovic-Cvijin, M. Fox, S. Adler-Golden, J. Cline, B. Gregor, J. Lee, A. C. Samuels, S. D. Higbee, L. S. Becker and T. Ooi</i>	121
Design & Optimization of an I-RTD Hybrid THz Oscillator Based Upon $\text{In}_{1-x}\text{Ga}_x\text{As}/\text{GaSb}_y\text{As}_{1-y}$ Heterostructure Systems <i>W. Zhang, D. Woolard, E. Brown, B. Gelmont and R. Trew</i>	133
Study of Transport and Devices Based on the Photo-Excited Two-Dimensional Electronic System <i>R. G. Mani</i>	149
Terahertz-Based Detectors Using Cold-Atom Optics <i>F. J. Crowne, W. M. Golding and C. Hazelton</i>	161
Detection of Gaseous Effluents from Airborne LWIR Hyperspectral Imagery Using Physics-Based Signatures <i>D. W. Messinger, C. Salvaggio and N. M. Sinisgalli</i>	173
Nonlinear Chemical Plume Detection Using Kernel-Based Matched Subspace Detectors <i>H. Kwon, N. M. Nasrabadi and P. Gillespie</i>	185
Vibrational Spectroscopy of Chemical Agents Simulants, Degradation Products of Chemical Agents and Toxic Industrial Compounds <i>S. P. Hernández-Rivera, L. C. Pacheco-Londoño, O. M. Primera-Pedrozo, O. Ruiz, Y. Soto-Feliciano and W. Ortiz</i>	199

Hyperspectral and Polarization Imaging with Double-Transducer AOTFS for Wider Spectral Coverage <i>N. Gupta</i>	217
Deep-UV Based Acousto-Optic Tunable Filter for Spectral Sensing Applications <i>N. S. Prasad</i>	229
Development of Miniature Acousto Optic Tunable Filter (AOTF) Spectrometer for Detection of Toxic Industrial Chemicals (TICS) <i>P. J. Stopa, J. I. Soos, R. G. Rosemeier, S. B. Trivedi and S. W. Kutcher</i>	239
Spin Dependent Wigner Function Simulations of Diluted Magnetic Semiconductor Superlattices – B Field Tuning <i>H. L. Grubin</i>	249
Indium Nitride: A New Material for High Efficiency, Compact, 1550nm Laser-Based Terahertz Sources in Chemical and Biological Detection <i>M. Wraback, G. D. Chern, E. D. Readinger, P. H. Shen, G. Koblmüller, C. Gallinat and J. S. Speck</i>	261
Fluctuation-Enhanced Chemical/Biological Sensing and Prompt Identification of Bacteria by Sensing of Phage-Triggered Ion Cascade (SEPTIC) <i>L. B. Kish, G. Schmera, M. D. King, M. Cheng, R. Young and C. G. Granqvist</i>	269
The Phenomenology of High Explosive Fireballs from Fielded Spectroscopic and Imaging Sensors for Event Classification <i>K. C. Gross and G. P. Perram</i>	277
Chirped-Pulse Fourier Transform Microwave Spectroscopy: A New Technique for Rapid Identification of Chemical Agents <i>J. J. Pajski, M. D. Logan, K. O. Douglass, G. G. Brown, B. C. Dian, B. H. Pate and R. D. Suenram</i>	289
Environmental Effects Influencing the Vibrational Modes of DNA: Nanostructures Coupled to Biomolecules <i>D. Ramadurai, T. Yamanaka, M. Vasudev, Y. Li, V. Sankar, M. Dutta, M. A. Strocio, T. Rajh, Z. Saponjic and S. Xu</i>	305
Multidimensional Identification of Chemical Warfare Agents Using Shaped Femtosecond Pulses <i>I. Pastirk and M. Dantus</i>	321

Surface-Plasmon-Resonance Based Optical Sensing <i>N. J. M. Horing and H. L. Cui</i>	329
Long-Wave Infrared and Terahertz-Frequency Lasing Based on Semiconductor Nanocrystals <i>V. I. Rupasov and S. G. Krivoslykov</i>	337
Ultrasensitive Wideband Integrated Spectrometer for Chemical and Biological Agent Detection <i>I. V. Vernik</i>	345
Compact Optical Characterization Platform for Detection of Bio-Molecules in Fluidic and Aerosol Samples <i>P. Kiesel, O. Schmidt, M. Bassler and N. Johnson</i>	357
Resonant Terahertz Spectroscopy of Bacterial Thioredoxin in Water: Simulation and Experiment <i>A. Bykhovski, T. Globus, T. Khromova, B. Gelmont and D. Woolard</i>	367
Study of Lennard-Jones Clusters: Effects of Anharmonicities Far from Saddle Points <i>L. Y. Chen and N. J. M. Horing</i>	377
van der Waals Interaction Between Adsorbate Layers/Molecules and a Substrate <i>N. J. M. Horing</i>	385
The Spin-Hall Effect in <i>p</i> -Type Bulk Semiconductors <i>S. Y. Liu, N. J. M. Horing and X. L. Lei</i>	395
Tunable Grid Gated Double-Quantum-Well FET Terahertz Detector <i>N. J. M. Horing, S. Y. Liu, V. V. Popov and H. L. Cui</i>	405
Laser-Ionization Mass Spectrometry of Explosives and Chemical Warfare Simulants <i>D. L. Huestis, C. Mullen, M. J. Coggiola and H. Oser</i>	419
In-Line Optical Fiber Structures for Environmental Sensing <i>A. Dhawan, M. D. Gerhold and J. F. Muth</i>	425

Progress in High Efficiency UV LED Research for Reagentless Bioagent Detection and Water Purification	437
<i>M. L. Reed, G. A. Garrett, A. V. Sampath, P. H. Shen, C. J. Collins, M. Wraback, J. Zhang, X. Hu, J. Deng, Alex, Lunev, Y. Bilenko, T. Katona and Remis Gaska</i>	
Towards Sensing Single or a Few Bio-Molecular Architectures: Design of Functional Surfaces	445
<i>P. Zhao, D. Woolard, J. M. Seminario and R. Trew</i>	
Coumarin Dye as a Fluorescence Sensor for Methanol Vapor	453
<i>N. Stevens and D. L. Akins</i>	
Nanoscale Imaging Technolgy for THz-Frequency Transmission Microscopy	463
<i>D. Woolard, P. Zhao, C. Rutherglen, Z. Yu, P. Burke, S. Brueck and A. Stintz</i>	
Predicting the Path of Electronic Transport through a Molecular Device via a Mountain-Pass Algorithm	481
<i>G. Recine and D. L. Woolard</i>	

ISSR 2006

**WATER SENSING AND
MONITORING SESSIONS**

OPERATIONAL AND LABORATORY VERIFICATION TESTING OF A HEURISTIC ON-LINE WATER MONITORING SYSTEM FOR SECURITY

DAN KROLL

*HACH HOMELAND SECURITY TECHNOLOGIES, 5600 Lindbergh Drive
LOVELAND, COLORADO 80539 USA
DKROLL@hach.com*

KARL KING

*HACH HOMELAND SECURITY TECHNOLOGIES, 5600 Lindbergh Drive
LOVELAND, COLORADO 80539 USA
KKING@hach.com*

Drinking water systems are vulnerable to attack. While homeland security initiatives commonly focus on aerial CW/BW attacks, they have tended to ignore the far more inexpensive and easy to orchestrate attacks on drinking water distribution networks. An approach that utilizes off-the-shelf broad-spectrum analytical instruments coupled with advanced interpretive algorithms to provide detection-response networks for water is described. This report summarizes the development of interpretive algorithms applied to drinking water instrumentation and the implications for immediately engineering and deploying distribution based detection-protection systems. Data obtained from in house testing along with testing at ECBC was used to produce fingerprint response data from traditional as well as non-military threat agents. Data obtained from a Battelle/EPA ETV study addresses issues such as long-term deployment and ability to detect and characterize contaminants. Loop testing at ECBC and Battelle demonstrates extrapolation of beaker testing fingerprints to flowing systems. Real world deployment data is used to demonstrate recognition and classification of actual events and heuristic capabilities of the system along with its potential role in enhancing water quality above and beyond its obvious security aspects. The system is shown to be a practical measure to detect and characterize backflow events involving both chemicals and bioagents.

Keywords: Water Distribution; Monitoring; On-Line; Algorithm, Multi-parameter

1. Our Water Systems are Vulnerable

The water that we use in our homes and businesses is supplied to us through a vast network of interconnected pipes known as the distribution system. The vulnerability of this system to attack is not a subject that most people dwell upon. Initially after the attacks of 9/11, government experts tasked with ensuring the safety of our water supply declared the system to be secure. EPA administrator Christie Whitman stated on

10/18/2001, *“People are worried that a small amount of some chemical or biological agent –a few drops for instance- could result in significant threats to the health of large numbers of people. I want to assure people –that scenario can’t happen. It would take large amounts to threaten the safety of a city water system. We believe it would be very difficult for anyone to introduce the quantities needed to contaminate an entire system.”*¹

Unfortunately these sorts of statements apply only to large reservoirs and water sources and not to the distribution system. Various industry experts soon pointed out the error of declaring our water supply to be safe. Researchers from the US Air Force, Army Corps of Engineers and Hach Homeland Security Technologies (HST) have calculated that an attack on drinking water distribution systems can be mounted for between \$0.05 and \$5.00 per death, using rudimentary techniques, and amass casualties in the thousands over a period of hours via a method known as a backflow attack.^{2,3,4,5}

Wherever water can be drawn from the distribution system, it would be possible to use a pump to inject material back into the system. Once the contaminant is present in the pipes, the normal movement of water in the system acts to disseminate the contaminant throughout the network. The introduction point can be anywhere in the system. Due to the vast expanse of the distribution network, physically securing the system against such attack becomes impossible. That leaves monitoring as the only viable method to protect against such events.

2. Distribution System Monitoring

Monitoring in the distribution system is a difficult proposition. The sheer number and diversity of potential threat agents that could be utilized in an attack against the system makes monitoring for them on an individual basis an effort that is doomed to failure from the start. What is needed is a broad-spectrum analyzer that can respond to any possible threat and even unknown or unanticipated events.

The difficult mission of detecting such a wide variety of potential threats is not the only challenge confronting a monitoring system for the distribution system. The environment that any such sensor would be exposed to is extremely harsh. Extreme variability in water conditions is routinely encountered in the pipes. Much of the existing water supply infrastructure is also aging and in poor condition. This results in conditions of corrosion and scaling that may cause the fouling of sensors that are not robust enough to operate under such conditions on an extended deployment timeframe. Biofilms may also form on exposed surfaces leading to sensor failure over time. What is needed is an extremely rugged sensor that is capable of withstanding long-term deployment and has the ability to respond to all types of threat agents.

3. The Hach HST Approach

Rather than attempting to develop individual sensors to detect contaminants or classes of contaminants, the Hach HST approach was to utilize a sensor suite of commonly available off-the-shelf water quality monitors such as pH, electrolytic conductivity, turbidity, chlorine residual and total organic carbon (TOC) linked together in an

intelligent network. The logic behind this is that these are tried and true technologies that have been extensively deployed in the water supply industry for a number of years and have proven to be stable in such situations. One of the difficulties encountered when designing such a device is that the normal fluctuations in these parameters found within the water can be quite pronounced.

The problem then becomes, can we differentiate between the changes that are seen as a result of the introduction of a contaminant and those that are a result of everyday system perturbation? The secret to success, in a situation such as this, is to have a robust and workable baseline estimator. Extracting the deviation signals in the presence of noise is absolutely necessary for good sensitivity. Several methods of baseline estimation were investigated. Finally, a proprietary, patent pending, non-classical method was derived and found to be effective.

In the system as it is designed, signals from 5 separate orthogonal measurements of water quality (pH, Conductivity, Turbidity, Chlorine Residual, TOC) are processed from a 5-parameter measure into a single scalar trigger signal in an event monitor computer system that contains the algorithms. The signal then goes through the crucial proprietary baseline estimator. A deviation of the signal from the established baseline is derived. Then a gain matrix is applied that weights the various parameters based on experimental data for a wide variety of possible threat agents. The magnitude of the deviation signal is then compared to a preset threshold level. If the signal exceeds the threshold, the trigger is activated.

The deviation vector that is derived from the trigger algorithm is then used for further classification of the cause of the trigger. The direction of the deviation vector relates to the agents characteristics. Seeing that this is the case, laboratory agent data can be used to build a threat agent library of deviation vectors. A deviation vector from the monitor can be compared to agent vectors in the threat agent library to see if there is a match within a tolerance. This system can be used to classify what caused the trigger event. This system can also be very useful in developing a heuristic system for classifying normal operational events that may be significant enough in magnitude to activate the trigger. When such an event occurs the profile for the vector causing it is stored in a plant library that is named and categorized by the system operator. When the event trigger is set off the library search begins.

The agent library is given priority and is searched first. If a match is made, the agent is classified. If no match is found, the plant library is then searched and, the event is identified if it matches one of the vectors in the plant library. If no match is found, the event is classified as an unknown and can be named if an investigation determines its cause. This is very significant because no profile for a given event need be present in the libraries for the system to trigger. This gives the system the unique ability to trigger on unknown threats. Also, the existence of the plant library with its heuristic ability to learn plant events results in a substantial and rapid decrease in unknown alarms over time. The developed system has been subjected to strenuous testing in both laboratory and field

scenarios as detailed in the remainder of this paper and has been found to be an effective tool for surveillance of the distribution system.

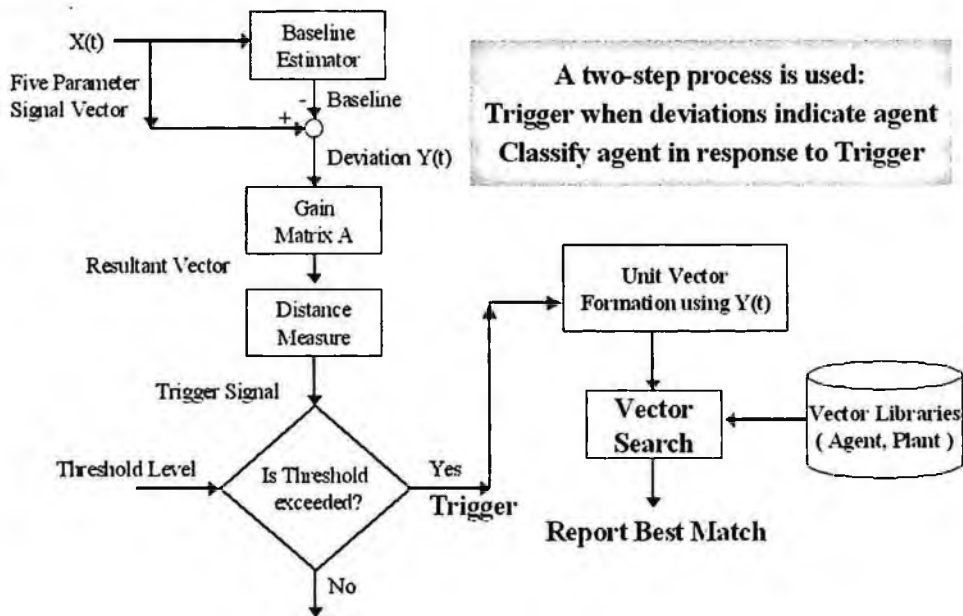


Fig 2. The use of intelligent algorithms with standard bulk parameter monitoring equipment allows for a robust system that is capable of triggering on and classifying a wide diversity of threat agents including unknown events.

4. Battelle EPA ETV Verification Testing

In the fall of 2004, the developed technology was submitted for testing to the EPA Environmental Technology Verification (ETV) testing program run by Battelle. The ETV Program develops testing protocols and verifies the performance of innovative technologies that have the potential to improve protection of human health and the environment. ETV was created to accelerate the entrance of new environmental technologies into the domestic and international marketplace. The category being tested was Multi-parameter Water Monitors for Distribution. The technologies verified needed to function on line and be capable of monitoring chlorine and at least one other parameter. The tested technologies were evaluated for accuracy of instrument readings versus reference methods, ability to maintain integrity during long term deployment, system to system variability, ability to respond to contaminants, and classification of unknown contaminants (The Hach system was the only one evaluated for ability to classify contaminants).

4.1 Accuracy⁶

The on-line instruments were evaluated on an individual basis versus standard laboratory methods for the given parameters.

Table 1. Accuracy

Evaluation Parameter		Cl	Turb	Cond.	pH	TOC
Stage I Accuracy	Units 1 & 2, range of %D (median)	-47.4 to 4.5 (-3.9)	-53.9 to -1.3 (-34.1)	-15.5 to 8.1 (2.2)	-6.6 to 3.1 (0.9)	-64.7 to 147.5 (-14.8)

Overall accuracy of all of the parameters measured was very good and correlation to laboratory method was strong. Small absolute values in turbidity and TOC lead to large %D. Due to sampling errors in some cases on-line instrument readings may be more reliable than the reference method e.g. chlorine.

4.2 Inter-unit Reproducibility⁶

Two separate but identical sets of instrumentation were deployed for the duration of the study. Measurements versus reference interments were compared through out the course of the study.

Table 2. Reproducibility

Parameter		Cl	Turb.	Cond.	pH	TOC
Inter-unit Reproducibility	Slope (intercept)	0.98 (0.03)	0.97 (0.005)	0.92 (4.19)	1.06 (-0.40)	0.97 (0.31)
	r ²	0.994	0.881	0.961	0.919	0.991
	p-value	0.779	0.884	0.006	0.517	0.374

With the exception of conductivity both units generated similar results.

4.3 Contaminant Injection⁶

During this phase of the test the instruments were installed in a recirculating pipe loop. Various contaminants were injected to determine if they altered the baseline response pattern of the instruments. Similar injections were performed after the long-term deployment test to ascertain if there had been any degradation of instrument response.

Total chlorine and TOC were dramatically affected by injections of nicotine, *E. coli*, and Aldicarb; and turbidity, pH, and conductivity were affected by some or all of the injections, but not as consistently as total chlorine and TOC. Aldicarb altered pH during testing after extended deployment but not before. Agreement in both cases with reference readings indicates that the instruments were functioning properly and, the difference was in the injection preparation.

Table 3. Response to Contaminant Injection

Parameter		Cl	Turb	Cond	pH	TOC	
Initial response to injected contaminants	Nicotine	Ref	-	(a)	NC	NC	+
		Hach	-	+	NC	NC	+
	Arsenic Trioxide	Ref	-	(a)	+	+	NC
		Hach	-	+	+	+	NC
	Aldicarb	Ref	-	(a)	NC	NC	+
		Hach	-	+	NC	NC	+
Response after extended deployment	E. coli	Ref	-	+(b)	+	-	+
		Hach	-	+	+	-	+
	Aldicarb	Ref	-	+(b)	NC	-	+
		Hach	-	+	NC	-	+

(a) Relatively large uncertainty in the reference measurements mad it difficult to detect a significant change

(b) Magnitude of change different between duplicate injections

+/- Parameter measurement increased/decreased upon injection

NC No change in response to contaminant injection

4.4 Long-term Deployment⁶

During this phase of the testing, the systems were operated continuously for 52 days with only normal maintenance, such as reagent replenishment, being performed. During the course of the test, instruments were regularly compared to reference instruments. See Table 4. At the end of the 52 days a second response to contaminant injection procedure was performed. See table 3. The results of the extended deployment study indicate that the system can be effectively deployed for long periods of time with only routine maintenance. Relative large %D in turbidity and TOC measurements are artifacts of the low total values for these parameters encountered during the testing procedure and do not indicate a problem with these sensors.

Table 4. Post-Extended Deployment Results

Parameter	Unit 1			Unit 2	
	Reference Average (SD)	Average (SD)	%D	Average (SD)	%D
Free chlorine	1.03 (0.03)	0.98 (0.02)	-4.9	0.98 (0.02)	-4.9
Turbidity	0.17 (0.02)	0.16 (0.03)	-5.9	0.15 (0.04)	-11.8
Temperature	22.66 (0.16)	22.61 (0.03)	-0.2	23.70 (0.06)	4.6
Conductivity	356 (1)	380 (1)	6.7	357 (1)	0.3
pH	8.59 (0.01)	8.40 (0.01)	-2.2	8.61 (0.00)	0.2
TOC	0.88 (0.01)	0.70 (0.01)	-20.5	0.91 (0.01)	3.4

4.5 Inter-Unit Reproduceability⁶

Two Hach units were compared, using data collected from reference samples throughout the verification test to determine whether similar results were generated. See table 5.

Table 5. Inter-Unit Reproducibility Evaluation.

Parameter	Slope	Intercept	r ²	t-test p-value
Free chlorine	0.98	0.03	0.994	0.779
Turbidity (outlier removed)	0.97	0.005	0.881	0.884
Temperature	0.72	7.68	0.758	5.5 × 10 ⁻⁶
Conductivity	0.92	4.19	0.961	0.006
pH	1.06	-0.40	0.919	0.517
TOC	0.97	0.31	0.991	0.374

Shading indicates a significant difference between the two units.

For free chlorine, pH, TOC, and turbidity, the linear regression had coefficients of determination greater than 0.91 and slopes within 6% of unity, indicating similar and repeatable results. The t-test p-values confirmed the sensors were generating statistically similar results. The conductivity meters had a linear regression coefficient of determination of 0.961 and a slope of 0.92, indicating that the data were highly correlated with one another. The t-test generated p-values significantly less than 0.05, which indicated that the results from the two conductivity sensors were significantly different. This difference was driven by the small amount of variability in the conductivity measurements; therefore, the small difference between the means of the two units was statistically significant. It is important to note that the offsets in the measured parameters do not affect the performance of the algorithm because the baseline is removed and the classification is performed based only on deviations from baseline.

4.6 Contaminant Classification⁶

During the final stage of the verification test thirteen contaminants (See Table 6) were injected at a concentration of approximately 15 mg/L, in duplicate, into a 1500 foot straight line pipe and allowed to flow past the monitoring sensors. Every contaminant injection resulted in the system exceeding the trigger threshold and producing a corresponding agent alarm. Each minute-by-minute search of the agent library can result in more than one agent being identified. For both Hach Units, the agent alarms occurred as few as eight times and as many as 79 times during the 20-minute injection periods. No agent alarms occurred outside of the 20-minute injection periods. If the system recognized deviations from the baseline, the agent library identified and recorded these deviations as “unknown” event. Due to the dynamic nature of the leading and trailing edges of the injected contaminant, it is possible that an injection event generated alarms other than the known injected contaminant. Table 6 shows all contaminant injections classified according to the fraction of agent alarms attributable to the correctly classified

injected contaminant. The data is depicted concisely by classification rates divided into five levels: Level 5 – greater than 70%, Level 4 – between 31% and 69%, Level 3 – between 1% and 30%, Level 2 – injected contaminant not identified but other contaminants were identified, and Level 1 – no injections detected.

Table 6. Contaminant Classification Results

Contaminant	Injection 1		Injection 2	
	Unit 1	Unit 2	Unit 1	Unit 2
Aldicarb	4	4	4	4
Arsenic trioxide	2	2	2	2
Colchicine	4	4	4	4
Dicamba	4	5	5	5
Dichlorvos	4	3	3	2
<i>E. coli</i>	3	2	4	2
Ferricyanide	5	5	5	5
Fluoroacetate	5	5	4	4
Glyphosate	4	3	2	2
Lead nitrate	5	5	5	5
Mercuric chloride	4	4	4	4
Methanol	4	4	4	3
Nicotine	2	2	2	2

From the tests conducted on Hach Version 1 System, weak results were obtained for Methanol and Dichlorvos, while poor results were obtained for Glyphosate, Nicotine, Arsenic Trioxide and *E. coli*. The data from the tests on VERSION 1 of the HHST technology at the EPA center were recorded at the time of the tests. The Event Monitor Trigger System also acts as a data logger and provides a copy of the sensor signals recorded during the tests. This situation afforded us the ability to analyze failures detected in the VERSION 1 tests, improve and upgrade software, and then replay new versions of the technology to test for efficacy. A variety of causes were found to affect the test results of VERSION 1.

Because of a misunderstanding, the original Version 1 threat agent library included Round-UP Herbicide (a form of glyphosate), while the ETV protocol used pure glyphosate. When pure Glyphosate was added to the agent library in Version 3, the system correctly classified the agent.

HHST had previously tested nicotine (in house, and at ECBC) with good results. However, the data from the ETV test facility revealed that the excessive mixing method employed prior to injection had caused the nicotine base to react with the carbon dioxide in the air, changing the chemical nature of the contaminant. The Agent Library was

improved by adding a signature for reacted nicotine, and Version 3 shows the positive test results. These two examples demonstrate the sensitivity of the system, and how the comprehensive data structure of its Agent Library derives its classification accuracy.

In addition while first developing the Agent Library, HHST employed bench-scale chlorine analyzers that contained EDTA (a metal sequestering agent) as a reagent component, whereas the EMTS sensor panel includes chlorine analyzers that do not use this substance. The EPA/ETV tests revealed this flaw, and VERSION 3 includes upgraded library signatures. Signatures for some other agents were examined for tabular errors and those were corrected as needed. This second set of test results could not be included in the ETV report, as any re-testing was not a part of the original test protocols. Following analysis and upgrades, two succeeding algorithm versions were produced; the test results from VERSION 3 is summarized in Table 7.

Table 7. Classification Results from Algorithm version 3

Contaminant	Injection 1		Injection 2	
	Unit 1	Unit 2	Unit 1	Unit 2
Aldicarb	4	4	4	4
Arsenic trioxide	4	5	4	5
Colchicine	4	4	4	4
Dicamba	4	3	5	5
Dichlorvos	2	2	3	3
<i>E. coli</i>	4	3	4	3
Ferricyanide	5	5	5	5
Fluoroacetate	5	5	5	5
Glyphosate	4	4	4	4
Lead nitrate	5	5	5	5
Mercuric chloride	5	5	5	5
Methanol	4	4	4	3
Nicotine	4	4	4	4

5. ECBC Testing

The purpose of this effort was to challenge water distribution systems and sensors, with agent simulants and real threat agents, in order to characterize the response of the distribution system and Early Warning System to agents. Agent concentrations and water solutions were varied to allow for the development and demonstration of distribution methodologies and performance data acquisition. In addition, this work evaluates the effectiveness of Hach Homeland Security Technologies real-time detection technology and provides important information necessary for the U.S. Army to perfect its theories of

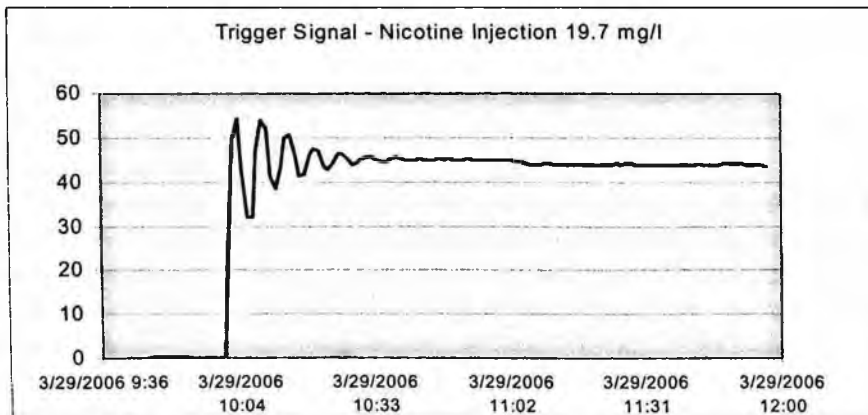
operation and response mechanisms. The scope of the work performed during these tests was two fold. The first part of the test was to perform beaker studies on agents that are not available for use in the Hach Laboratories in Colorado such as VX, Sarin, Soman, Ricin and Anthrax etc. The second part of the testing protocol called for verification of signatures in a flowing loop to validate the transfer of the beaker signature data to real world scenarios.

It is known that in a real attack on a water distribution system the concentrations of the agents would vary throughout the distribution system. The concentrations tested were to either infectious amounts, ID_{50} , for replicating agents, or LD_{50} amounts for chemical agents. Each agent was tested at three dose values, as defined in a test plan matrix. Two types of disinfectant are commonly used in water distribution systems: free chlorine, and monochloramine. It is necessary to test agents in both types of media to have information representative of each type. It is also clear that variable amounts of chlorine would be in the distribution system water, so tests used different solutions: 0.2 ppm Free Available Chlorine (FAC), 1 ppm FAC, and 2 ppm Monochloramine, as these are the limits of typical system concentration. Monochloramine is usually less variable in concentration and can be tested at typical values. Real-time loop tests were run on nicotine, ricin, BA, and methanol.

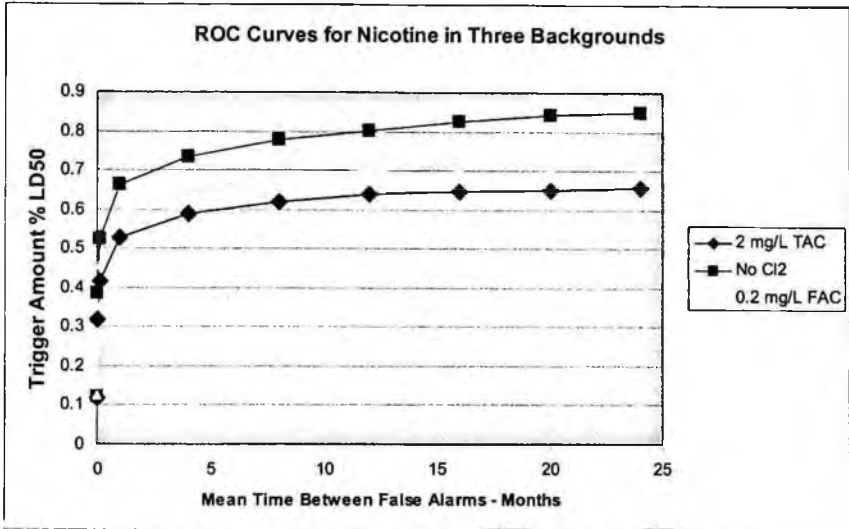
All fingerprints were successfully developed and ROC curves were generated for all agents tested. It was also found that the fingerprints developed from the lab work could be successfully transferred to a flowing system by successfully triggering and classifying agents in the flow loops. Because of security concerns and confidentiality, only selected nicotine test data is provided in graphs 1 and 2.

The test concentration of nicotine was 19.7 mg/L. In a single test run, nicotine was recognized by the system, with detection angles ranging from 0.94 degrees (essentially a perfect match) to 9.84 degrees (a weak match). The fluctuating trigger signal is due to the mixing dynamics of a recirculating loop.

Graph 1 Nicotine Injection in Loop



Graph 2 ROC Curves for Nicotine in Three Backgrounds

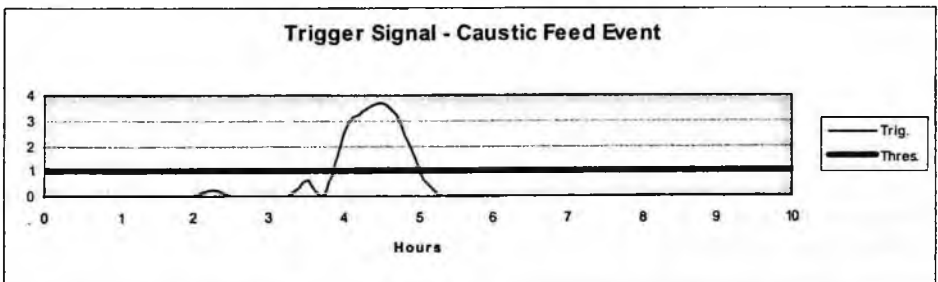


Threshold trigger levels range from 0.36 to 1.26 for all curves

6. Real World Testing

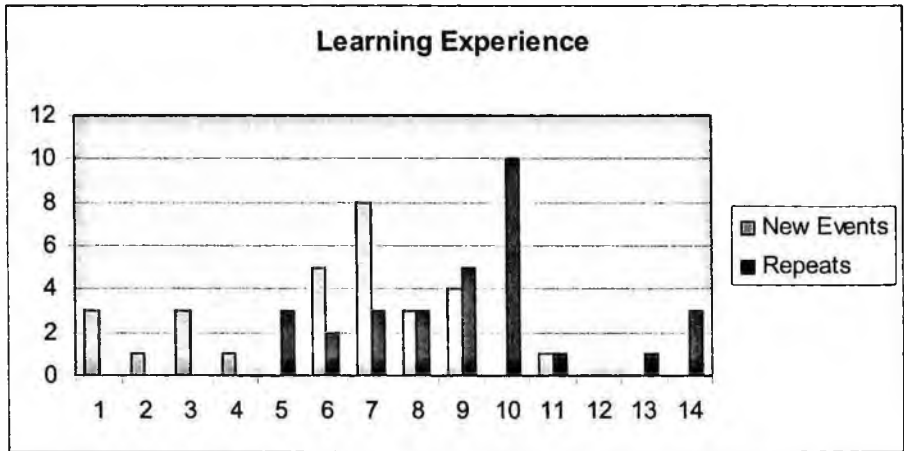
The described system has been deployed in a variety of real world venues across the United States to determine theories of deployment and response and to verify robustness of the trigger and learning ability. To date over 120,000 hours of real world data has been collected. Several actual incidents have been recorded and learned. An example is the caustic over feed event depicted in Graph 3. Graph 4 shows a very noisy real world situation and demonstrates how quickly the system can learn commonly occurring unknown events to help reduce their occurrence.

Graph 3 Caustic Over Feed



In this situation, the water utility plant used caustic feed to control the water pH. Misdelivery of a more concentrated form of the caustic resulted in the feed of excess caustic. The pH and conductivity of the water deviated enough to trigger a Plant Event for the system to “learn” and identify its reoccurrence.

Graph 4. Learning Experience Rate at a Field Site



The data in this case represent a real world deployment situation that had very noisy water quality. In this scenario there were 26 unique trigger events in the first 11 days of operations. All were fingerprinted and learned by the system. 11 of the events were repeated. This demonstrates that common events are rapidly learned by the system resulting in a rapid decrease in unknown alarms.

7. Conclusion

The designed and tested system makes use of an integrated array of robust common water quality monitoring sensors coupled with interpretive algorithms to recognize and classify significant water quality deviations. Extensive in house and 3rd party verification testing as well as extensive deployment at field sites has demonstrated the suites ability to fill the analytical gap that currently exists for distribution network monitoring and serve the purpose of an early warning system in the water distribution network. Hopefully the systems unique ability to learn will result in not only increased safety from terror related events but will morph into an operational tool that will find everyday use in improving water quality operations and ensure a better quality drinking water product to consumers.

References

1. C. Whitman. Environmental News: Whitman allays fears for water security; Possibility of successful contamination is small. EPA Press Release. Thursday, October 18th (2001).
2. D. Kroll, Mass casualties on a budget. Confidential paper. Hach HST (2003).
3. Army Corps of Engineers, Calculations on threat agents, requirements, and logistics for mounting a successful backflow attack. (2003)
4. T. P. Allman, Drinking water distribution system modeling for predicting the impact and detection of intentional contamination, Masters Thesis, Colorado State University, Dept of Engineering, Fort Collins Colorado, Summer (2003).
5. M. Ginsberg, V. Hoch and D. Kroll. Terrorism and the security of water distribution systems. *Sword and Ploughshares*. January. (2005)

REAL-TIME, ON-LINE MONITORING OF DRINKING WATER FOR WATERBORNE PATHOGEN CONTAMINATION WARNING

JOHN A. ADAMS

*Sensor Products Group, JMAR Technologies, Inc., 10905 Technology Place,
San Diego, CA 92127, United States
jadams@jmar.com*

DAVID MCCARTY^a

*Sensor Products Group, JMAR Technologies, Inc., 10905 Technology Place,
San Diego, CA 92127, United States
dmccarty@jmar.com*

The increasing value of safe, purified water in today's global environment presents many challenges for improving water purification processes while accelerating water quality monitoring technologies. These challenges include: escalating worldwide security threats, increasing demands by consumers and regulatory agencies for higher water quality standards, and proliferation of contamination events.

A major concern in protecting water purity is the detection and control of harmful microorganisms in water sources, treatment, and distribution. Due to the length of time for standard laboratory methods to yield results, typically 24-72 hours, current analytical techniques have not been able to reliably detect microorganisms in real time and provide an immediate warning of biological contamination.

Through years of research and development, an emerging technology called MALS (multi-angle light scattering) can now achieve on line, real time detection and classification of microorganisms in water. The method utilizes a laser beam focused into a slip-stream of water and the light scattered by microorganisms generates unique patterns. These "bio-optical signatures" are instantaneously captured by photodetectors, matched against catalogue profiles in an on-board data base, and subsequently classified within minutes of detection. The system can then immediately alert personnel to the contamination event while extracting a water sample for further confirmation.

^a 8264 Woodruff Avenue, Cincinnati, OH 45255

1. Introduction

Today's water utilities face many challenges when it comes to protecting the public health. These challenges include: increasing demands by consumers and regulatory agencies for higher water quality standards; escalating worldwide security threats; and proliferation in the number of contamination events from accidental discharge, treatment malfunction, or operator oversight.

A major concern for all water utilities is the detection and control of pathogenic microorganisms, both known and emerging, in potable water treatment and distribution. Not only are there a number of chlorine resistant pathogens such as *Cryptosporidium* that can contaminate drinking water systems, but there are also potentially harmful microorganisms that can be introduced, either accidentally or intentionally, that will quickly propagate under suitable environmental conditions. Due to the length of time for standard laboratory methods to yield results, typically 24-72 hours, there has not been a reliable system to detect microbial contaminants in real time and provide an immediate warning of pathogen contamination events. Because of these expanding challenges, there has been an accelerated development of rapid tests and real-time methods to address the pressing needs of the water treatment community.

2. Microbiology and Emerging Technologies

Analytical methods in microbiology were developed more than 100 years ago and are almost identical today. These methods are comprised of the following steps: sampling; culturing and isolating the microbes in a suitable growth media by incubation; identifying the organisms through microscopic examination or stains; and quantification. Although still the 'Gold Standard' for microbial analysis, these procedures found in Standard Methods or ASME require ascetic technique in sampling and handling, skilled technicians to perform the analysis, and a number of reagents, materials, and instruments to obtain results in one to three days. Due to the undetectable, adaptable, and insidious nature of microorganisms, intermittent grab samples are often 'hit-or-miss' and do not provide continuous surveillance as compared to a number of chemical on-line instruments.

In order to reduce the amount of time to obtain microbiological results, a number of methods have been developed and can generally be categorized as:

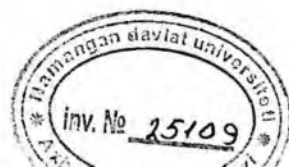
- Accelerated/Automated Tests
- Rapid Tests
- Real-Time Contamination Warning Systems (CWS)

Each test category has different characteristics in terms of the time to obtain results, specificity, sample frequency, accuracy, enumeration, sensitivity, operating complexity, capital and operating costs, and operating environment as illustrated in [Table 1](#).

Table 1. Comparison of rapid microbial methods

PARAMETER	MALS (BioSentry™)	Chemical & Physical Monitors	Rapid PCR	ATP Lumines- cence	Immuno- assays
Sample Rate	on line, continuous	on line, continuous	intermittent sample	grab sample	grab sample
Response Time	2-5 minutes	5-60 minutes	1-3 hours*	1-12 hours	5-30 minutes
Ease of Use	very good	good	average	good	excellent
Automation	excellent	very good	good	no	none
Communication	excellent	very good	no	no	none
Remote Operation	yes	yes	field use - operator	no	field use
Rugged	good	good	yes	no	yes
Broad Spectrum	yes	yes	limited	yes	limited
Qualitative	yes	no	yes	yes	yes
Quantitative	relative concentration	no	no	good	no
Low False +/-	average	no	interferences	average	cross reactions
Specificity	detects and classifies	no	very good	poor	good, but high false +
Reproducibility	average	poor	very good	good	good
Sensitivity	average	poor	poor - average	good	poor
Capital Cost	low-medium	medium	medium	low	low
Operating Cost	low	low-medium	high	medium	low

The remainder of this article will focus on real-time detection and classification of waterborne microorganisms for continuous water surveillance (noted above in [Table 1](#) as MALS).



3. MALS History and Technology

The phenomenon of light scattering by small particles originated with the pioneering work done by John Tyndall¹, Lord Rayleigh², Gustav Mie³, and Chandrasekhara Raman⁴ around the turn of the 20th Century. MIE theory is the model often used to describe light scattering by small particles where the particle size is similar to the wavelength of light used in the scattering measurements. Scattering calculation programs developed by Laven⁵ and Bazhan⁶ are based on MIE theory and are most often used in light scattering instruments⁷.

The theory of using MIE scattering for the detection of microorganisms was first proposed in 1969, but a number of problems existed. In 1985, Phil Wyatt and Greg Quist⁸ used an empirical approach to solve the inverse scattering dilemma and subsequently used this model to identify homogeneous and isotropic refractive particles. The technique, called multi-angle light scattering (MALS), uses simultaneous measurement of various scattered light angles and intensities and was demonstrated using *E. coli* and *Staphylococcus epidermis*⁹. Wyatt and Jackson¹⁰ later used MALS to develop unique 'signatures' for 12 phytoplankton species and this was significant in demonstrating how the technology could be used to classify complex biological particles.

MALS is based on laser technology, particle light scattering, photo-detection, and computer signal processing. The development of a real-time, on-line water monitoring system has evolved through several generations of system design. Due to the fact that computing power, laser development, and optics technology have all greatly improved over the past 5 years, MALS has now become a viable technology for commercial use as a waterborne contaminant instrument. The first public demonstration of the system was at the 2003 Super Bowl as a security device and was commercially introduced by JMAR Technologies as BioSentry™ in 2005.

4. BioSentry System Overview

The BioSentry™ Water Monitoring System is an early warning system (EWS) that is intended to alert the user of potential harmful contamination, much like a smoke alarm alerts the homeowner of a potential household fire. It utilizes the described MALS technology to provide continuous, real-time monitoring and classification of harmful pathogens in water.

4.1. System Operation

Figure 1 shows the basic schematic of a MALS system and illustrates how the laser beam is directed at a particle. The laser generates a scattered-light pattern that is registered onto a detector. The light pattern is much like a ‘fingerprint’ since it is unique to the internal and surface features of the particle (including size, shape, morphology, and material composition). These ‘fingerprints’, termed Bio-Optical Signatures (BOS), are used to differentiate particles into several probable classifications.

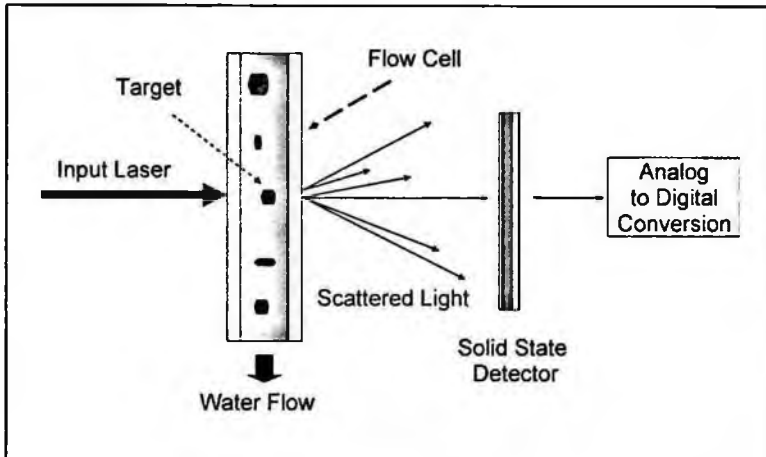


Figure 1. Basic MALS system

Differentiating microbes from non-biological particles in water was a challenge for bacteria ranging in size from 500 nanometers to two microns, depending on the size and water particle background, until the two-dimensional, multi-channel array was introduced into the system.

* This system, BioSentry v1.75, is scheduled for release in August, 2006.

Figure 2 shows scattering patterns achieved using a 2-D, multi-channel detector array. The MALS scattering patterns clearly show that different types of particles have unique scattered-light ‘fingerprints’. Calibrated, micron-sized plastic spheres show symmetrical scattering patterns, while an inorganic ‘dirt’ particle scatters light in a very random pattern and pathogens such as *Cryptosporidium*, *Giardia*, and *E.coli* have distinctive patterns that differ greatly from ‘dirt’. This phenomenon demonstrates that Bio-Optical

Signatures can be generated and installed in a database that can then be compared using pattern recognition and matching algorithm software for detection and classification in a real-time operating function.

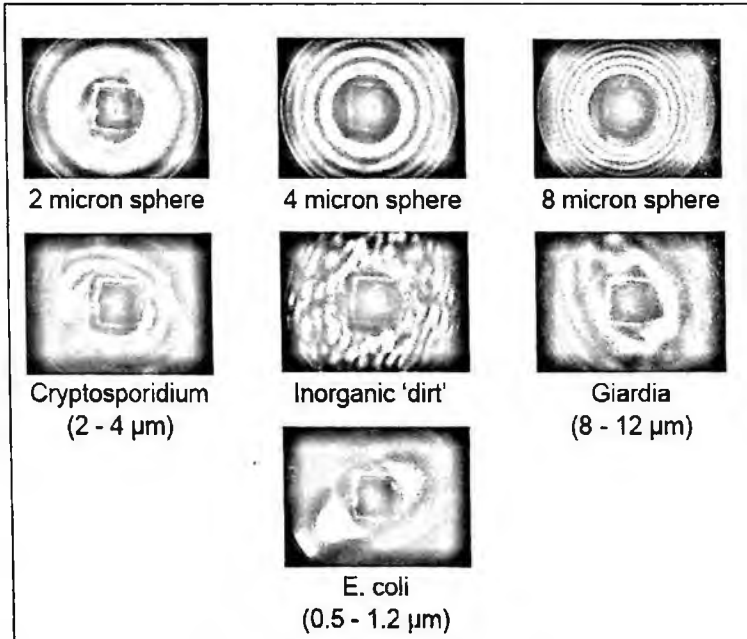


Figure 2. Particle Scattering Patterns

5. Relationship Between Time, Detection Limits, and Interferences

Water is monitored continuously as it passes through the flow cell. The system's detection time and performance are characterized by:

1. **Microbial Concentration:** The probability of 'seeing' and 'classifying' a particle is proportional to the concentration level since the higher the concentration, the more likely it will pass through the surveillance volume within a given time.
2. **Water Quality:** The higher the 'clarity' of the water, the fewer the interferences which will result in better detection, classification, and lower (improved) levels of detection (LOD).
3. **Particle Size:** The larger the particle, the more distinct the scattering pattern will be to provide a better 'match' to a Bio-Optical Signature in the classification library.

4. Bio-Optical Signature Database versus 'Unknown': The more BOS classifications and configurations in the library to match the scattering patterns, the fewer the unknowns.
5. False Positive: False positives are particles that are mis-classified as an organism in the detection library. False positives are a result of:
 - a. Water clarity - the higher the clarity, the lower the false positives
 - b. Flow cell cleanliness - the cleaner the cell, the less light distortion
 - c. A False Positive is not the same as a false alarm that can be managed by adjusting the Warning and Alert levels on the system.
6. Completeness of the BOS: the BOS is formed from processing thousands of the same microorganisms through the system and recording patterns that represent various sizes, shapes and health of the organisms. However, it is not possible to characterize every form that a microorganism can take with current technology.
7. Limit of Detection: A measurement of the system's ability to detect microorganisms at specified concentration levels over specified periods of time. The lower the number, the better the sensitivity.

6. First Generation System Development and Performance

The BioSentry's performance was first tested on treated drinking water. Inactivated pathogens were used to establish the Bio-Optical Signatures. Calibration and testing were often performed by using standardized, inert polystyrene latex (PSL) spheres to mimic live microorganisms in potable water. In lab testing, using 700 nanometer PSL spheres, a limit of detection (LOD) sensitivity of eight particles per 100 ml was achieved in four hours, which translates to one particle per 100 milliliters in 24 hours of continuous monitoring. Field test results at a municipal water facility indicated that the sensitivity was slightly less, with an LOD of 16 bacteria per 100 ml in four hours.

Water quality greatly affects the limit of detection and number of false positives. Since there is less interference in high purity water compared to potable quality water, the LOD's are lower, the probability of detecting and classifying microorganisms is much higher, and the false positive rate is minimized. As expected, experiments performed in the laboratory show that performance is significantly better. Using filtered tap water (0.2 micron filter), spike tests performed using *P. aeruginosa* yielded detection results of <1 bacteria per 100 ml in four hours, and 22 bacteria per 100 ml in five minutes. Results can vary depending on the type of microorganism and the different particle backgrounds contained in each specific water sample. It should also be noted that potable water quality (i.e. interferences) can vary depending on geographical location and the type of treatment

process used. Spike test results for *Cryptosporidium* at two different water treatment facilities are shown in Table 2.

Table 2. Spike Test Results for Drinking Water - First Generation BioSentry

WATER DISTRIBUTION SPIKE TESTS - <i>Cryptosporidium</i>				
SPIKE CONCENTRATION	FALSE POSITIVES	RESPONSE TIME - LIMIT OF DETECTION		
		LOD 1 min (#/ml)	LOD 5 min (#/ml)	LOD 4 hrs (#/ml)
# / ml	%			
0	1.98			
2 micron spheres @ 2000/ml	0	4.66	0.93	0.02
2000	0	12.92	2.58	0.05
2000	0	61.59	32.64	13.71
0	3.00			
2 micron spheres @ 4000/ml	0	77.45	40.35	16.78
500	0	268.16	149.5	64.57
500	0	500.4	280.6	121.2
500	0	342.6	189.9	81.78
500	0	800.7	451.5	196.2
150	0	459.6	258.8	112.35

7. BioSentry™ in Operation

7.1. Factory Calibration

In calibrating a system for an application, JMAR first characterizes the water from a specific site and optimizes the system to reduce background interference. The system is then programmed with a customized list of microorganism BioOptical Signatures using the appropriate algorithms. This procedure alleviates 'normal' water particulates from impairing the System's classification capability.

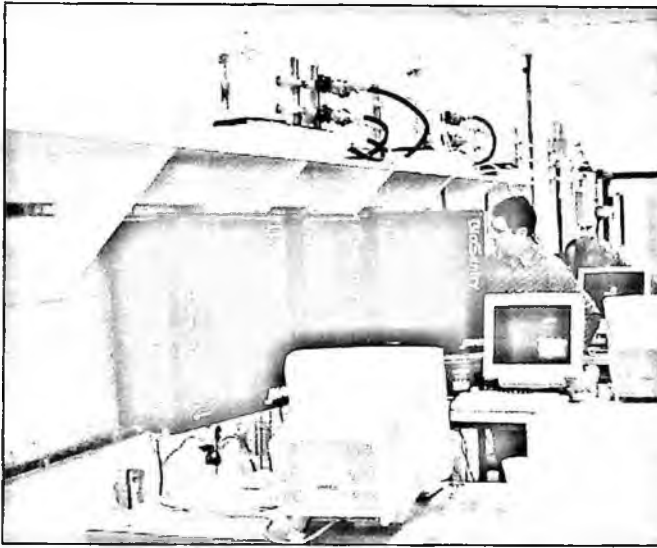


Figure 3. Production unit calibration at the JMAR BSL2 lab.

7.2. System Installation and Operation

The BioSentry is easily installed by mounting the system to a wall or free-standing frame, tapping into to the water line, connecting to an electrical supply, and setting up remote communication services. After calibrating the system for the specific site location and water, BioSentry continuously (24/7) monitors the water line. As particles pass through the system, the instrument detects the particles and compares the scattering patterns to bio-optical profiles in the onboard computer library. 'Matches' are classified, while unclassified biological and inorganic contaminants are categorized as 'Unknowns'. If microbial contaminants exceed the customer's pre-set thresholds for warning and alert, the system automatically provides an alarm (communicated remotely) and can, optionally, extract a sample for further analysis and confirmation.

As shown below in [Figure 4](#), a graphic user interface (GUI) screen depicts the continuous water surveillance and illustrates a biological contamination event. BioSentry includes records of the water monitoring over time and also includes maintenance and problem logs for reporting.

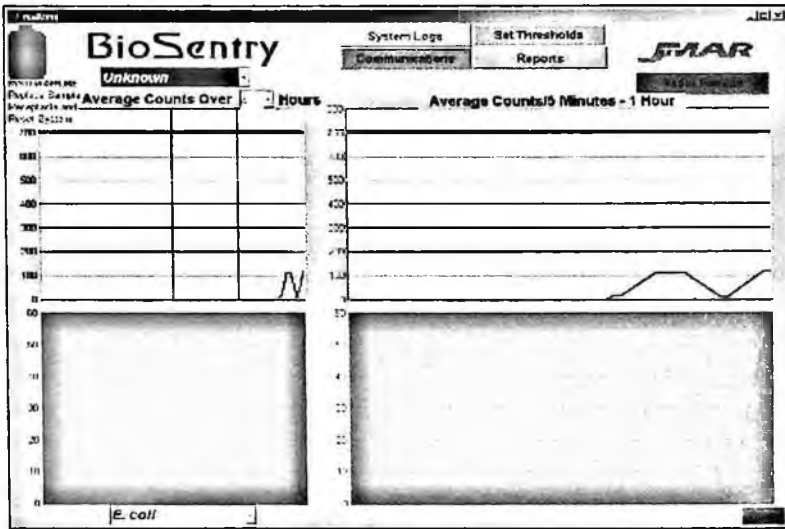


Figure 4. Graphic User Interface screen shows surveillance in real-time. The alert threshold for *E. coli* is shown by the red alert screen with sample bottle icon in the upper left illustrating that a sample was extracted for confirmation.

7.3. Alert Response

Since BioSentry is a real-time contamination warning system designed to “warn” of a potential biological threat, it does not provide the accuracy and sensitivity of some rapid tests or laboratory methods requiring a grab sample for analysis. Therefore, validation is needed to confirm the event, which can be done using an automatically-generated sample taken at the time a warning is given.

Recommended action in response to an Alert:

1. Check Water Quality Changes or Process Variance: Examine other water quality parameters such as pH, conductivity, chlorine concentration, and turbidity to confirm that the water quality has changed. An intrusion alarm or an unusual variance in the water system operation should also be investigated.
2. Rapid Test Confirmation: From the extracted sample, a number of rapid confirmation methods or test kits are available to confirm if the contamination is a biological agent (ATP). Even more specific tests such as immunoassays can be utilized to corroborate the BioSentry probable classification. PCR is another possible confirmation test with high specificity¹¹.

3. **Response:** A contingent plan should be developed while the confirmation is taking place. This should include the release of precautionary notices and readiness to implement an emergency response plan. Should a biological threat be confirmed, a procedure (planned and practiced) should be implemented for remedial action. This would include authority and public notification, containment, and/or diversion of the water, and medical alert and response.

7.4. Emergency Planning and Response

It is important to have an emergency response plan devised and implemented should an intentional or accidental contamination event occur. A summary of the plan should include defined actions to respond to the following:

1. **Warning and Stand-by:** In the event of a warning, personnel should be aware of a potential water contamination event and begin precautionary steps.
2. **Alert:** Upon alert, BioSentry will notify personnel of the probability that a biological contamination event has occurred. It is possible that a change in water quality, especially turbidity or suspected solids, has entered the system.
3. **Containment:** Water diversion or filtration should be actuated for a temporary containment of the water to protect consumers.
4. **Validation / Investigation:**
 - d. If other water quality devices are installed, these should be checked for parameter changes outside of the normal range.
 - e. The BioSentry unit should be checked for conditions leading up to the alarm.
 - f. A rapid biological method, such as ATP or Immunoassay, can be used to determine if a biological contaminant exists.
 - g. A chemical water quality test kit can be used to determine changes in water quality if on line chemical or particle sensors are not in use.
5. **Findings:** If the findings indicate a biological or water quality change, an emergency action plan should go into effect.
6. **Action:** This can include notification of safety and emergency personnel, disruption of water service or diversion to a secondary water source, implementing plans for a confirming test at a laboratory, and flushing and sanitizing the water system.

8. Applications

From source water to point-of-use, potable water production and distribution has many possible points for contamination. These can be from intentional or accidental incidents, treatment process errors, or post treatment contamination. There are numerous examples of major outbreaks of *Cryptosporidium* in potable water systems worldwide, as well as operator error causing pathogens such as *E. coli* to enter the water distribution system, often infecting thousands of people. It is also known that some terrorist groups have considered introducing microbial pathogens into potable water systems to cause outbreaks and panic. Continuous, real-time surveillance has the potential for preventing widespread damage. Suggested applications are shown in [Table 3](#).

Table 3. BioSentry Early Warning System Applications

Category	Applications	BioSentry Placement	BioSentry Advantage
Water Quality Assurance	<ul style="list-style-type: none"> • Protection from natural contamination • Treatment system malfunction • Membrane or filter breakdown 	Post Treatment & Distribution	Continuous monitoring, real-time detection & notification
Security	<ul style="list-style-type: none"> • Protection from deliberate contamination at distribution 	Distribution	Continuous monitoring, real-time detection & notification
Cost Reduction	<ul style="list-style-type: none"> • Possible reduction in treatment chemicals • Higher lab analysis efficiency - BioSentry can extract sample upon contamination alert 	Finished Water Prior to Distribution	Continuous monitoring, real-time detection /classification & low FP with High LOD Auto sample collection when Alert occurs

9. Security Protection Case History

In early 2006, the City of Anaheim and JMAR Technologies entered into a two-month pilot project to test the effectiveness of an on-line, water quality, Early Warning System (BioSentry™) as part of a Homeland Security initiative.

The project required JMAR to install the system in the Public Utilities Department of the City of Anaheim to monitor the water for the presence of *Cryptosporidium* and other unknown pathogens. *Cryptosporidium* was used as it is a harmful pathogen that has been known to cause fatalities at relatively low concentration levels. *Cryptosporidium* is one of the major pathogens identified by the EPA / Department of Homeland Security¹¹ as a target contaminant that could be deliberately introduced into a water supply by terrorists, thereby causing severe facility disruption and possible fatalities.

The parameters used to assess performance were:

1. Limits of Detection statistical performance – a measurement of the system’s ability to detect microorganisms at specified concentration levels. The lower the number, the better the sensitivity.
2. False Positive Rate statistics report those Events that were identified incorrectly – e.g. a piece of dirt was mistaken for *Cryptosporidium*. It should be noted that false positives will only initiate a false alarm if the threshold is exceeded.
3. Identification Rate (ID) statistical consistency – this measures the percentage of actual versus theoretical microorganisms that were correctly identified during the monitoring process. Those detected but not identified did not create a scattering pattern recognizable by the on-board computer database.
4. Maintenance Performed
5. System Operational Performance (Uptime)

The results from this test project exceeded expectations. It was found that the LOD will actually improve (achieve better results at lower microorganism concentrations) over a longer period of time since the BioSentry™ is a continuous-monitoring device. Figure 5 below shows the Limits of Detection for the BioSentry™ in this application. The LOD was measured at ~243 organisms/mL, well below the standard EPA figure of 10,000 organisms/mL. As shown in Figure 6, the BioSentry also exceeded performance parameters for False Positive Rates, with a mean of 3.25% for this project. The BioSentry’s normal operating parameter for False Positive Rate Performance is less than ten percent.

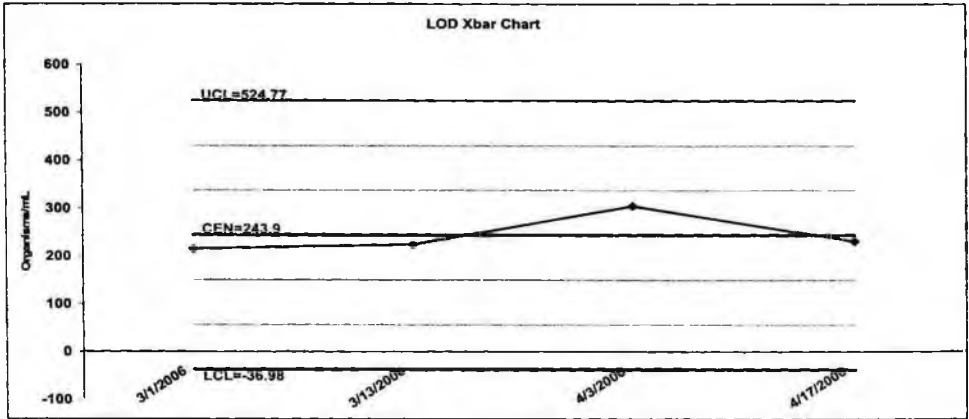


Figure 5. Limit of Detection Performance

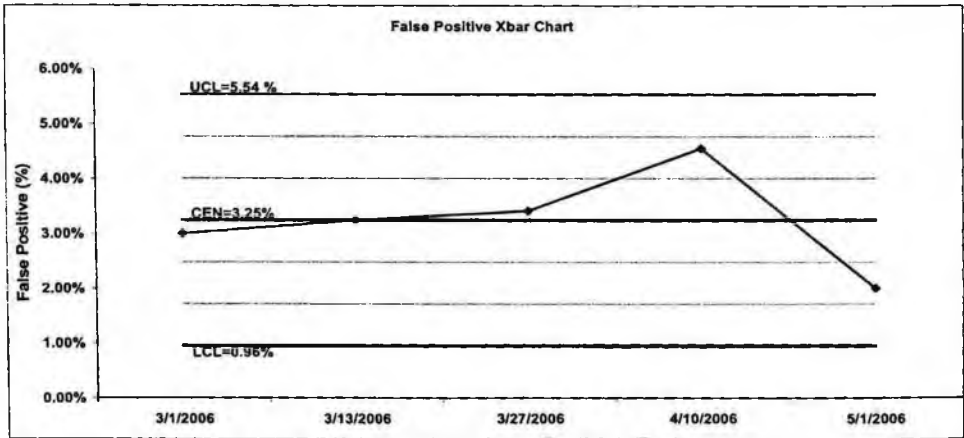


Figure 6. False positive %

The BioSentry™ System performed successfully and proved that it would have notified the City of an Event if *Cryptosporidium* had been introduced into the water supply at average LOD concentration levels above 243 organisms/ml. It further demonstrated that, with frequent calibration, the System could achieve performance levels below 100 organism/ml.

10. Next Generation BioSentry

System improvements in optics, photo-detection, and improved algorithm technologies are expected to provide an even faster response at lower detection limits and offer higher specificity. Figures 7 and 8 show the comparison between the first generation system (BioSentry v1.5) and the predicted performance of the second generation system (BioSentry v1.75) for *Cryptosporidium* and *Bacillus subtilis* in potable water. As the detection library expands and the limitations of interferences are minimized, the BioSentry will provide increasingly higher classification accuracy and lower limits of detection.

With current system improvements in development, a device that automatically detects, classifies with high probability, and extracts a sample which is diverted directly to an automatic rapid test verification system, all without operator intervention, will soon be commercially available. Combined with a suite of other chemical, toxin, and radiological real-time sensors and an integrated data analysis and alert response system, the ultimate water security system is quickly becoming a reality.

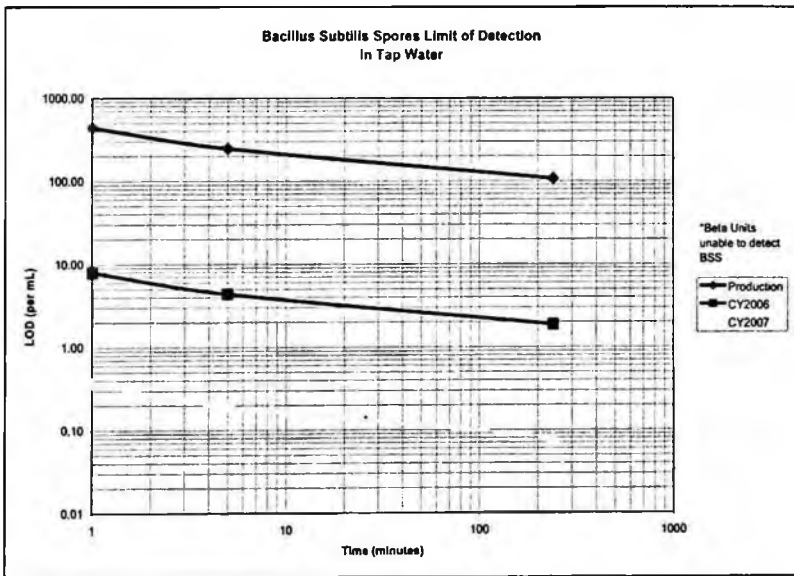


Figure 7. BioSentry 1.5

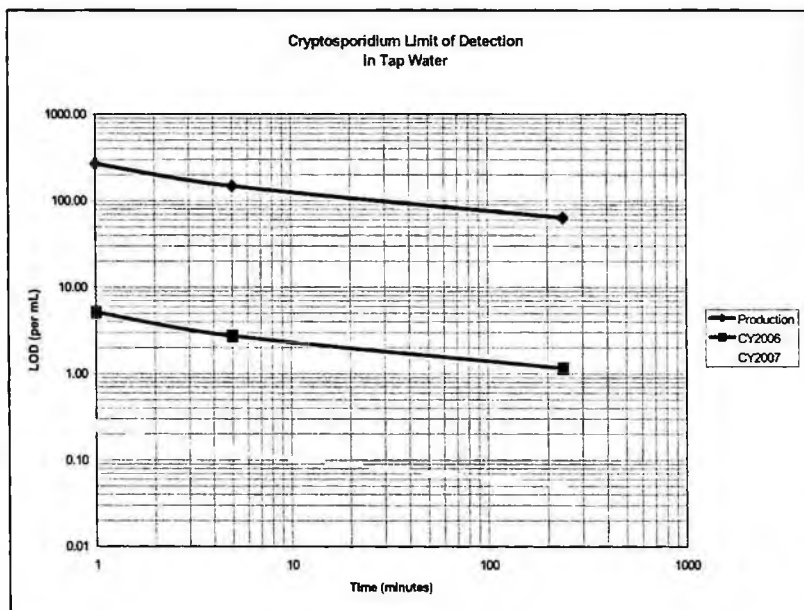


Figure 8. BioSentry 1.75

11. Summary and Current Status

Detecting and classifying bacteria continuously, on-line and in real time for contamination warning and security in the municipal drinking water market offers a valuable defense against widespread illness. BioSentry is a pathogen alert system that is both synergistic and complementary to other water quality sensors and enhances current lab quality control methods. By implementing BioSentry into a distribution water monitoring program, an integrated approach to detecting a biological contamination event can be realized.

Released for production in November 2005, BioSentry™ is currently used in water utilities, beverage plants, cruise ships, and Homeland Security applications.

REFERENCES

1. Tyndall, John. [Phil. Mag. 37, 384; 38, 156, 1869]
2. Rayleigh, Lord. [Phil. Mag. 47, 375, 1899]
3. Mie, Gustav. "Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen. *Annalen der Physik*", *Vierte Folge*, Band 25, 1908, No. 3, S. 377- 445.
4. Raman, Chandrasekhara. *Observed weak inelastic scattering of light from liquids, an effect arising from the scattering of light by vibrating molecules* [Indian J. Phys. 2, p387, 1928]
5. <http://www.philiplaven.com/index1.html>
6. <http://www.scatlab.com/index.html>
7. Bohren, Craig F., Huffman, Donald R., *Absorption and Scattering of Light by Small Particles*, John Wiley & Sons Inc., New York, 1983.
8. Quist, G.M., Wyatt, P.J., "Empirical Solution to the Inverse-scattering Problem by the Optical Strip Map Technique", *JOSA A*, Vol. , Issue 11, pp. 1979- (November 1985) and *JOSA A*, Vol. 3, Issue 5, pp. 671- (May 1986)
9. Wyatt, P. J., "Differential light scattering: a physical method for identifying living bacterial cells", *Applied Optics*, 7:1879-1896
10. Wyatt, P. J. C. Jackson, "Discrimination of phytoplankton via light scattering properties", *Limnology and Oceanography*, 34(1), 96-112 (1989)
11. U.S. Environmental Protection Agency, Office of Water, Office of Science and Technology Health and Ecological Criteria Division; *Technologies and Techniques for Early Warning Systems to Monitor and Evaluate Drinking Water Quality: A State-of-the-Art Review*, December 2005.

CHIP-SIZE WAVELENGTH DETECTORS

OLIVER SCHMIDT, PETER KIESEL, MICHAEL BASSLER, and NOBLE JOHNSON

*Palo Alto Research Center Inc., 3333 Coyote Hill Road,
Palo Alto, CA 94304, USA
michael.bassler@parc.com*

Chip-size wavelength detectors are described that can resolve the spectrum of the incident light with high accuracy over a broad spectral range. The devices can be configured either as compact spectrometers or precise wavelength shift detectors. Applications that are anticipated to benefit from the spectrometers include reagentless optical identification of analytes in fluidic and aerosol samples. The wavelength shift detector is widely applicable as read-out instrument for optical sensors in which a stimulus (e.g., temperature, strain, PH-value, etc.) results in a wavelength shift of an optical output signal, examples include the interrogation system for Fiber Bragg Grating sensors, photonic crystal sensors, Fabry-Perot type sensors, and sensors in which an analyte influences a laser cavity and thereby its emission wavelength.

Keywords: Spectrometer, Wavelength-Shift Detector, Optical Sensor.

1. Introduction

Wavelength detection and monitoring are important in many applications such as optical spectroscopy, optical sensing and wavelength demultiplexing in telecommunications. In recent years there is an increasing demand for compact, cheap and yet precise spectrometers that can replace bulky and expensive bench-top instruments. Two particular areas of interest will be addressed in this paper: (1) monitoring the spectral position of a monochromatic light source for optical sensing applications and (2) integration of spectrometers into lab-on-a-chip devices.

1.1. *Application of chip-size wavelength-shift detectors in optical sensing*

Many new applications of optical sensors have emerged. The automotive sector has created new demands for robust and advanced optical sensing techniques as the amount of optical information processing expands in modern cars.¹ In addition, optical sensors are widely used in structural health monitoring of civil structures² and aircrafts³ as well as in biological, chemical and environmental sensing applications⁴. In comparison to their electronic counterparts, optical sensors offer many distinct advantages. They are very sensitive, allow for remote and distributed sensing, can be used in harsh environments, and are immune to electromagnetic interference.

One prominent example of an optical sensor is the Fiber Bragg Grating (FBG) sensor.^{5,6} The reflection spectrum of a FBG sensor changes in response to an appropriate external stimulus. Similarly, changes in the reflection properties of one and two-dimensional photonic crystals can be used for optical sensing applications. In recent years several applications in biosensing have emerged which are based on a spectral shift of the transmission and reflection properties of a surface or object. Surface Plasmon Resonance (SPR) sensors are widely used for screening biochemical interactions⁷, whereas other groups have developed optical biosensors based on Fabry-Perot cavities in porous silicon⁸ or guided-mode resonance reflectance filters⁹. Another application is to use the shift of optical resonances (whispering gallery modes) excited in silicon spheres to quantify DNA¹⁰.

We have developed a low-cost and compact wavelength-shift detector suitable for the read-out of all kinds of optical sensors that are based on a wavelength shift. Key features of this technology include compactness (chip-size), no mechanical parts, customized spectral resolution, fast read-out and monolithic integration.

1.2. Application of chip-size spectrometers in lab-on-a-chip devices

There is an urgent need for fast, sensitive, and specific characterization and detection methods in point-of-care testing. Examples include health care service, environmental monitoring, bio-agent detection, fresh and waste water analysis, industrial process control, and gas analysis. It is generally a necessity to detect the analyte of interest “on-the-fly”, that is, as it is moving in order to allow for continuous and real-time detection. Optical detection methods offer high sensitivity, but current approaches capture only a “snapshot” of the moving particle which yields limited information. Furthermore, most existing detection schemes are bulky and expensive.

Chip-size wavelength detectors are composed of readily available components and leverage existing technologies from consumer products. The devices can be customized for high spectral resolution or broad spectral range.

2. Concept of Chip-size Wavelength Detectors

Our chip-size wavelength detectors are composed of a detector array which is coated with a linear variable band-pass filter (LVF). The filter converts the spectral information into a spatially dependent signal which is analyzed by the detector array. Thus, the output signal of a coated detector array yields the wavelength spectrum of the incident light.¹¹ One can select from a large variety of light sensing elements and design coatings covering a broad spectral range from the deep-UV to the far-IR. Systems demanding either very high wavelength resolution (e.g., for absorption or Raman spectroscopy) or a wide spectral range (e.g., fluorescence spectroscopy) can be accommodated by choosing the appropriate coating. The size of the spectrometer is only slightly larger than that of a detector array itself, and fabrication requires only conventional, readily available processing technology, which should enable a cost effective manufacturing process.

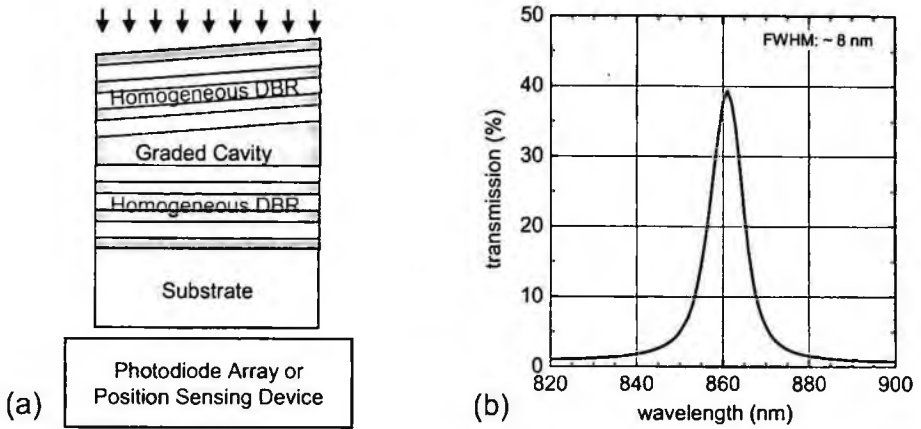


Fig. 1. (a) Concept of linear variable filter and wavelength detector. (b) Transmission properties of linear variable filter at a small location.

2.1. Linear variable filter transmission properties

The transmission properties of a LVF vary across the surface of the filter. For example, a broad band LVF for the visible range transmits at one edge the blue components of the spectrum and blocks all other components, whereas it transmits the red components at the opposite edge. This kind of filter can be realized by creating an inhomogeneous cavity between two mirrors as illustrated in Fig. 1a. A graded Fabry Perot cavity is sandwiched between two homogeneous distributed Bragg reflectors (DBRs).¹² Figure 1b shows the transmission spectrum of such a structure at a well-defined position. The wavelength λ of the transmission peak is given by the thickness $d(x)$ and the refractive index n of the cavity:

$$N\lambda = 2nd \quad N=1,2,3,\dots \quad (1)$$

As the thickness of the cavity $d=d(x)$ changes across the filter in x -direction, the transmitted wavelength $\lambda = \lambda(x)$ varies as well. Alternatively, the thickness d can be held constant and the refractive index $n=n(x)$ varied across the filter. The amount by which the transmission peak wavelength λ changes across the filter is characterized by the gradient g :

$$g = \frac{\Delta\lambda}{\Delta x} \quad (2)$$

A gradient of 30 nm/mm indicates that the spectral position of the transmission peak changes by 30 nm when the transmission spectrum is compared between two locations 1 mm apart on the filter.

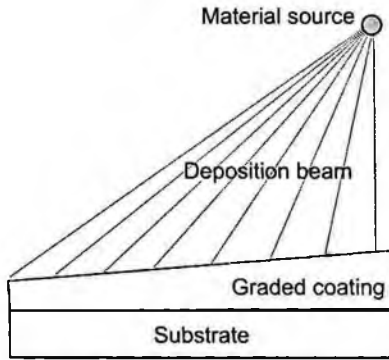


Fig. 2. (a) Setup for fabrication of inhomogeneous coatings

2.2. Fabrication of inhomogeneous coatings

There are many potential approaches to fabricate inhomogeneous Fabry-Perot cavities. We have fabricated various coatings that exhibit the properties described above, by metal-organic vapor phase epitaxy (MOVPE), molecular beam epitaxy (MBE) and electron beam deposition.

In one realization we have installed a substrate at a position off-axis in an electron beam deposition chamber as indicated in Fig. 2. The material is emanating from a point source and needs to travel a longer distance r to one edge of the substrate than the other. Since the material flux decreases as $1/r^2$, the substrate edge closest to the source receives more material than the far edge. A similar effect can be achieved with a MBE system where the cells for different materials are located at different positions inside of the MBE chamber. In this case homogeneous layers are obtained by rotating the substrate during growth, and inhomogeneous layers can be achieved by stopping the substrate rotation during the growth of the cavity.

3. Wavelength-Shift Detection

The wavelength detectors can be designed to resolve small wavelength changes over a broad spectral range. The differential readout from adjacent pixels of a detector array or from a position sensing device enables relative wavelength resolutions down to ~ 1 pm.¹³ The wavelength shift detectors are applicable for any optical sensor in which a stimulus (e.g., temperature, strain, PH-value, etc.) results in a wavelength shift of an optical output signal, such as the interrogation system for Fiber Bragg Grating sensors, photonic crystal sensors, Fabry-Perot type sensors, and sensors in which an analyte influences a laser cavity and thereby its emission wavelength.

3.1. Sub-pm wavelength-shift detection

As one illustration of our wavelength-shift detector, a GaAs/AlAs LVF with a gradient of about 3 nm/mm and a narrow spectral transmission band (FWHM) of 1 nm was

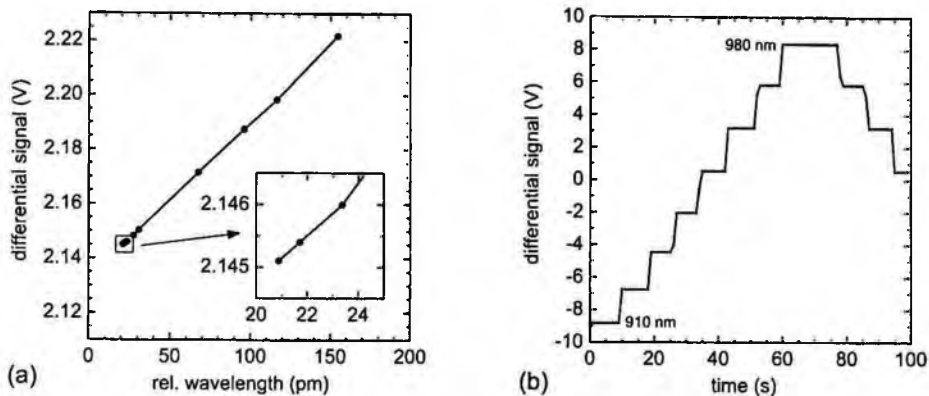


Fig. 3. Response of wavelength shift detector to (a) small and (b) large wavelength shifts of the incident light.

combined with a low-cost position sensor as depicted in Fig. 1a. A position sensor is a large-area photo detector which outputs two currents I_1 and I_2 from which the centroid of the incoming light beam can be determined very precisely. We have used a commercial 2.5 mm position sensing device that is able to detect the position of a light spot with an accuracy of about $0.1 \mu\text{m}$. The product of the coating gradient and the position sensor accuracy shows that a wavelength resolution of about 0.3 pm is feasible.

We have characterized this combination with a wavelength-tunable VCSEL. The emission wavelength can be tuned very precisely by adjusting either the laser current or temperature. Figure 3a shows the amplified differential current $(I_1 - I_2)/(I_1 + I_2)$ delivered by the position sensor converted to a voltage by a differential amplifier. The position sensor signal follows very precisely the wavelength of the incoming light. Wavelength changes in the sub-pm range can be resolved as shown in the inset of Fig. 3a. Note that the differential output signal depends only on the wavelength of the incoming light and is independent of its intensity. The largest observed wavelength shift for this particular device is about 5 nm .

3.2. Broad-band wavelength-shift detection

We chose a different configuration to demonstrate a wavelength-shift detector that can monitor wavelength shifts over a broad spectral range. A 2.5 mm position sensing device was covered with a coating that has a steep gradient of about 35 nm/mm . In order to characterize this combination we used spectrally filtered light from a halogen lamp as a tunable light source. The time sequence in Fig. 3b shows the response of the wavelength detector on wavelength changes of the incident light in 10 nm steps over a range of up to 70 nm . In this experiment the range was limited by the length of the position sensing device. The wavelength resolution in this case was about 3.5 pm , which was confirmed by using a tunable VCSEL.

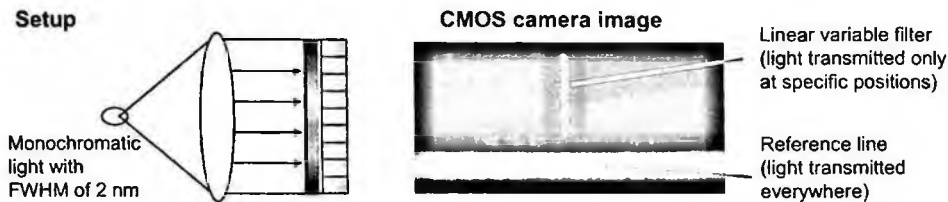


Fig. 4. (a) Setup for chip-size spectrometer calibration. (b) Snapshot of CMOS camera.

4. Spectrometer-on-a-Chip

The wavelength detectors can also be integrated into lab-on-a-chip systems due to their compactness. Applications that are anticipated to benefit from the spectrometers include spectroscopy-on-a-chip applications for reagentless identification of analytes (e.g., biomolecules or chemicals) in fluidic and aerosol samples.

4.1. Description of prototype

We have realized a compact spectrometer by combining a 12bit CMOS camera and a linear variable band-pass filter (LVF). The camera had a resolution of 1024×1024 pixels with a pixel size of $10.6 \times 10.6 \mu\text{m}$. The exposure time could be varied between $10 \mu\text{s}$ and 0.4 s . The LVF covered a spectral range from 380 to 720 nm with a FWHM of the transmission peak of $\sim 2\%$ of the transmission peak wavelength and a spectral gradient of 30 nm/mm . The LVF had dimensions of $\sim 12 \times 3 \text{ mm}$ and was positioned over a designated zone of the camera. Parallel to this zone a $12 \times 1 \text{ mm}$ reference line was designated to collect photons of all wavelengths in the spectral range. Unused parts of the camera were blocked by an aperture. The reference line is used to monitor the intensity distribution across the detector and is an important feature of this technology, as explained below.

4.2. Calibration and reference line

Light from a Halogen lamp was spectrally filtered by a monochromator and coupled into a 1 mm polymer optical fiber. At the other end of the fiber the light was collimated and directed onto the detector such that the whole detector was illuminated as illustrated in Fig. 4a. This setup was used to calibrate and demonstrate the chip-size spectrometer. Figure 4b features a snapshot of the camera illuminated with 550 nm . In the upper half of the image light propagates through the filter at a single sharp position. The position of the line is determined by the transmission properties of the LVF. The filter blocks the light at all other positions. In contrast, light is received by the detector over its entire width through the reference line in the lower half of the image. The horizontal intensity profile of the filter region is plotted in Fig. 5a for five different input wavelengths between 450 nm and 650 nm . Correlating the pixel position of the intensity maximum with the

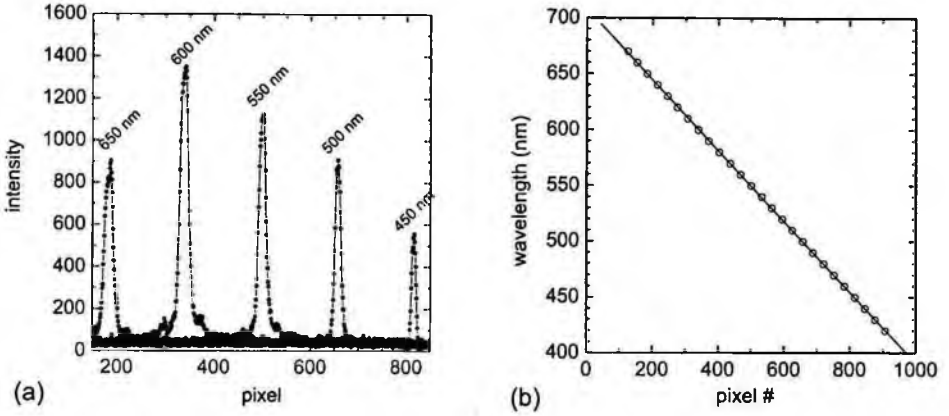


Fig. 5. (a) Intensity profiles behind LVF at different incident wavelengths. (b) Wavelength calibration curve.

known input wavelength results in a calibration of pixel number into wavelength (Fig. 5b).

The peaks in Fig. 5a have different heights for several reasons: (1) the Halogen lamp spectrum is not constant within the observed wavelength range, (2) the camera sensitivity is not constant within the entire spectral range, (3) the peak transmission of the filter varies across the filter, and (4) the illumination of the chip-size spectrometer is not homogeneous. The effects (1), (2) and (4) can be eliminated by recording the intensity profile of the reference line and normalizing the filter profiles $I_{LVF}(\lambda)$ with the corresponding reference profile $I_{ref}(\lambda)$:

$$I(\lambda) = \frac{I_{LVF}(\lambda)}{I_{ref}(\lambda)} \quad (3)$$

Figure 6a shows the calibrated and normalized spectra. The peak height increases from short to long wavelengths because the peak transmission of the LVF is lower for short wavelengths and higher for long wavelengths. The effects (2)-(4) have to be considered when measuring “real” spectra. The wavelength-dependent peak transmission of the LVF $T_{LVF}(\lambda)$ and the camera sensitivity $S_{cam}(\lambda)$ are used to correct the spectra. Furthermore, the illumination inhomogeneity I_{ill} obtained from the reference channel is used to correct the intensity:

$$I(\lambda) = \frac{I_{LVF}(\lambda)}{I_{ill} \cdot T_{LVF}(\lambda) \cdot S_{cam}(\lambda)} \quad (4)$$

In the experiment above the light was collimated with a lens. Therefore, the Gaussian shaped beam profile introduced weaker illumination of the camera at the edges than in the center. Figure 6b shows the reference profile I_{ill} for this geometry. For optimum performance the reference line should be positioned in close proximity to the LVF line.

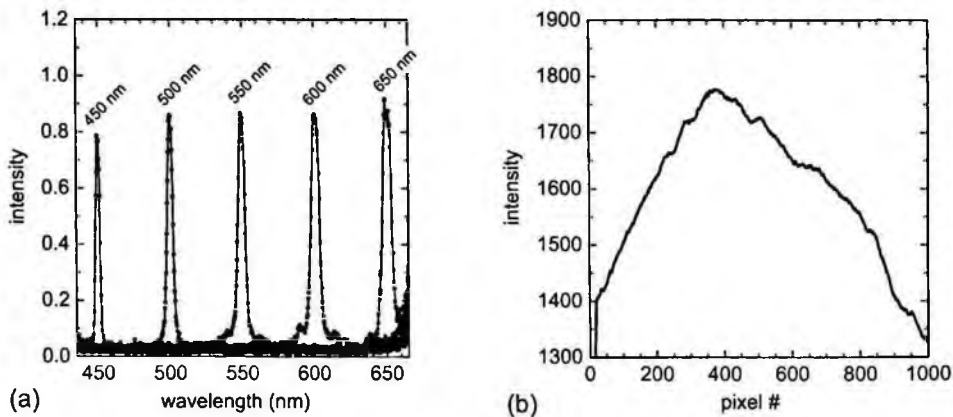


Fig. 6. (a) Spectra of various light inputs (calibrated and referenced). (b) Intensity profile of reference line.

4.3. Large area spectroscopy

The photosensitive area of the spectrometer can be as large as the size of the detector array. This area is large compared to conventional spectrometers where the light has to be focused onto a small slit in order to receive good wavelength separation. Therefore, the chip-size spectrometer is especially favorable for large light emitting areas or moving particles. The geometry of the detector enables an optical detection method that can actually take advantage of the motion of a moving optical source for large integration times and improved detection without losing throughput capacity. The spectral information of the light-emitting particle is gathered step-by-step while it is moving across the wavelength detector. Figure 7 shows the conceptual cross section of a lab-on-a-chip system that is described in further detail elsewhere in these proceedings¹⁴. Fluorescent particles move along a fluidic channel and are continuously excited by the excitation laser. The emission spectra were recorded while the particle was traversing the detector. First the red components of the spectrum were recorded, then the green and finally the blue components.

5. Conclusion

Chip-size wavelength detectors were designed by combining linear variable band-pass filters with photo-detectors. The filter converts the spectral information of the incident light into a spatially dependent signal that is analyzed by the detector. The filters can be designed either to cover a broad spectral range or to enable sub-pm wavelength resolution.

High-resolution wavelength-shift detection has been demonstrated down to the sub-pm level. For this purpose we have used simple position sensing devices in combination with a differential amplifier. We have investigated the performance of these wavelength detectors depending on the gradient of the filter. This technology can be used

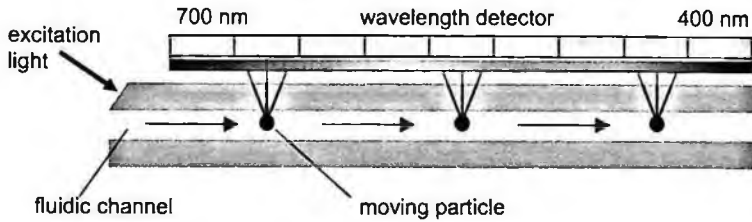


Fig. 7. (a) Conceptual cross-section of lab-on-a-chip device with integrated chip-size spectrometer.

for the read-out of optical sensors that are changing their output wavelength upon a stimulus.

Compact broad-band spectrometers enable integration of spectroscopic techniques onto lab-on-a-chip devices. We have demonstrated a spectrometer for the visible spectral range and have discussed calibration and referencing techniques. Inclusion of a reference line that monitors the intensity distribution of the incoming light is essential for this kind of wavelength detectors to eliminate errors from inhomogeneous illumination. Due to their extended detection area, chip-size spectrometers are especially favorable for large light emitting areas or moving particles. The spectrometer was illustrated by recording fluorescence spectra from objects as they traversed the detection area.

Acknowledgements

The authors are pleased to acknowledge technical support and helpful discussions from Mark Teepe, Setu Mohta, and Stefan Malzer. Financial support was provided by ONR under grant N00014-05-C-0430 monitored by Jeremy Walker, Janet Jensen, and Paul Armistead.

References

1. W. J. Fleming, *IEEE Sensors Journal* **1**, 296 (2001)
2. C. I. Merzbachery, A. D. Kersey, E. J. Friebele, *Smart Mater. Struct.* **5**, 196 (1996)
3. E. Udd, *Proc. SPIE* **6167**, 61670C (2006)
4. O. S. Wolfbeis, *Anal. Chem.* **76**, 3269 (2004)
5. B. Lee, *Optical Fiber Technology* **9**, 57 (2003)
6. A. Othonos, K. Kalli, *Fiber Bragg Gratings* (Artech House, Boston, 1999)
7. J. Homola, S. Yee, and D. Myszka, *Surface Plasmon Resonance Biosensors*, Chapter 7 in *Optical Biosensors, Present and Future* edited by F. Liger and C. A. Rowe Taitt (Elsevier Science, New York, 2002)
8. H. Ouyang, C. C. Striemer, and P. M. Fauchet, *Appl. Phys. Lett.* **88**, 163108 (2006)
9. B. Cunningham, P. Li, S. Schulz, B. Lin, C. Baird, J. Gerstenmaier, C. Genick, F. Wang, E. Fine, L. Laing, *Journal of Biomolecular Screening* **9** 481 (2004)
10. F. Vollmer, D. Braun, A. Libchaber, M. Khoshshima, I. Teraoka, S. Arnold, *Appl. Phys. Lett.* **80**, 4057 (2002)
11. J. A. Wahl, J. S. Van Delden, S. Tiwari, *Appl. Opt.* **44**, 5190 (2005)

12. J. A. Wahl, J. S. Van Delden, S. Tiwari, *IEEE Photon. Technol Lett.* **16**, 1873 (2004)
13. O. Schmidt, P. Kiesel, S. Mohta, N. M. Johnson, "Resolving pm wavelength shifts in optical sensing" submitted to *Appl. Phys. B*
14. P. Kiesel, O. Schmidt, M. Bassler, N. M. Johnson, "Compact optical characterization platform for detection of bio-molecules in fluidic and aerosol samples", in *same issue as this article*

CLASS IDENTIFICATION OF BIO-MOLECULES BASED ON MULTI-COLOR NATIVE FLUORESCENCE SPECTROSCOPY

MICHAEL BASSLER, OLIVER SCHMIDT, PETER KIESEL, and NOBLE M. JOHNSON

*Palo Alto Research Center Inc., 3333 Coyote Hill Road,
Palo Alto, CA 94304, USA
michael.bassler@parc.com*

Laser-induced native fluorescence is measured on a set of bio-molecules from different classes (bacteria, proteins, fungi) for excitation at 266nm and 355nm. A method of preprocessing the spectra to obtain an inherently normalized set of data is described. Class identification on the normalized data set is demonstrated.

Keywords: class identification, bio-molecules, laser induced native fluorescence

1. Introduction

Biological molecules are composed of only a few basic building blocks and, therefore, exhibit similar physical properties. In particular, aromatic amino acids can be found in biological molecules and exhibit native fluorescence. Among the amino acids tryptophan, phenylalanine and tyrosine show fluorescent emission in the ultraviolet at 350nm, 280nm and 300nm, respectively. In addition, many enzymes or cofactors, NADH (reduced b-nicotinamide adenine dinucleotide, 460nm) and Riboflavin (550nm) being the most prominent examples, exhibit pronounced native fluorescence in the visible range.

Laser-induced fluorescence (LIF) is a promising tool for differentiating fluorescing molecules.^{1,2} However, the variety of biological molecules is huge compared to the number of basic building blocks. Therefore, the fluorescence spectra of different analytes are often quite similar, and high spectral resolution of the emission spectra is required to reveal differences. Additionally, multi-wavelength LIF is essential to improve the specificity of the spectra and enable class identification of bio-molecules.

In order to test class identification based on multi-color native fluorescence we have measured LIF spectra for representative bio-agent simulants in solution for various classes of bio-molecules (i.e., toxins, viruses and bacteria) with a conventional laboratory set-up. For a proof-of-concept, the library of spectra was subjected to a mathematical procedure to extract and evaluate distinguishing features of the analytes. In a second step, the extracted distinguishing features were used to teach a multidimensional classifying

algorithm to find a classification rule for the set of analytes. A classification rule defines the procedure to identify an unknown analyte from a single measurement.

2. Experimental

2.1. Fluorescence measurement setup

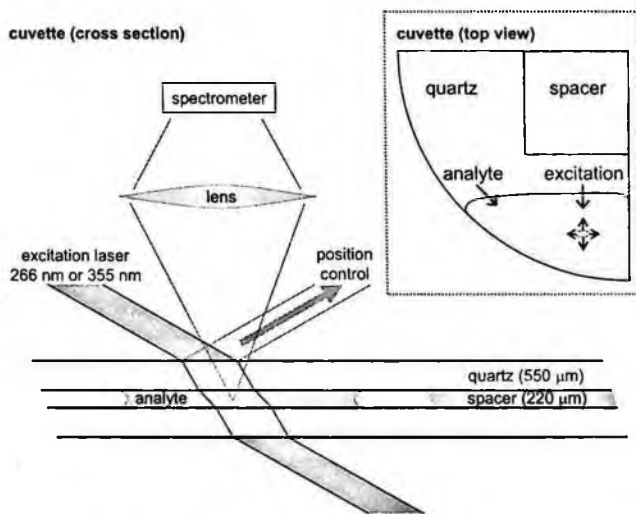


Fig. 1. Experimental setup for the measurement of laser-induced native fluorescence on bio-molecules in aqueous solution.

Figure 1 depicts the setup for the measurement of native fluorescence. An analyte containing the bio-molecule under investigation is injected between two quartz wafers (thickness 550 μm each) which are separated by a spacer (220 μm); this setup is referred to as a cuvette in the following. The analyte is contained within the cuvette by capillary forces. As the excitation source either a 266nm laser (frequency-quadrupled Nd:Yg, JDS Uniphase) or a 355nm laser (frequency-tripled Nd:Yg, CryLas) was used and the beam was coupled through the quartz surface of the cuvette at an angle of 51 deg with respect to the surface normal. To ensure constant excitation intensity in the volume under investigation, the reflected portion of the excitation light was used as a position control for the beam. The fluorescence signal was collected by a conventional spectrometer (Oriel) with a CCD chip in the detection plane (Andor DV 480-BV). Dielectric filters (Asahi Spectra Corporation, Japan) with a cut-on at either 275 nm and 385 nm were used to suppress scattered excitation light by four orders of magnitude. The cuvette was mounted on a two-axis positioning stage to select the volume under investigation. To reduce the background fluorescence due to surface contamination, the cuvettes were subjected to the following cleaning procedure: (1) 5 min methanol, (2) 2 min rinsing with water, (3) 10 min etching in an mixture of $\text{H}_2\text{O}_2:\text{NH}_4:\text{H}_2\text{O}$ (1:1:5) at 70 deg C, (4) 2 min rinsing in water and (5) drying under nitrogen flow. The background fluorescence was

monitored for each cuvette prior to the injection of the analyte for both excitation wavelengths.

The simulants (see tab. 1) were chosen to provide representation in each of the classes of bacteria, fungi and proteins. All simulants were desolved in DI water and, to aid comparison, the concentration level was chosen to result in similar fluorescence peak intensities for all analytes.

Table 1. List of simulants investigated

Class	Simulant (synonym)	Concentration
Amino acid	Tryptophan (TP)	hc: 2mg/ml, lc: 0.02mg/ml
Protein	Bovine Serum Albumin (BSA)	hc: 1mg/ml, lc: 0.1mg/ml
	Horse Heart Cytochrome C (HHC)	1mg/ml
	Nicotinamide Adenine Dinucleotide (NADH)	1mg/ml
Bacterium	Bacillus Thuringiensis (BT)	1×10^8 /ml
	Bacillus Globii (BG)	1×10^8 /ml
	Escherichia Coli (E.coli)	1×10^9 /ml
Fungus	Yeast (YE)	hc: 1mg/ml, lc: 0.1mg/ml

The reproducibility of the fluorescence spectra was tested by varying the lateral position of the excitation spot. Artifacts were observed when the beam was near the edges of the cuvette and the solution. These were mostly caused either by fluorescence of the edges themselves or by increased analyte fluorescence caused by enhanced excitation intensity due to internally scattered or reflected light. To effectively suppress these artifacts the distance between the edge of the analyte and the excitation spot had to be larger than 2mm. For each analyte, the spectra were taken for five different excitation spots to confirm lateral homogeneity and the absence of surface contamination which might be located on the outer surface of the cuvette.

2.2. Native fluorescence spectra

Figure 2 shows the fluorescence spectra of BT (see table 1) when excited at 266nm (dashed lines) and 355 nm (solid lines). For each excitation wavelength five spectra are plotted which were taken at five positions on the cuvette. The small deviations within one set of spectra confirm the lateral homogeneity and reproducibility of the fluorescence measurement. The sharp peak at 355nm is caused by residual transmission of the filter for scattered excitation light. The spectra have a Raman peak superimposed at 293nm (266nm excitation) and 404nm (355nm excitation) with a corresponding Raman shift of $3,460\text{cm}^{-1}$ and $3,420\text{cm}^{-1}$, respectively, which can be attributed to the water.³

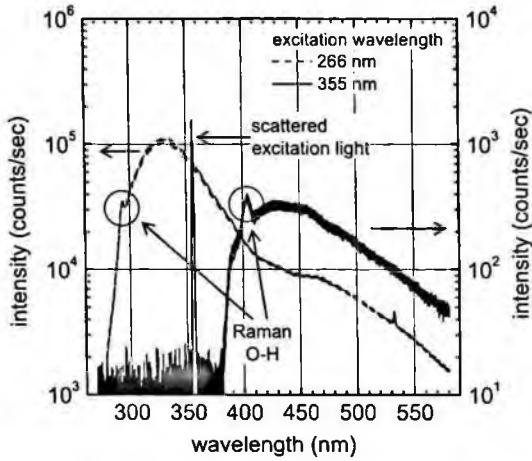


Fig. 2. Fluorescence spectra of BT for 266nm (dashed lines) and 355nm (solid lines) excitation. Each set of curves contains five measurements at different positions.

Figure 3 displays for both excitation wavelengths representative spectra for a selection of the analytes. BG and E.coli are not shown for the sake of clarity. BSA (dotted line in Fig. 3a and 3b) and Yeast (dashed line in Fig. 3a and 3b) were measured in two concentrations “hc” and “lc” which differed by one order of magnitude. The resulting fluorescence intensities scale linearly with the difference in concentration. This reveals that internal re-absorption in the analyte and inhomogeneous absorption of the excitation light over the excitation path play minor roles in the chosen concentration regime. Overall BSA and Yeast reveal the most similar spectra in the set of analytes. The first obvious distinctive feature in the 266nm-spectra is the slightly different position of the maximum at 333nm for Yeast and 337nm for BSA. The second distinctive feature is the higher intensity at the maximum for BSA in comparison to the intensity in the tail of the spectrum between 450nm and 580nm. In comparison to Yeast and BSA, the maximum of HHC is much broader and spans from 300nm to 340nm. The small local minimum in that range is caused by the filter transmission curve and is not specific to HHC. No significant feature can be seen in the 355nm-spectra in Fig. 3b.

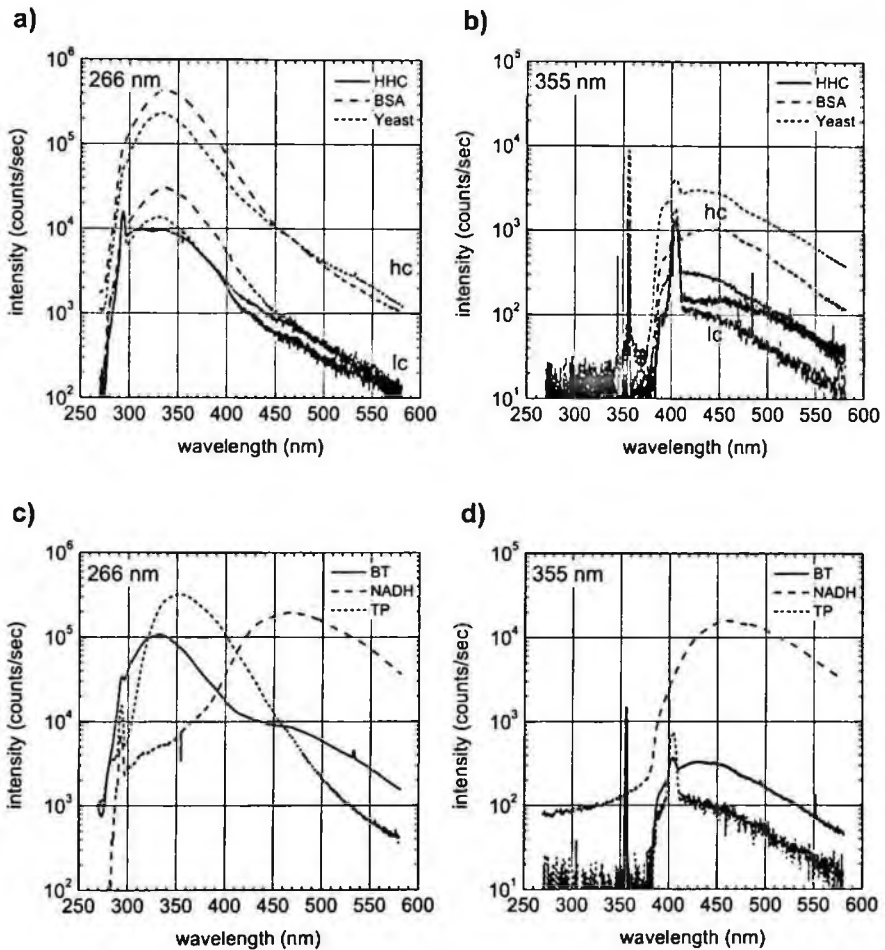


Fig. 3. Comparison of LIF spectra for HHC, BSA, Yeast, BT, NADH and TP for excitation with 266nm (a, c) and 355nm (b,d). BSA and Yeast were measured in two concentrations marked by "lc" and "hc"; for details see Tab. 1.

For BT, TP and NADH the distinctive features are much more obvious (Fig. 3c). BT (solid curve) exhibits a pronounced shoulder above 450nm which is not observed in other analytes. The spectra for BG (not shown), which is a close relative to BT, have the same shoulder. TP (dotted curve) and NADH (dashed curve) can clearly be characterized by the position of the maximum which is at 352nm for TP and 468 for NADH. Again, no significant feature can be seen for the 355nm excitation (Fig. 3d).

3. Class identification

Even though it is possible to extract distinguishing features by carefully analyzing the spectra for a small number of analytes, it turns out that the identification of an analyte cannot be accomplished by using a single attribute, for example, peak position. In the following we will describe the results of a more general and automated approach to extract distinguishing features from the spectra.

3.1. Data normalization

Besides the spectral response of the analyte, the measured intensity of a fluorescence spectrum is controlled by the analyte concentration, the intensity of the excitation light, and the excitation and emission efficiencies. As the concentration fluorescence efficiency of the analyte is typically not known and the intensity of excitation may vary for various reasons, a method for normalization is required to preprocess the data prior to the classification. In the following we describe the method we used for normalization.

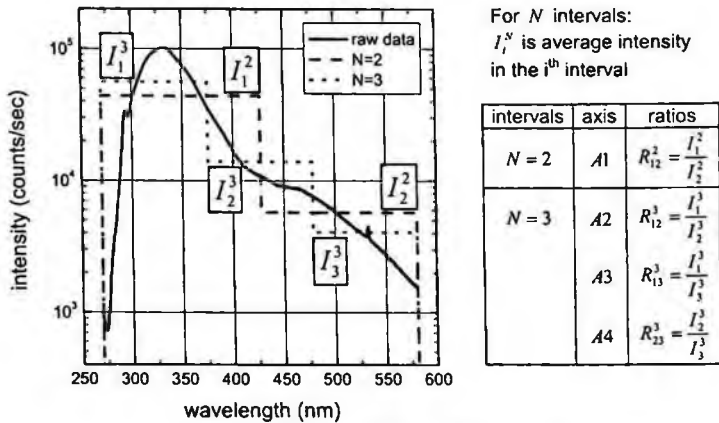


Fig. 4. Preprocessing the data leads to one ratio for $N=2$ and three ratios for $N=3$; the corresponding axes are numbered sequentially.

For our normalization method, each spectrum was first subdivided into N (with $N=2, 3, \dots, N_{\max}$) intervals of equal length, where the maximum number of intervals is N_{\max} . For each subdivision N and each interval i the average intensity I_i^N ($i=1, 2, \dots, N$) was calculated (compare Fig. 4) and for each pair (i, j) with $i < j$ the following ratio was determined:

$$R_{ij}^N = \frac{I_i^N}{I_j^N} \quad (1)$$

The ratios for all pairs (i, j) for which $i \geq j$ were not used as they are simply reciprocals of other ratios or identity ($i = j$) and, therefore, do not contain additional information

about the spectrum. Thus, a set of characteristic ratios for each spectrum was acquired where the spectral resolution required to actually measure the particular ratio is given by the length of the interval. Note that each ratio is inherently normalized because both excitation intensity and analyte concentration act as a multiplicative factor which cancels out. To extend this procedure to multi-wavelength excitation, the normalization for excitation intensity has to be performed first for all spectra.

In the following we refer to a particular set of (N, i, j) as an “axis” on which the ratios for different measurements of the same or different analytes can be compared. This procedure leads to a total number of axes:

$$N_{ax} = \sum_{N=2}^{N_{max}} \binom{N}{2} \quad (2)$$

3.2. Meaningful axes

Identifying the “meaningful” axes, this is, those axes that contain the most distinguishing features, is important for three reasons:

- (1) The maximum N from among the meaningful axes establishes the maximum spectral resolution that is needed to identify the analyte in an unknown sample,
- (2) the contributing intervals define the spectral region needed to identify the analyte, and
- (3) identifying the minimum number of essential axes minimizes the computational effort to define the classification rule.

Here we describe one method to select meaningful axes. For each axis the ratios obtained from several measurements on *one* analyte “A” can be characterized by the average ratio

$$\langle R_{ij}^N(A) \rangle \quad (3)$$

and the standard deviation of the distribution of ratios

$$\sigma_{ij}^N(A). \quad (4)$$

Two analytes can be distinguished on one axis when the distribution of their ratios does not overlap. To assess the distance $d_{ij}^N(A, B)$ of two distributions for analytes A and B we chose the following metric

$$d_{ij}^N(A, B) = \frac{|\langle R_{ij}^N(A) \rangle - \langle R_{ij}^N(B) \rangle|}{\sigma_{ij}^N(A) + \sigma_{ij}^N(B)}. \quad (5)$$

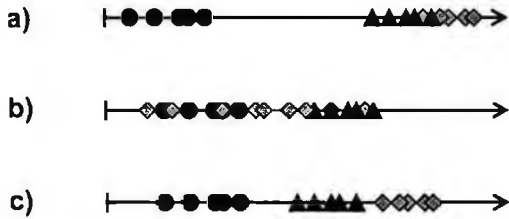


Fig. 5. Basic criteria for the selection of meaningful axes; a) one analyte is best separated from all others, b) best separation for a pair of analytes, and c) all analytes are separated on one axis.

Analytes A and B can be considered to be distinguishable on axes where $d_{ij}^N(A, B) > 1$, that is, the distance between the average ratios exceeds the sum of the standard deviations. For identification of multiple analytes it is not sufficient to consider the distance between each pair of analyte. It is also necessary to take into account the position of the remaining analytes and define selection criteria. Figure 5 displays three criteria we have used to select axes: (a) one analyte is best separated from all others, (b) best separation for each pair of analytes, and (c) all analytes are separated. The last criterion, of course, describes the ideal situation where one axis is sufficient to perform class identification on the selected set of analytes.

3.3. Data analysis

The preprocessing of data and selection of meaningful axes were conducted on a total of 85 spectra taken on the set of analytes listed in Table 1 for excitation with 266nm and 355nm. In Fig. 6 a) and b) the ratios are plotted for the four meaningful axis that were closest to fulfilling criterion c) (“all analytes are separated on one axis”, see Fig. 5c), although none met this criterion. Nevertheless, two axes (A865 and A375) were sufficient to separate all analytes in two dimensions (Fig. 6a). The ratios for axis A865 were formed from the average intensities at the intervals [340nm±8.5nm] and [357nm±8.5nm], and for axis A375 the intervals were [283nm±11nm] and [525nm±11nm]. Both axes utilize spectra from 266nm excitation. Note that no meaningful axis was obtained from 355nm excitation. Thus, for the chosen set of analytes, class identification was achieved with excitation only at 266nm, with the spectra recorded in the range from 280nm to 530nm with a spectral resolution of 17nm.

Figure 6b illustrates that additional axes contain independent useful information: in this plot the separation for HHC is much clearer than in Fig. 6a, where it almost overlaps with BSA. It is obvious that a multidimensional classification algorithm would be favorable to perform the class identification on the set of meaningful axes.

3.4. Multidimensional classification

Figure 6a gives rise to an exemplary classification rule in two dimensions: assignment of a measured analyte follows from association with a specific zone in a two-dimensional plot of the axes A865 and A375. This specific “rule” applies only for the selected set of

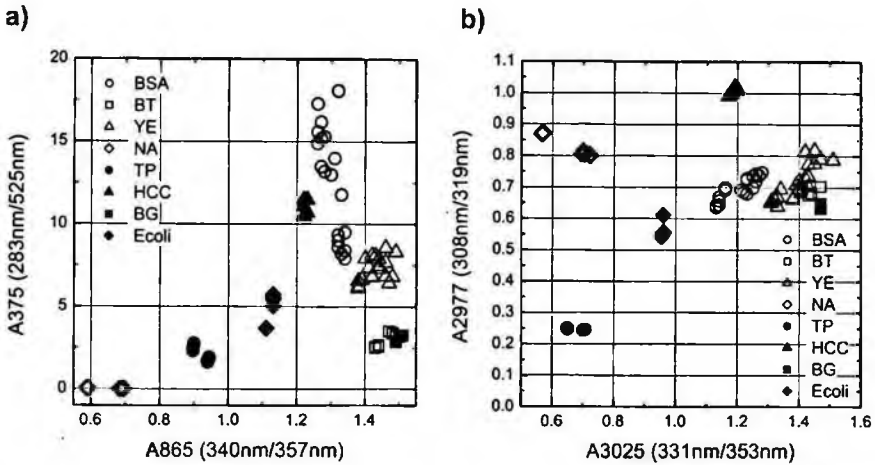


Fig. 6. 2-dimensional plot of ratios using meaningful axis. The underlying spectral resolution is 22nm (A375), 17nm (A865), and 11nm (A2977, A3025).

analytes and is presented for illustrative purposes. For large numbers of analytes, one can rely on readily available multidimensional tools for classification. To test our library of spectra we chose WEKA 3.4.7, which is an open source software available under the GNU General Public License. WEKA contains a broad selection of tools, and its use and application is described in the book of I.H. Witten and E. Frank⁴.

We used several classifiers described in WEKA and tested the results with the built-in cross validation capability, which works as follows:

- (1) Test data are generated by taking 10% of the analytes out from the library
- (2) The rest of library is used to gain classification rules from the classifier
- (3) The set of test data is used to check the classification rules

We compared the classification for three libraries: (1) the original normalized spectra, (2) the full set of axis and (3) the selection of meaningful axis. We observed that the classification works best on the set of meaningful axes where the cross validation resulted in no classification error for all applicable WEKA classifiers. Wrong classifications occurred for the normalized spectra and the full set of axes.

4. Summary

We successfully demonstrated class identification via laser-induced native fluorescence on a selection of bio-molecules from different classes (bacteria, proteins, fungi). We presented a method for processing the spectra to achieve a normalized set of data for each measurement. Further, we described routines to select the coordinates with the most distinguishing features between the different analytes. This information can be used as direct input for classification algorithms. Excitation with a 266nm laser and recording the fluorescence spectra with a moderate spectral resolution of 17nm were sufficient to reliably identify the analytes under investigation.

5. Acknowledgments

The authors are pleased to acknowledge technical support and helpful discussions from Ben Hsieh and Armin Voelkel. This work was supported by ONR through the EUWP Program, contract # N00014-05-C-0430, monitored by Jeremy Walker, Janet Jensen, and Paul Armistead.

References

1. V. Sivaprakasam, A. Houston, C. Scotto and J. Eversole, Multiple UV wavelength excitation and fluorescence of bioaerosols, *Opt. Express* **12**, 4457-4466. (2004)
2. Y.S. Cheng, E.B. Barr, B.J. Fan, P.J. Hargis, D.J. Rader, T.J. O'Hern, J.R. Torczynski, G.C. Tisone, B.L. Preppernau, S.A. Young, R.J. Radloff. Detection of bioaerosols using multiwavelength UV fluorescence spectroscopy, *Aerosol Sci. Tech.* **30**, 186-201 (1999)
3. G.E. Walrafen, Raman spectral studies of the effect of temperature on water structure, *J. Chem. Phys.* **47**, 114-126. (1967)
4. I.H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques* (2nd Edition, Morgan Kaufmann, San Francisco, 2005)

CHESAPEAKE BAY WATER QUALITY MONITORING USING SATELLITE IMAGERY

BHARATH RAMAKRISHNA and CHEIN-I CHANG

*Remote Sensing Signal and Image Processing Laboratory
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County, Baltimore, MD 21250*

BRUCE TROU and JERRY HENQEMIHLE

Microtel LLC, 111 Centerway, Greenbelt, MD 20770

Abstract

The Chesapeake Bay is a valued ecological, economic, recreational, cultural and scenic resource. The Bay watershed States and the District of Columbia, in conjunction with the EPA Chesapeake Bay Program, have worked and teamed together over the past 20 years to protect and restore the Bay ecosystem. A key component of this effort is water quality and habitat monitoring to assess the impact of management actions and natural processes, and evaluate habitat parameters on living resources such as submerged aquatic vegetation (SAV), oysters, and fisheries. Using aerial and satellite remote sensing imagery has become a practical and effective means of monitoring water quality in a timely manner. Of particular interest in evaluation of water clarity are several initiative measures. Specifically, Secchi-Disk Transparency (SDT) and Chlorophyll *a* (Chl-*a*) have been widely accepted as critical indicators of water quality and their reliable estimation using satellite imagery provides a cost effective and speedy means for water quality monitoring. Work done at Water Resources Center, University of Minnesota has demonstrated the feasibility of performing regional assessment of lake water quality using LANDSAT image data. This paper investigates an approach similar to their work but uses a different type of satellite imagery, EO-1 ALI imagery where the SDT and Chl-*a* are also used as indicators to estimate water quality for the Chesapeake Bay and DC area (Potomac River). In doing so, three major issues are investigated, which are (1) the study site that is an open Bay area, not a self-contained lake; (2) investigation of applicability of equations that are used to specify the SDT and Chl-*a* to our Bay area study; (3) the use of a different type of satellite imagery for water quality monitoring. This paper develops techniques to address these three issues and presents preliminary experiments which show encouraging results.

Key words: *Chesapeake Bay, EO-1 ALI Imagery, Secchi Depth (SDT), Chl-*a*.*

1. Introduction

The Chesapeake Bay is a valued ecological, economic, recreational, cultural and scenic resource. The Bay watershed States and the District of Columbia, in conjunction with the EPA Chesapeake Bay Program, have teamed for the past 20 years to protect and restore the Bay ecosystem. A key component of this effort is water quality and habitat monitoring to assess the impacts of management actions and natural processes, and

evaluate habitat parameters for living resources such as submerged aquatic vegetation (SAV), oysters, and fisheries. The Maryland Department of Natural Resources (DNR), Tidewater Ecosystem Assessment (TEA) Division is responsible for the collection, analysis and reporting of ambient water quality and habitat measurements throughout the tidal waters of Maryland. The TEA currently collects a full suite of biological, chemical and physical parameters at fixed stations in open, deep water, and shallow water habitats; and uses ship board monitoring to provide measurements in tributary waters.

Work performed at Water Resources Center, University of Minnesota has demonstrated the feasibility of performing regional assessment of lake water quality using Landsat imagery [1-3]. The aim was to fine tune the algorithms available from this research to process imagery collected by NASA Advanced Land Imager (ALI) to generate water quality mappings (secchi depth transparency and chlorophyll *a*) of the Chesapeake Bay. Extensive ground truth measurements collected routinely and specifically during satellite overflights by Maryland DNR were utilized to verify results. Accurate and timely water quality maps of the Chesapeake Bay will provide a valuable source of information for water management and policy decision makers. Remote sensing could prove useful in the evaluation of these criteria. Satellite and aerial overflight remote sensing techniques have been evaluated on a limited scale, but have not been fully assessed due to limited resources and expertise.

2. Water Quality Metrics

2.1. Secchi Depth Transparency

The Secchi disk is a device used to measure water transparency in open waters of lakes, bays, and the ocean. It's a patterned disk which is lowered slowly down in the water and the depth at which the pattern on the disk is no longer visible is taken as a measure of the transparency of the water. Secchi disk measurements have been an integral component of water quality assessment programs for some time they are used to reveal general trends in water quality. In spite of it being little subjective and being affected by sun glare etc Secchi disk is an inexpensive and straight-forward method of measuring water clarity.

2.2. Chlorophyll – a

Chlorophyll-*a* data can be employed to give an estimate of primary production, but the relationship between biomass and primary productivity. The main cause of excessive algae growth appears to be increased nutrient inputs. It is also affected by declines in the abundance of filter-feeders reduced aquatic sediments i.e. increased levels of light penetration, increased water temperature and changes in hydrodynamics. Procedure for Satellite Image Analysis

3. Regression Analysis to link satellite imagery to ground samples

3.1. Data Collection and Pre-processing

- (a) Satellite Image - Suitable Landsat, ALI Imagery of the Bay, DC area were chosen to do the analysis.
- (b) Geo Referencing - Geo-referencing maps the satellite imagery pixels to latitude and longitude values.
- (c) Ground Truth - Ground Truth samples obtained were +/- 2 days and +/- 1 day window of the image date for SDT and Chl-a respectively.
- (d) Water Only Image – Water only image is easily created by carrying out simple image classification.
- (e) Morphological Operations - Morphological operations like closing, erosion and mask out sections close to shoreline etc. Pixels which are close to the shore, influenced by presence of vegetation are removed.

3.2. Secchi Depth Transparency (SDT) Mapping

This section uses two well known equations [1-3] estimating Secchi Disk Transparency (SDT) using Satellite Imagery. Both of these equations are used to estimate the SDT for the given sample points.

$$\ln(\text{SDT}) = a(\text{band1}/\text{band3}) + b(\text{band1}) + c \quad (1)$$

$$\ln(\text{SDT}) = a(\text{band1}) + b(\text{band2}) + c(\text{band3}) + d \quad (2)$$

3.3. Chlorophyll (Chl-a) Mapping

Two regression equations were used to estimate Chlorophyll-*a* (chl-*a*) using Satellite Imagery to assess their suitability – one providing reasonable results in Michigan lakes [1-3], equation 3, and an alternate equation sited as used effectively to estimate inland water quality at respective sites [4], equation 4.

$$\ln(\text{Chl-a}) = a(\text{band2}) + b(\text{band3}) + c(\text{band7}) + d \quad (3)$$

$$\ln(\text{Chl-a}) = a(\text{band2}) + b(\text{band3}/\text{band1})^2 + c \quad (4)$$

We will investigate two different readily available techniques that have been successfully used to measure chlorophyll a concentration. These techniques are both based on reflectance in the red and near infra red (NIR). The basic concept of these algorithms [5] is the inclusion of the spectrum range, which show the maximal sensitivity to changes in chlorophyll concentration (~700 nm) and the range with the minimum sensitivity (~670 nm) to variation of chlorophyll concentration.

4. Experiments

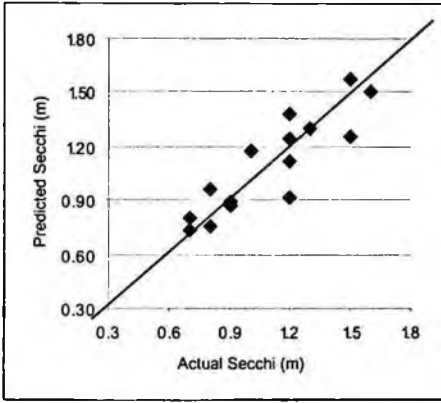
4.1. Study Area

Study Area	Image Date/Description	Ground Sampling Distribution	Cloud Cover
1	DC / Main Bay 5/9/05 Landsat 7	DC & Bay samples (5/9 -5/11)	5 %
2	DC / Main Bay 5/9/05 Landsat 7	DC (5/9 only)	5 %
3	Main Bay 6/28/05 ALI	Bay (6/28 – 6/30)	45 %
4	DC 5/16/05 ALI	Potomac (5/16)	30 %

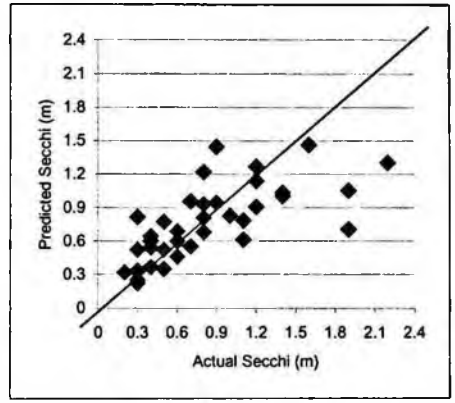
The Landsat-7 image was used because we were concerned that we might not be able to acquire any suitable imagery of the main stem of the Chesapeake Bay. This is actually the best available image and covers DC and essentially the Maryland region of the Bay. Because it received the most analysis, it is presented here. The 6/28/05 EO-1 image has significant clouds (~45%), does not have a lot of sample points available due to clouds, samples points are in shadow, and has a clear image registration problem that was not address. The 5/9/05 DC image is clear, but suffers from image registration problem that we did not address. Rain fall, re-suspension of particles in shallows, and other effects bring about a more stringent requirement to collect samples as close as possible to satellite collects as possible. Therefore, samples taken from multiple days will provide diminished results. Even for the 6/28/05 ALI Main Bay image, which really is not a suitable image for analysis due to its poor quality, a clear trend between predicted and measured values is apparent.

4.2. Secchi Depth Transparency (SDT) and Chlorophyll-a Mapping

Figures 1 and 2 show the scatter plots actual vs. predicted for SDT and Chl-*a* respectively. The presented results indicate that the usage of EO-1 ALI or Landsat is a viable means to perform Secchi Disk Transparency prediction. The high r^2 values (0.77 – 0.80) observed using same day data indicates the strong correlation of secchi depth with ALI and Landsat 7 spectral bands. The bay is a very dynamic environment. Rain fall, re-suspension of particles in shallows, and other effects bring about a more stringent requirement to collect samples as close as possible to satellite collects as possible. Therefore, samples taken from multiple days will provide diminished results. Even for the 6/28/05 ALI Main Bay image, which really is not a suitable image for analysis due to its poor quality, a clear trend between predicted and measured values is apparent. Secchi and Chlorophyll maps for DC area are shown in figures 3 and 4.

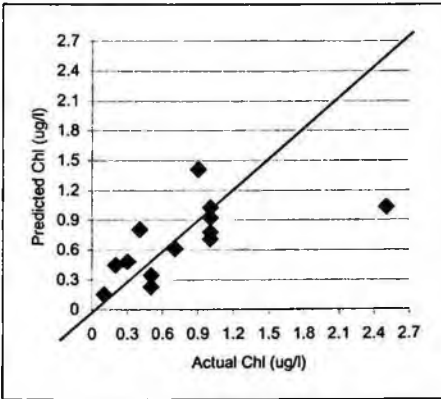


(a)

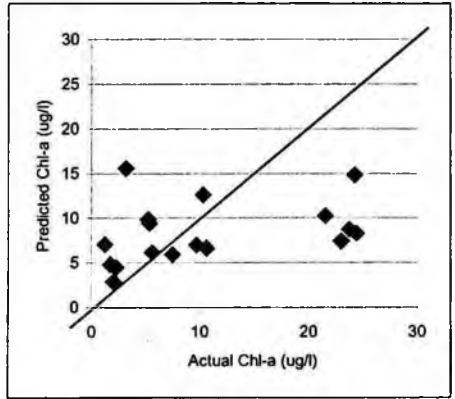


(b)

Fig. 1. SDT Scatter Plot with $r^2=0.8$ for 5/16/05 DC ALI Image and $r^2=0.6$ for 5/9/05 DC Landsat Image



(a)



(b)

Fig. 2. Chl-a Scatter Plot with $r^2 = 0.58$ for 5/16/05 DC ALI Image and Chl-a Scatter Plot with $r^2 = 0.20$ for 5/9/05 DC Landsat Image

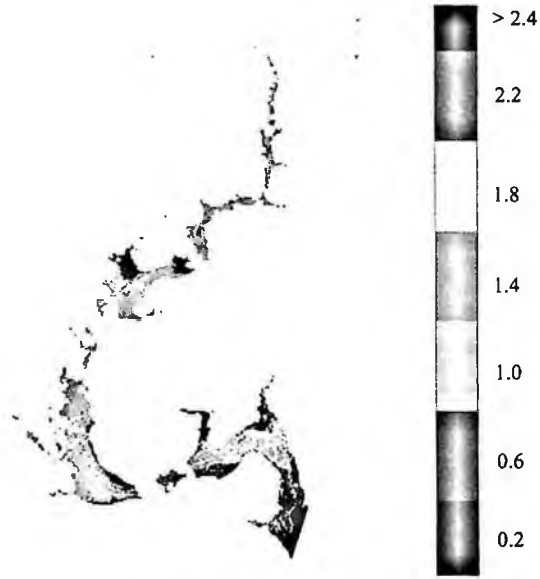


Fig. 3. SDT Map of the DC area for 5/16/05 DC ALI Image

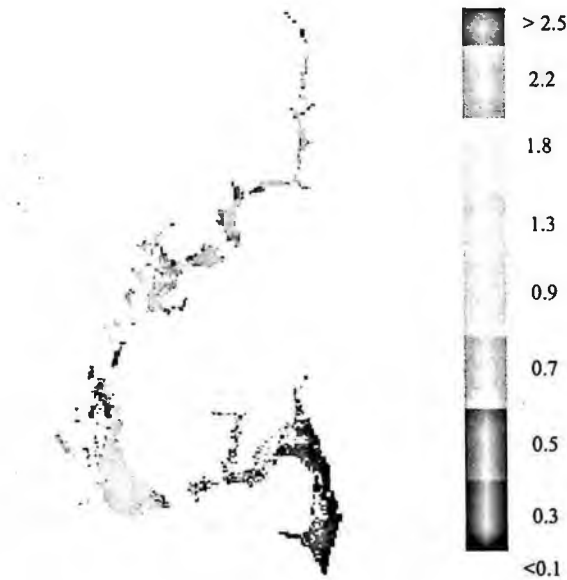


Fig. 4. Chl-a Map of the DC area for 5/9/05 DC ALI Image

5. Conclusion

Coordination of satellite imagery and ground truth measurement collection was carried out. Algorithms from the University of Minnesota lake monitoring program were used to

utilize collected satellite imagery to provide water quality assessments (secchi depth and chlorophyll a) of the bay. Water quality maps of Maryland Chesapeake Bay region utilizing collected satellite imagery and ground truth measurements are also provided. The obtained results demonstrate the feasibility of using EO-1 ALI or Landsat 7 imagery in prediction of Secchi Disk Transparency. The high r^2 values of (0.77 – 0.80) observed using same day data indicates the strong correlation of secchi depth with ALI and Landsat 7 spectral bands. The results also show the feasibility of using EO-1 ALI or Landsat 7 imagery to provide Secchi Disk Transparency prediction for the Chesapeake Bay. The work done so far is merely a proof of concept, and further work is required. Some factors surrounding the SDT that need to be investigated are size of the pool of images, better image registration using GCP's, sample data with varying latencies to understand sampling timeline requirements, better characterization of regression coefficients varying over time for the same location and at different points in the bay, and different and varying water environments in order to understand water sampling requirements, hone discrimination of areas effected by submerged aquatic vegetation, bottom. Additionally, shoreline effects, pixel averaging, spatial filtering, smoothing should be also investigated in an attempt to improve SDT prediction.

Chlorophyll results presented in this paper only provide a preliminary study. The analysis worked reasonably well on a water area having very low chlorophyll (samples < 5 mg/l), but fell apart for other images having moderately high chlorophyll (up to 40 mg / l). There were several factors that potentially have held down results that we would have liked to have investigated further. Chlorophyll estimation is highly sensitive both temporally and spatially. In reviewing Chl-*a* sample data from continuous monitor sites, large variances can be seen over periods on the order of an hour. Current chlorophyll estimation techniques take advantage of very narrow absorption and reflectance regions in the visible and near infra red regions which are not discernable using a multispectral imager such as Landsat or ALI.

Acknowledgements

Bruce Trout, Jerry Henqemihle, (301) 345-4005, Microtel LLC, 111 Centerway, Greenbelt, MD 20770 Fax: (301) 345-4005 for providing the data sets for the study and support through funds from the NASA SBIR grant.

References

1. Brezonik, P.L., Kloiber, S.M., Olmanson, L.G., and Bauer, M.E. 2002. Satellite and GIS tools to assess lake quality. Tech. Rept. 145, Water Resources Center, Univ. of Minnesota, St. Paul, 8 pp.
2. Kloiber, S.M., Brezonik, P.L., Olmanson, L.G., and Bauer, M.E, 2002a. A procedure for regional lake water clarity assessment using Landsat multispectral data, *Remote Sens. Environ.* 82(1), 38-47.
3. Predicting Water Quality by Relating Secchi-Disk Transparency and Chlorophyll *a* Measurements to Satellite Imagery for Michigan Inland Lakes, August 2002.

4. Baruah, P. J., et al., Neural Network Modeling of Lake Surface Chlorophyll and Sediment Content from Landsat TM Imagery, 2001.
5. Jerry C Ritchie, Paul V Zimba, Hydrological Application of Remote Sensing: Water Quality: Suspended Sediment and Algae, Encyclopedia of Hydrological Sciences. Edited by M. Anderson., 2005 John Wiley & Sons, Ltd.

LIF DETECTION OF TRACE SPECIES IN WATER USING DIFFERENT UV LASER WAVELENGTHS

ANNA V. SHARIKOVA^a

*Department of Physics, University of South Florida,
Tampa, Florida 33620, USA
ashariko@mail.usf.edu*

DENNIS K. KILLINGER

*Department of Physics, University of South Florida,
Tampa, Florida 33620, USA
killinge@cas.usf.edu*

We have conducted studies of deep UV laser-induced fluorescence (LIF) for the reagentless detection of trace species and Dissolved Organic Compounds (DOC's) in water. Our LIF detection system had two interchangeable UV lasers, 266 nm and 355 nm, illuminating a flow cell containing a water sample. The fluorescence emitted at 90 degrees to the laser beam was collected by focusing optics, passed through cut-off and interference filters with 21 optical bandpass channels (240–680 nm), and detected by a photomultiplier tube (PMT). The samples analyzed by the system included bottled, tap and river water; we have also worked with biological and chemical species (*Bacillus Globigii*, malathion). In terms of the excitation wavelength, it was observed that the deep UV excitation resulted in spectra that contained more features, and had better separation of the LIF from the Raman peak, thus enhancing the detection of unique spectral features.

Keywords: UV; LIF; laser-induced fluorescence; water.

1. Introduction

Laser-induced fluorescence is a sensitive, reagentless method of monitoring water quality that can be used to detect trace amounts of organic species. The choice of the laser excitation wavelength plays an important role in obtaining good LIF spectra, and therefore reliable detection of trace organics. As our research has shown, using shorter, deep-UV wavelengths can be beneficial in that the resultant fluorescence spectrum emitted is separated from other interfering emission processes (Raman, Mie).¹⁻³

^a University of South Florida, 4202 East Fowler Ave, PHY 114, Tampa, FL 33620, USA.

2. Experimental Apparatus

Our LIF system (Fig. 1) used two interchangeable microchip Nd:YAG lasers (JDS Uniphase) working at 266 nm and 355 nm (fourth and third harmonic) as UV sources. A tiny fraction of UV radiation went to the triggering detector (New Focus Inc., Model 1621). The greater portion of the laser light illuminated the sample inside the quartz flow cell (Spectrocell Inc., Model RF-3010-F). Multiple passage of excitation radiation, implemented by placing two plane mirrors at a slight angle before and after the cell, increased the total fluorescence. The fluorescent light was collected at an angle of 90 degrees with respect to the laser radiation. A concave mirror placed on the opposite side of the sample cell, as well as two quartz lenses used for collimation and focusing, increased the light collection efficiency. The five cut-off and 19 interference filters were mounted on filter wheels (CVI Laser, Models AB-302 and AB-304) which selected the range of wavelengths passing to the PMT detector (Hamamatsu, Model H6780-03); 21 optical bandpass channels were used to cover the range of 240–680 nm. The PMT signal was processed by the boxcar averager and gated integrator system (SRS, Model SR 250), and went through the computer interface module (SRS, Model SR 245) to a notebook computer. The water sample was contained in the 1cm×1cm×5cm quartz flow cell. The sample could be circulated through the cell using an external peristaltic pump or be introduced in a static manner.

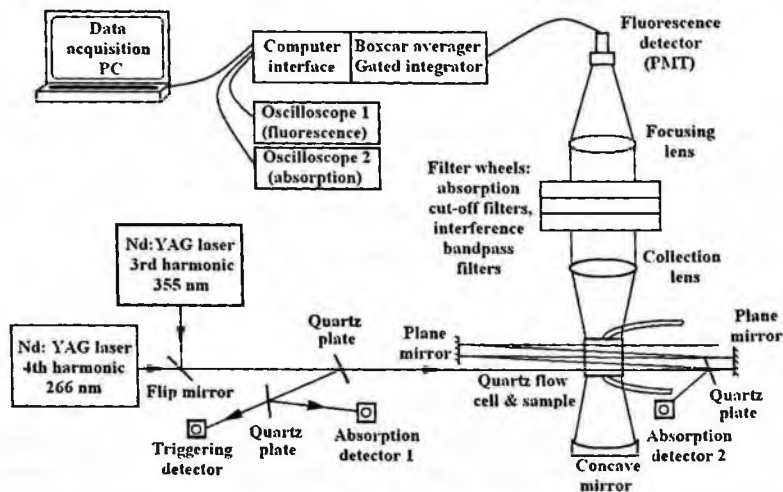


Figure 1. Diagram of LIF system.

The entire system was contained in two plastic cases. Fig. 2 shows the case holding the optical equipment; the electronics box is shown in Fig. 3.

The system response was found to be linear from 0.05 to 20 ppb for quinine sulfate in sulphuric acid, and sensitivity limit for the same analyte approaching ten parts per trillion.¹

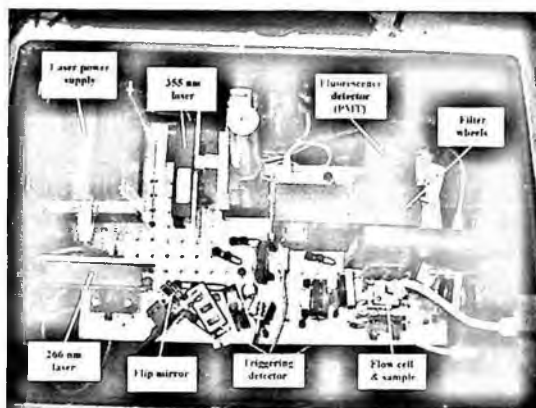


Figure 2. Photograph of LIF system optics.

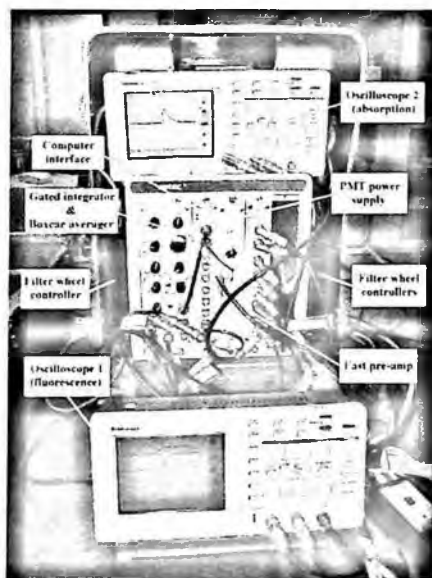


Figure 3. Photograph of LIF system electronics.

3. LIF Data of Water Samples

The LIF system was used to measure different water samples by placing the sample in the quartz cell of the apparatus. The goal was to obtain fluorescence patterns typical for the ground and drinking water.

Figs. 4–5 show the difference in water fluorescence excited by 266 nm and 355 nm laser radiation. The large fluorescence emission near 450 nm is due to Dissolved Organic Compounds (DOC) in the water.^{1,4} The tap and ground water were from the University of South Florida water supply, and the distilled water were commercial bottled water brands

sold by local grocery store. As can be seen, the signal was strongest for the ground water sample, and weakest for the distilled one. An extended set of water samples have been analyzed and reported previously.²

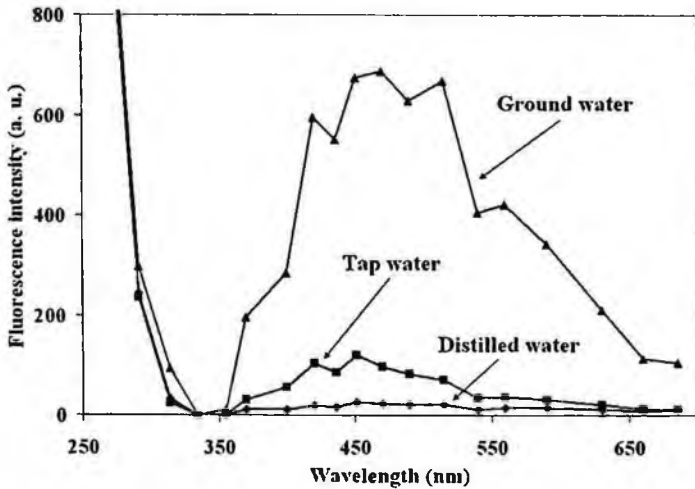


Figure 4. Distilled, tap and ground water fluorescence excited by 266 nm laser.

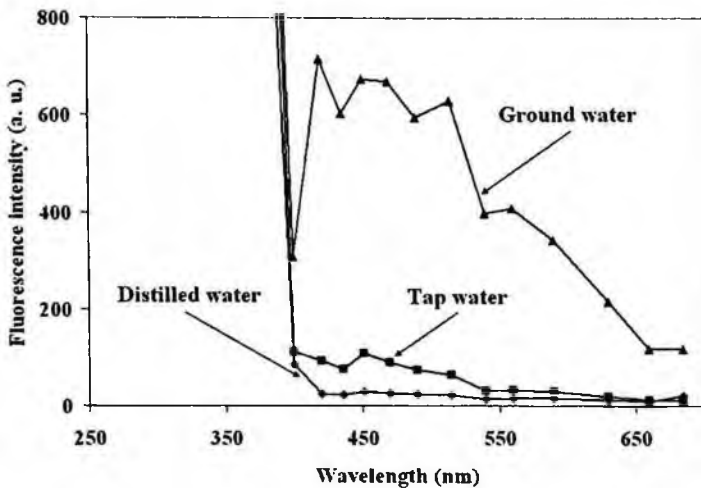


Figure 5. Distilled, tap and ground water fluorescence excited by 355 nm laser.

For the fluorescence spectra shown, the typical data acquisition time per interference filter was set to 1.5 seconds. Each data point represented the average of 300 signal pulses. The complete spectrum collection for both excitation wavelengths, including the time for filter switching, usually took two to three minutes.

4. LIF Data of *Bacillus Globigii* (BG) Endospores and VX Simulant

The LIF system was used to perform initial measurements on other interesting compounds, including BG endospores and an insecticide, malathion. Figs. 6–7 show the signal generated by BG (anthrax simulant) in water for the two laser wavelengths. For high concentration of 3.5 grams per liter, we observed fluorescence of the spores without terbium doping typical in such measurements.⁵ As can be seen, the signal had a broad peak centered around 450 nm, similar to the fluorescence of the DOC's.

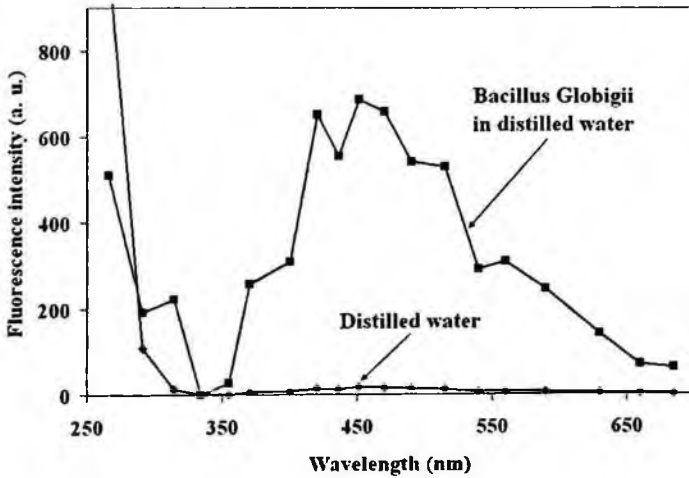


Figure 6. Distilled water and BG fluorescence excited by 266 nm laser.

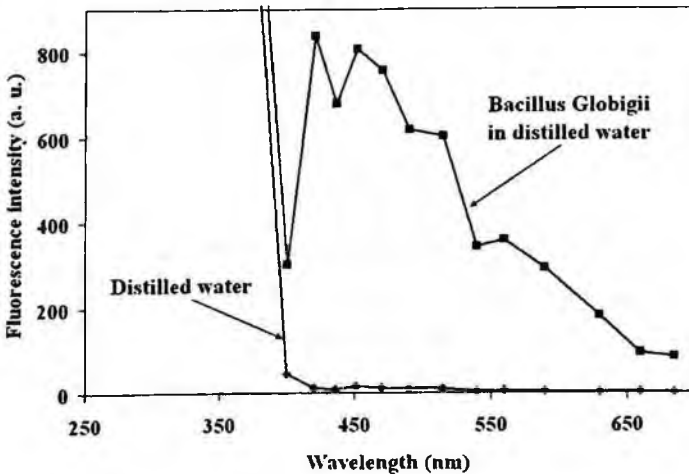


Figure 7. Distilled water and BG fluorescence excited by 355 nm laser.

A sample of the insecticide malathion (VX simulant) was obtained by diluting 4 mL of Dexol in 1 L of distilled water. The spectra of its fluorescence, excited by 266 nm and 355 nm ultra-violet lasers, are shown in Figs. 8–9. As can be seen, there appears to be a distinct fluorescence emission near 360 nm. Further work is being conducted to better quantify our results.

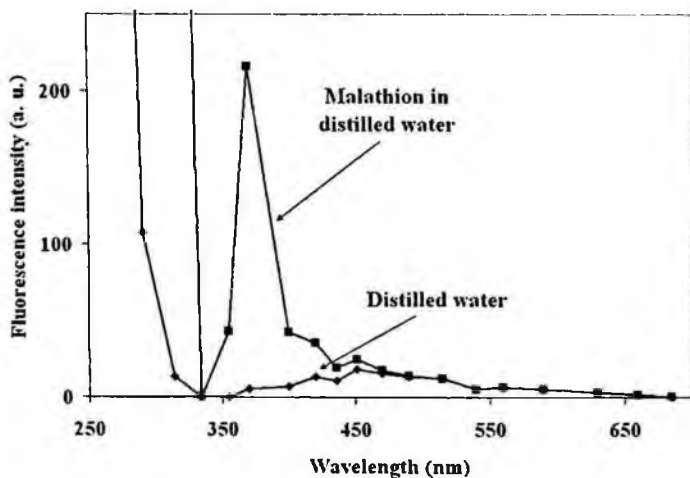


Figure 8. Distilled water and malathion fluorescence excited by 266 nm laser.

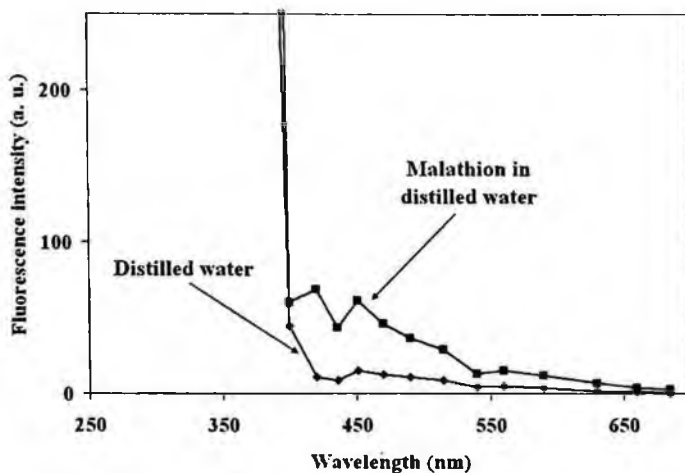


Figure 9. Distilled water and malathion fluorescence excited by 355 nm laser.

5. Analysis and Conclusions

From a comparison of Figs. 4–9 it is clear that the shorter excitation UV wavelength leads to more spectral features in the fluorescence, in part due to the shift of the Raman peak. The peak fluorescence intensity, however, is typically comparable for the two wavelengths. Our research also indicated that using the deeper UV lasers for water LIF studies results in more distinct spectra with quantitative features.

Currently, we are investigating the feasibility of using deep UV LEDs as an alternative for the laser sources and applying our system to the real-time monitoring of reverse osmosis processed water.

6. Acknowledgments

The material is based upon work supported by the US Army Research Office under contract/grant number W911NF-05-1-0431.

References

1. V. Sivaprakasam and D. K. Killinger, Tunable ultraviolet laser-induced fluorescence detection of trace plastics and dissolved organic compounds in water, *Appl. Opt.* **42**, 6739–6746 (2003).
2. A.V. Sharikova and D. K. Killinger, Laser-induced fluorescence studied of water processed by a reverse osmosis purification unit, *Proc. of SPIE* **5994**, 59940B-1–59940B-5 (2005).
3. V. Sivaprakasam and D. K. Killinger, Effect of polarization and geometrical factors on quantitative laser induced fluorescence-to-Raman intensity ratios of water samples and a new calibration technique, *J. Opt. Soc. Am. B* **20**, 1980 (2003).
4. P. G. Coble, Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy, *Mar. Chem.* **51**, 325–346 (1996).
5. D. L. Rosen, C. Sharpless and L. B. McGown, Bacterial Spore Detection and Determination by Use of Terbium Dipicolinate Photoluminescence, *Anal. Chem.* **69**, 1082–1085 (1997).

DEVELOPMENT OF MATERIALS AND SAMPLING METHODS FOR IR-BASED DETECTION OF TOXIC COMPOUNDS IN WATER

BRIAN J. NINNESS¹, LUKE D. DOUCETTE¹, BEN MCCOOL², CARL P. TRIPP^{1,2}

¹ *Orono Spectral Solutions, 20 Godfrey Drive Orono, ME 04473*
Brian.Ninness@umit.maine.edu

² *Laboratory for Surface Science & Technology, University of Maine, Orono, ME 04469*

A detection approach based on the principles of Fourier Transform Infrared Spectroscopy (FTIR) is presented for the trace level detection of toxic compounds in water. The main advantages of this approach are that it operates in heterogeneous aqueous environments, provides fast detection (< 10 min), and exhibits high sensitivity/selectivity to nonvolatile toxic materials with minimal false alarms. The key enablers to using FTIR for aqueous-based detection is the development of a selective and robust sampling protocol coupled to a miniaturized portable FTIR unit. The sampling approaches involve synthesizing and tailoring microporous, mesoporous, and nonporous metal oxide powders/films that are amenable for *in situ* FTIR measurements. In this paper we provide an overview of the material synthesis and surface modification strategies, and present results obtained using these materials for the low level detection of the organophosphate pesticide phosmet. Phosmet is used as a surrogate for the nerve agent VX.

Keywords: FTIR; silica; phosmet; detection.

1. Introduction

The Joint Service Agent Water Monitor (JSAWM) is a funded Department of Defense program that is intended to identify, mature and develop the technology needed for specific and sensitive detection of chemical warfare agents (CWA) in source, treated, stored and distributed water supplies. According to the JSAWM criteria, a successful detection system will function in heterogeneous liquid environments, have less than a 10 minute response time with low false alarms, work in both continuous (on-line monitoring) and batch modes, be portable, small, easy to operate, and have very low power requirements.¹

The mid-infrared region (4000 – 400 cm⁻¹) of the electromagnetic spectrum finds widespread use in both qualitative and quantitative analysis. Spectra of organic molecules give rise to a unique fingerprint that is compound specific. Fourier Transform Infrared Spectroscopy (FTIR) therefore provides the inherent selectivity needed for a system designed for the detection and identification of toxic compounds in water. The main hurdle for an FTIR detection system of compounds in water is the lack of sensitivity at the parts-per-billion (ppb) level. For example the detection limit for the nerve agent VX is 5ppb (based upon a daily water consumption rate of 15 L/day), according to the current U.S. Tri-Service Water Quality Standards.²

We have developed several approaches to silanize oxide materials from the gas phase,⁶⁻¹² solution^{6,8,9,12-15} and supercritical CO₂.^{16,17} The preferred method of synthesis is to pass a gaseous stream of an organosilane at high temperatures (>300°C) over the silica.¹⁸ However, this gas phase treatment is limited to volatile organosilanes. A more common approach is to disperse the oxide powder into a non-aqueous liquid solution containing the organosilane at temperatures below 100°C. At these lower solution temperatures, we have shown that there is no direct reaction of an organosilane with the surface hydroxyl groups.^{6,13} Water (present either adsorbed upon the surface or in trace amounts in the non-aqueous liquid) is needed to convert the chloro or alkoxy organosilane to an organosilanol which, in turn, reacts with surface silanol groups.

The main drawback of using silanes to modify the surface of metal oxides is that it is plagued by a lack of day-to-day reproducibility. This variability is attributed to the self-polymerization of the organosilane giving rise to an ill-defined, polymerized silane coating. The competition between a surface reaction and self-polymerization is dictated by several factors including the nature of the organosilane,^{19,20} the choice of solvent,²¹ reaction temperature²² and in particular, to the level of water adsorbed on the surface.^{14,23-26} In addition to variability issues, polymerization with adsorbed water is a source for low silane density as the polymers can plug access to the inner pore structure.

One method we have developed for avoiding polymerization with the adsorbed water layer is to use a base catalyzed reaction in supercritical CO₂ (SCF CO₂). We have shown that SCF CO₂ has the unique property of removing all adsorbed water from the silica surface. Once removed the amine base catalyses the reaction of the silane with the surface.⁷ As shown in Figure 1, the preadsorbed amine hydrogen bonds to the surface silanol and renders the Si-O group of the silanol more nucleophilic for reaction with the silicon atom of the incoming chloro or alkoxy silane. Triethylamine is the preferred amine as it binds to all isolated silanols from SCF CO₂ and is not removed with subsequent exposure to flowing SCF CO₂ solvent. Furthermore a tertiary amine should be used as primary and secondary amines such as ammonia are known to react with CO₂ to produce carbamates.²⁷⁻²⁹

A second hurdle is that direct silanation of mesoporous silica using conventional solution based protocols has lead to poor levels of silane incorporation.^{4,30} In this respect, SCF CO₂ has an important characteristic of high diffusivity and low viscosity and therefore can be used as a carrier to bring reagents into the pore structure of the oxide.³¹ With these attributes in mind, we have used SCF CO₂ to modify the surface chemistry of the synthesized mesoporous materials. Specific surface modifications applied for the detection of organophosphate pesticides have included the use of hexamethyldisilazane (CH₃), octadecyldimethylchlorosilane (C-18), and trifluoropropyldimethylchlorosilane (CH₂CH₂CF₃). The results shown in this paper are for a methylated (CH₃) surface.

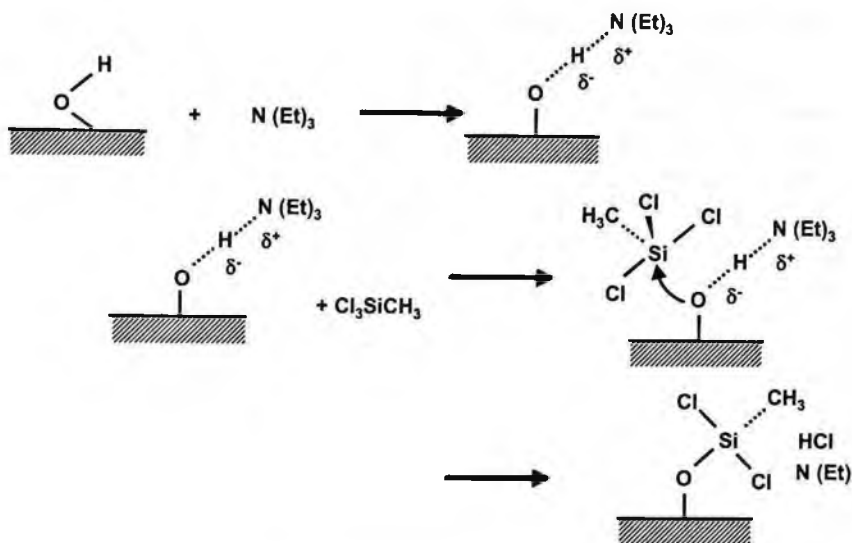


Figure 1. Two step base catalyzed reaction.

3.1 Experimental

The substrates with the mesoporous silica films were loaded into a 25 cc heated reaction chamber. The silica film is first purged with SC-CO₂ at 50 ° C and 200 atmospheres for a period of 1 hour to remove all adsorbed water. The SC-CO₂ is also delivered to a 10 cc delivery vessel where the hexamethyldisilazane (HMDS) is dissolved. After the purge period the HMDS is delivered to the reaction chamber with a flow rate of 0.5 mL/min. After the HMDS exposure, the reaction vessel is closed off from the HMDS delivery vessel and purged with pure CO₂ for 5 minutes. This process was repeated for a total of three exposures.

4. Development of Batch & Continuous Sampling Methods

One difficulty to using infrared spectroscopy arises from the fact that the collection of infrared spectra is very much dependent on sample properties (optical, mechanical, morphological, etc.). For this reason infrared spectroscopy is notorious for its many sampling methods and accessories. Therefore the sampling procedure is a critical element in the development of an IR-based detection system. For example, a direct measurement of organophosphate pesticides in an aqueous solution by simple IR transmission is difficult because of the strong absorption of infrared radiation by the solvent. In the sampling methods discussed in the next section absorption of the infrared light by the solvent is avoided because the mesoporous films provide a means of the concentrating the target species and separating them from the solvent.

4.1. *Batch detection of VX surrogate phosmet.*

In order to demonstrate the ability of functionalized, mesoporous adsorbents to concentrate and present for identification via FTIR, we performed a thorough investigation of the detection of the VX surrogate phosmet (pesticide). Figure 2 represents a typical infrared spectrum of adsorbed phosmet on a coated Si wafer obtained after exposure to a 2 ppm solution (10 minute exposure). Included in this figure is a reference spectrum of pure phosmet in a KBr matrix. VX and phosmet are organo-phosphorous compounds and both are acetylcholinesterase inhibitors. In addition to their chemical similarities phosmet and VX have identical extinction coefficients ($4.4 \times 10^{-4} \text{ m}^2/\text{mg}$) for the strongest absorption band in each of their respective mid-IR spectrum.

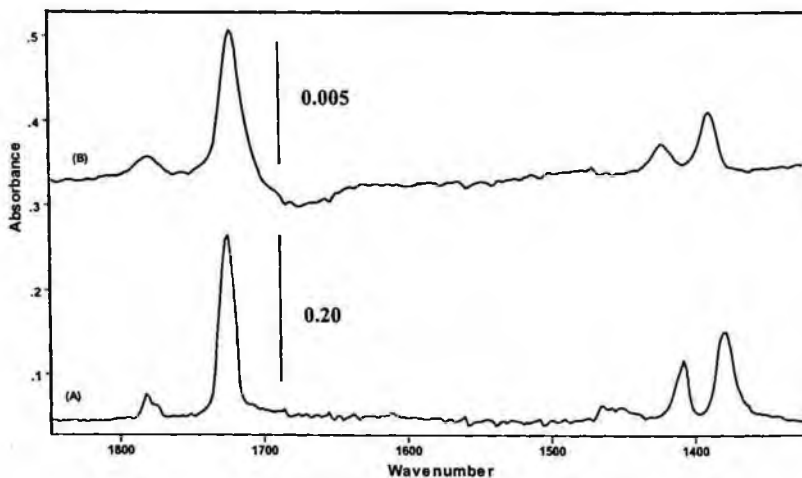


Figure 2. (A) Spectra of pure phosmet (KBr Matrix) & (B) Phosmet adsorbed to HMDS Modified Silica Transmission substrate.

4.2. *Continuous detection of VX surrogate phosmet*

One FTIR sampling technique that has been postulated as a means of conducting continuous monitoring of water streams is Attenuated Total Reflection (ATR). With ATR the evanescent field produced at the surface of an internal reflection element only penetrates approximately 1 – 5 μm into the bulk water stream. Therefore one is able to acquire IR spectra of surface species in the presence of the water. The spectra shown in Figure 3 shows that this is indeed possible. In this case, the silica coated Ge ATR crystal was in contact with a 5 ppm phosmet solution. Spectra were recorded in situ with the solution flowing across the crystal at 4 minute intervals. Clearly the presence of water does not impede the identification of phosmet in this continuous monitoring approach.

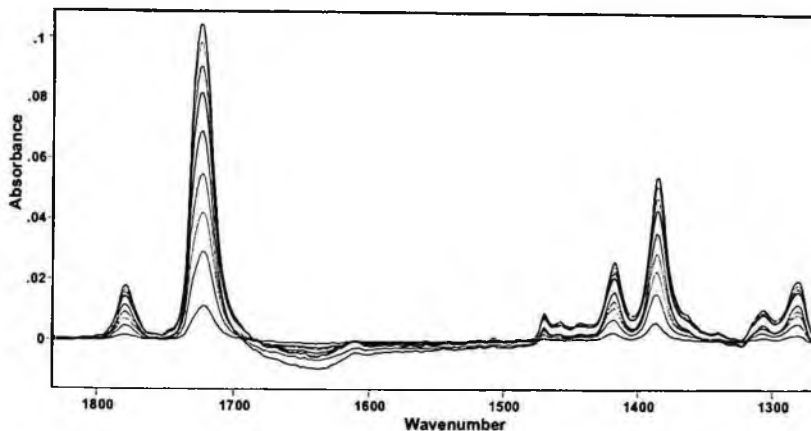


Figure 3. ATR spectra of phosmet, 5 minute intervals.

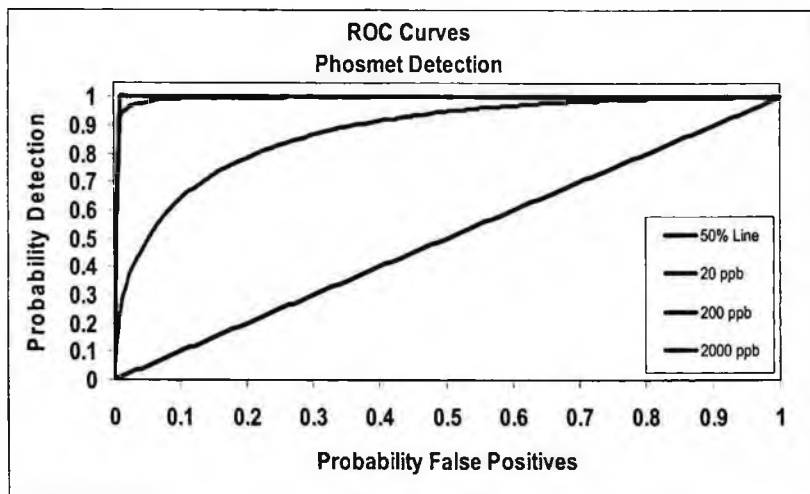
4.3 Detection sensitivity and reproducibility

As mentioned earlier, the key hurdle is to achieve detection sensitivity at the low ppb level. The intensity of the bands in absorbance units (ordinate scale) is proportional to the amount of phosmet probed by the infrared beam. The spectrum of phosmet adsorbed on the Si wafer in Figure 2 shows that detection at 2 ppm is easily accomplished. In order to determine the lower detection limit, a statistical design approach using Receiver Operator Curves (ROC) is required. To generate a ROC curve, a total of thirty samples were tested over the course of five days. For each day of testing, the same stock solution of 0.25% phosmet in methanol was used to prepare serial dilutions in DI water over 5 orders of magnitude (20,000 ppb, 2,000 ppb, 200 ppb, 20 ppb, 2 ppb and BLANK). The samples were exposed to a 250 ml volume of the solution for 10 minutes. The order of sampling was varied over the course of the 5 days from low – high and high-low.

Table 1 demonstrates the experimental matrix which was followed, and the plot below (Figure 4) shows the ROC curves that were generated from these experiments. The carbonyl peak centered at 1720 cm^{-1} is used as the metric for phosmet detection. For clarity, 20,000 ppb is omitted since it overlays 2,000 ppb; 2 ppb is not shown since this data was not statistically different than the blank. The separation factors, or ROC K-factors, for the curves shown were 6.75, 3.75, and 1.63 for 2,000 ppb, 200 ppb, and 20 ppb respectively. The minimum value of K that meets the Joint Service Agent Water Monitor (JSAWM) requirement of 95% detection at 5% false positives is 3.29. Therefore, these test results clearly demonstrate the detection of a VX surrogate at 200 ppb using these modified mesoporous adsorbents in a simple batch sampling mode.

Table 1: Phosmet Test Matrix

	2 ppb	20 ppb	200 ppb	2 ppm	20 ppm	BLANK
Day 1	A1	B1	C1	D1	E1	F1
Day 2	F2	E2	D2	C2	B2	A2
Day 3	A3	B3	C3	D3	E3	F3
Day 4	F4	E4	D4	C4	B4	A4
Day 5	A5	B5	C5	D5	E5	F5

**Figure 4.** ROC curves for Phosmet Detection in DI water background.

5. Pathway To Achieve Lower Detection Limits

The inability to lower the detection limits from 200 ppb to the single ppb range when using coated ATR crystals or transmission disks in contact with aqueous solutions is due to two factors; (1) equilibrium partitioning and (2) mass transport limitations. The first factor is due to the increase in solubility of phosmet (or VX) with decreasing concentration which, in turn, leads to a lower partitioning of the target compound on the hydrophobic adsorbent. This problem can be circumvented by using a surface chemistry which leads to formation of a strong chemical bond with the target molecule. However this approach does not circumvent the mass transport limitation. The problem with the mass transport limitation is that as the concentration is lowered the incubation time required to reach maximum adsorption exceeds the JSAWM 10 minute detection requirement. This is clearly seen in the ATR spectrum shown in Figure 2. The intensity of the phosmet bands increase with contact time and the maximum in intensity is achieved after 2 hrs contact.

A simple approach to overcome both limitations can be found in the solid phase extraction (SPE) sampling process. SPE is a conventional method of extracting organic

substances from aqueous solutions. It has become popular due to the simplicity of the method, the reduced use of solvents as compared to liquid-liquid extraction, and its low overall costs. After extraction of the target compounds the traditional technique uses solvent extraction followed by solvent evaporation. One of the major requirements for the JSAWM program is that any detection system must be reagentless, due to the logistical burdens of supplying and stocking the requisite consumables that may be needed in the field. Therefore the purpose here was to extract the phosmet from an aqueous solution using SPE, followed by directly analyzing the sorbent material.

The following simple experiment demonstrates how this extraction process, adapted to our materials, may provide the pathway to the needed detection limits. The same solution volume (250 ml) used to generate the ROC curves for phosmet, as shown in Figure 3, was passed through an SPE cartridge containing 50 mg of hydrophobic silica at a total concentration of 5 ppb. The adsorbent was subsequently removed from the cartridge, dried and scanned using a typical bench-top FTIR using diffuse reflectance (DRIFT). The resulting spectrum of phosmet is shown in Figure 5. The carbonyl peak at 1720 cm^{-1} is used as the metric for phosmet detection. Therefore by using this sampling approach it has been shown that detection limits as low as 5 ppb are indeed possible using FTIR.

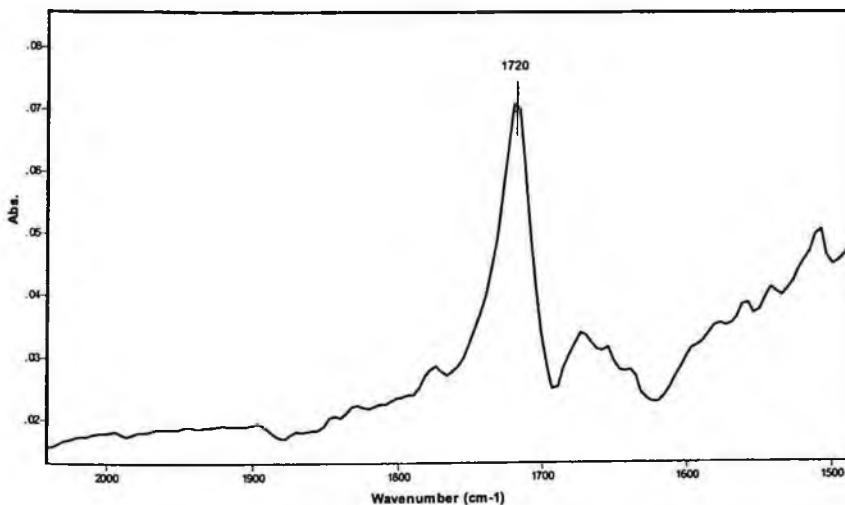


Figure 5. 5 ppb detection of Phosmet using SPE sampling.

The SPE sampling process effectively addresses both factors that limit the effective detection limits of the batch sampling techniques. The mass transport limitations are overcome by ensuring that the entire solution volume comes into intimate contact with the adsorbent bed, and the single equilibrium point of batch sampling is overcome by effectively increasing the number of theoretical plates. The obvious

limitation of the standard SPE adsorbent material for simple infrared transmission analysis is the large particle size. In order to achieve uniform and rapid flow through the packed bed the particle size of the silica is typically around 40 μm . This particle size causes intense scattering of the IR radiation when doing simple transmission scanning, hence the requirement of using DRIFT for the analysis.

6. Conclusions

High surface silica adsorbent films have been designed to concentrate toxic compounds in water and to be amenable for direct analysis by an IR- based detection system. This approach has been shown to provide high detection selectivity but the films coated on IR transparent substrates are limited by mass diffusion and equilibrium partitioning to a detection limit of 200 ppb of phosmet in water. In order to achieve single digit ppb detection the high surface areas films need to be fabricated on a porous substrate that will mimic the extraction efficiencies of SPE while providing an analysis platform for simple FTIR transmission. Work in developing these supports is currently underway.

7. Acknowledgements

We gratefully acknowledge the assistance and guidance from Janet Jensen and Dr. James Jensen of the Edgewood Chemical and Biological Center, Aberdeen Proving Grounds, MD. This work was funded by the Department of Defense – Army, Contract # W911SR-04-C-0029.

References

1. Private conversation: Janet Jensen, Edgewood Chemical Biological Center, Aberdeen Proving Ground, MD.
2. U.S. Tri-Service Water Quality Standards, U.S. Army Research Office, Durham, N.C.
3. B.A. McCool, N. Hill, J. DiCarlo and W.J. DeSisto, Synthesis and characterization of mesoporous silica membranes via dip-coating and hydrothermal deposition techniques, *J. Memb. Sci.*, **218**, 55-67 (2003).
4. G.E. Fryxell and J. Liu, *Designing Surface Chemistry in Mesoporous Silica*, in Adsorption on Silica Surfaces, E. Papirer, Editor (Marcel Dekker, New York, 2000).
5. R.P.N. Veregin, C.P. Tripp, M.N.V. McDougall and D. Osmond, The Role of Water in the Triboelectric Charging of Alkylchlorosilane-Treated Silicas as Toner Surface Additives, *J. Img. Sci. Tech.* **39**, p. 429 (1995).
6. C.P. Tripp and M.L. Hair, The reaction of alkylchlorosilanes with silica at the solid/liquid and solid/gas interface, *Langmuir*, **7**, p. 1961 (1992).
7. C.P. Tripp and M.L. Hair, The chemical attachment of chlorosilanes to silica: a two step amine promoted reaction, *J. Phys. Chem.* **97**, 5693-5698 (1993).
8. C.P. Tripp, R.P.N. Veregin and M.L. Hair, The effect of fluoro-alkyl substituents on the reaction of alkylchlorosilanes with silica surfaces, *Langmuir*, **9**, 3518-3522 (1993).

9. C.P. Tripp and M.L. Hair, Reaction of methylsilanols with hydrated silica surfaces: the hydrolysis of trichloro-, dichloro- and monochloro-methylsilanes and the effects of curing, *Langmuir*, **11**, 149-155 (1995).
10. L.D. White and C.P. Tripp, A low-frequency infrared study of the reaction of methoxymethylsilanes with silica, *J. Colloid Int. Sci.*, **224**, 417-424 (2000).
11. L.D. White and C.P. Tripp, An infrared study of the amine catalyzed reaction of methoxymethylsilanes with silica, *J. Colloid Int. Sci.*, **227**, p.237 (2000).
12. K.M.R. Kallury, M. Thompson, C.P. Tripp and M.L. Hair, Interaction of silicon surfaces silanized with octadecanoic acid and octadecanamine studied by ellipsometry, XPS and FTIR, *Langmuir*, **8**, p.947 (1992).
13. C.P. Tripp and M.L. Hair, The reaction of octadecyltrichlorosilane with silica, *Langmuir*, **8**, p.1120 (1992).
14. C.P. Tripp and M.L. Hair, Direct observation of surface bonds between self-assembled monolayers of octadecyltrichlorosilane and silica surfaces: a low frequency IR study at the solid/liquid interface, *Langmuir*, **11**, 1215-1219 (1995).
15. B.J. Ninness, D.W. Bousfield and C.P. Tripp, An in situ infrared technique for studying adsorption onto particulate silica surfaces from aqueous solutions, *App. Spec.*, **55**, p.655 (2001).
16. C.P. Tripp and J.R. Combes, The interaction of supercritical CO₂ with the adsorbed water and surface hydroxyl groups on silica, *Langmuir*, **14**, p.7348 (1998).
17. J.R. Combes, L.D. White and C.P. Tripp, Chemical modification of metal oxide surfaces in supercritical CO₂: in situ infrared studies of the adsorption and reaction of organosilanes on silica, *Langmuir*, **15**, 7870-7875 (1999).
18. C.P. Tripp and M.L. Hair, The reaction of chloromethylsilanes with silica: a low frequency infrared study, *Langmuir*, **7**, p.923 (1991).
19. S.M. Kanan, W.T.Y. Tze and C.P. Tripp, A method to double the surface concentration and control the orientation of adsorbed (3-aminopropyl)dimethylethoxysilane on silica powders and glass slides, *Langmuir*, **18**, p.6623 (2002).
20. S.M. Kanan, C.P. Tripp, R.N. Austin and H.H. Patterson, Photoluminescence and raman spectroscopy as probes to investigate silver and gold dicyanide clusters doped in A-type zeolite and their photoassisted degradation of Carbaryl, *J. Phys. Chem. B*, **105**, p.9441 (2001).
21. M.E. McGovern, K.M.R. Kallury and M. Thompson, Role of solvent on the silanization of glass with octadecyltrichlorosilane, *Langmuir*, **10**, 3607-3614 (1994).
22. C.G. Golander and E. Kiss, Quantitative measurements of amino silylation on plasma-activated mica surface, *Colloids and Surfaces*, **74**, 217-222 (1993).
23. A.N. Parikh, D.L. Allara and F. Rondelez, The role of interfacial water in grafted and non-grafted octadecylsiloxane monolayers on solid substrates, *Nature*, (1994).
24. D.L. Angst and G.W. Simmons, Moisture absorption characteristics of organosiloxane self-assembled monolayers, *Langmuir*, **7**, 2236-2242 (1991).
25. P. Silberzan, D. Ausserr and J.J. Benattar, Silanation of silica surfaces, a new method of constructing pure or mixed monolayers, *Langmuir*, **7**, 1647-1651 (1991).
26. J.D. LeGrange, J.L. Markham and C.R. Kurkijan, Effects of surface hydration on the deposition of silane monolayers on silica, *Langmuir*, **9**, 1749-1753 (1993).
27. K.P. Battjes, A.M. Barolo and P. Dreyfuss, New evidence related to reaction of aminated silane coupling agents with carbon dioxide, *J. Adh. Sci. and Tech.*, **5**, 785-799 (1991).

28. N. Wen and M.H. Brooker, Rate constants for cyanate hydrolysis to urea, a raman study, *Can. J. Chem.*, **72**, 1099-1106 (1992).
29. B.R. Ramachandran, A.M. Halpern and E.D. Glendening, Kinetics and mechanism of the reversible dissociation of ammonium carbamate, *J. Phys. Chem. A.*, **102**, 3934-3941 (1998).
30. C.P. Jaroniec, M. Kruk, M. Jaroniec and A. Sayari, Tailoring surface and structural properties of MCM-41 silicas by bonding organosilanes, *J. Phys. Chem. B.*, **102**, 5503-5510 (1998).
31. T.S. Zemanian, G.E. Fryxell, J. Liu, S. Mattigod, J.A. Franz and Z. Nie, Deposition of self-assembled monolayers in mesoporous silica from supercritical fluids, *Langmuir*, **17**, 8172-8177 (2001).

THz SPECTROSCOPY OF PROTEINS IN WATER: DIRECT ABSORPTION AND CIRCULAR DICHROISM

JING XU

Department of Physics, UC Santa Barbara, Santa Barbara, CA 93106

KEVIN W. PLAXCO

Department of Chemistry and Biochemistry, UC Santa Barbara, Santa Barbara, CA 93106

S. JAMES ALLEN

Department of Physics, UC Santa Barbara, Santa Barbara, CA 93106

Many of the functionally relevant collective vibrations of proteins and other biopolymers are expected to occur at terahertz frequencies. Precise absorption measurements combined with careful titration of biopolymers in water have allowed us to directly measure the terahertz absorption spectra associated with these motions, despite the strong background absorption of the solvent. We have also explored the circular dichroism spectroscopy of biomolecules over this same frequency range. Since circular dichroism requires the presence of net chirality in a molecule and chirality is present in nearly all biomaterial, it has the potential to capture the background free spectral features in biopolymers. To undertake these studies we have developed a broad band terahertz spectrometer suitable for both direct absorption and circular dichroism measurements of proteins in water between 0.75 – 3.72 THz. Direct terahertz absorption spectra of prototypical proteins bovine serum albumin (BSA) and hen egg white lysozyme have been documented and are described here. We have also successfully demonstrated the magnetic circular dichroism in semiconductors, and placed an upper bound on the terahertz circular dichroism signature of solvated BSA. In the terahertz frequency range, it appears that circular dichroism signatures are exceedingly small and detection remains a challenge.

Keywords: biopolymer, far infrared, free electron laser, collective modes, sub-millimeter wave, terahertz

1. Introduction

Biological polymers are expected to exhibit broad spectral features in the terahertz regime, corresponding to functionally relevant, global and sub-global vibrational modes with periods on the picosecond timescale. The ability of biomolecules to deform is critical to their participation in biological activity, and mutations that alter the conformational rigidity of proteins and nucleic acids can lead to loss of function^{1,2}. Theoretical studies paint a picture consistent with the importance of the inter-domain collective dynamics in biopolymer structure and function and suggest the existence of a large number of terahertz collective modes in biomaterials³⁻¹³.

Terahertz absorption spectroscopy directly probes such collective vibrational modes, and provides valuable experimental data to test the validity of these theoretical calculations. Previously we have reported the first measurements of the terahertz

absorption spectra of prototypical proteins -bovine serum albumin (BSA) and hen egg white lysozyme- in aqueous solution^{14,15}. By precisely measuring of terahertz absorption spectrum of concentrated protein solutions and careful assessing the solvent background, we have successfully extracted the terahertz absorption of proteins in their natural water environment despite overwhelming terahertz absorption from water.

While, with sufficient care, the direct terahertz absorption spectra of the protein can be extracted from the large absorbance background of aqueous solutions, it is clear that approaches that discriminate against non-biological absorption are desired. Terahertz circular dichroism may present such a unique potential. Circular dichroism, the differential absorption between left and right handed circular polarized radiation appears only in the presence of chiral symmetry. Since chirality almost universally and uniquely a feature of biological material (and since the collective vibrational modes of a folded, chiral biomolecule are themselves fundamentally chiral), terahertz circular dichroism spectrum can be uniquely associated with the biopolymer component of a mixture. In addition, terahertz spectral features are directly related to charge distributions and large-scale macromolecular conformations and low frequency collective motions, thus the corresponding terahertz circular dichroism spectra will relate to unique and distinguishable biomolecular structures and be potentially very useful for biochemical "fingerprinting"^{16,17}.

Circular dichroism has been well understood and established in the ultra-violet, visible, near- and mid-infrared frequency range as an integral part of contemporary biophysics with numerous, excellent, turnkey instruments commercially available¹⁸. However, there have been no documented measurements of terahertz circular dichroism. We have developed a simple physical model of circular dichroism in the terahertz frequency range, and build upon our broad band absorption spectrometer to explore the terahertz circular dichroism signatures of prototypical proteins in aqueous solution.

2. Terahertz Direct Absorption Spectroscopy

We have developed a 0.075 - 3.72 THz absorption spectrometer suitable for studying strongly attenuating media such as proteins in aqueous solution. Radiation is provided by harmonic multipliers (up to 210 GHz) and the UCSB free-electron lasers (up to 4.8 THz). The spectrometer combines these sources with a cryogenic detector and a variable path length sample cell to precisely extract the terahertz extinction of protein solutions.

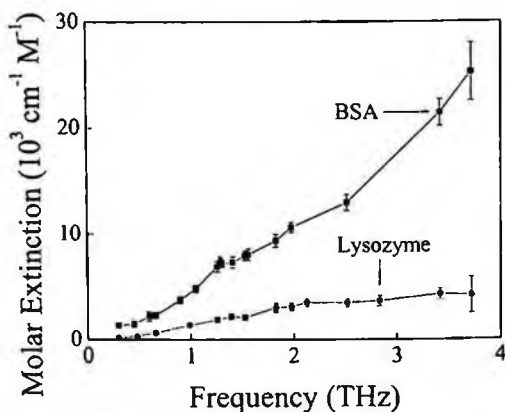


Figure 1. Molar extinction of solvated BSA and lysozyme.

For our study, we have employed the proteins bovine serum albumin (BSA) and chicken hen egg white lysozyme (HEWL, or lysozyme) as the representative biomolecules. BSA is a monomeric protein of molecular weight 66.4 kDa and shares 76% sequence homology with human serum albumin (protein data bank file code 1AO6, Figure 1.3.2.A)¹⁹. Lysozyme is a small monomeric protein with molecular weight 14.3 kDa and PDB file codes 1LYZ and 1BWH. Both BSA (Cohn Fraction V, lyophilized powder) and lysozyme (lyophilized powder) were obtained from Sigma (St. Louis, MO) and used without further purification.

To overcome the dominating terahertz absorption of water, we employ intense radiation sources (up to 10 kW) and sensitive detectors ($\sim \text{pW/Hz}^{1/2}$) to precisely measure the extinctions of both protein solutions and their associated buffer blanks. The terahertz absorption measurements are dominated by water absorption. We then carefully assess the amount of bound water in the protein's hydration shell, allowing us to obtain an accurate estimation of the dominating solvent background and extract the molar extinction of the solvated protein^{14,15} (Figure 1).

3. Terahertz Circular Dichroism Spectroscopy

Terahertz circular dichroism offers an opportunity to explore the biologically relevant collective terahertz dynamics undeterred by presence of non-biological medium.

3.1. Simple Physical Model

We have developed a simple physical model suggesting that for most, if not all, chiral macromolecules, their terahertz absorption will be accompanied by terahertz circular dichroism. *Ab initio* methods have been used to successfully calculate the vibrational circular dichroism spectra of small molecules, however, these models have largely focused on spectra in the near-infrared, spectra dominated by bond stretching and bending, and the wagging of small subunits in a macromolecule²⁰⁻²³. The direct calculation of circular dichroism features arising from the collective vibrational modes of macromolecules is computationally difficult. In an effort to develop our intuition about terahertz circular dichroism, we resort to a simple physical model to roughly approximate the α -helical secondary structure common among proteins, and the relevant molecular motions.

Given the prevalence of helical substructures in terrestrial biopolymers such as proteins, DNA and RNA, we are compelled to explore a “mass and spring” model of a helical molecular structure as a simple physical representation that captures the essential features of biological macromolecules. Here we eschew as much detail as possible and strive for the simplest model that produces terahertz circular dichroism

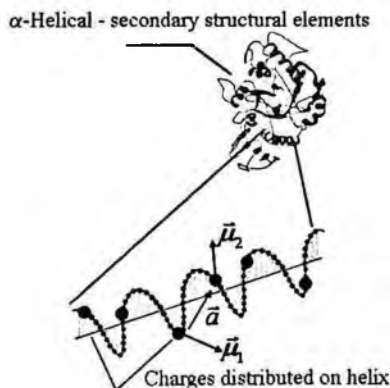


Figure 2. Simple model of terahertz circular dichroism.

signatures and that we can manipulate to develop some intuition about spectral signatures and their order of magnitude.

Our model is comprised of a collection of masses distributed uniformly in a helical pattern (Figure 2). Each mass represents more than a simple atom or ion; it represents the covalently bonded atom on the backbone and whatever is attached to it. We assume that these units interact only with their nearest neighbors. For the sake of simplicity we assume all masses are the same and we treat only the acoustic-like torsional motion of the masses about the axis of the helix. The model is further simplified by not considering radial and longitudinal acoustic modes. To couple to the propagating terahertz electromagnetic field and produce a polarization that attenuates and alters the speed of the electromagnetic wave, charges are distributed along the helix.

To obtain numerical estimations of both terahertz absorption and circular dichroism in a prototypical helix-containing macromolecule, we employed the following parameters in our analytical solution of our modeled chiral medium: period of the helix, 0.55 nm; masses per period, 10; charge distribution, $-5e$ every 14 sites; highest frequency acoustic mode, 10 THz; volume density or helix period/ m^3 , $0.9 \times 10^{27} m^{-3}$; relaxation time, $0.53 \times 10^{-12} s$. The torsional acoustic modes supported by our specific model of the helix are shown in Figure 3 (Top). As in linear chain models of acoustic modes in solids, the highest frequency mode corresponds to an acoustic wavelength of $\lambda_{max} = 2a$, twice the site spacing. This particular calculation assumes that the helix is infinitely long. A finite helix will limit the number of modes on the helix and introduce a long wavelength cut-off. Guided by typical terrestrial examples, we take the helix to be 7 periods long and estimate that the longest wavelength is approximately $\lambda_{max} = 140a$, corresponding to a frequency of ~ 0.5 THz.

In reality, the excitation of the helix in the biomolecule is not uniform; there are a variety of masses, extraneous coupling to the environment, different charges and different local field corrections. The electromagnetic field, very nearly uniform on the length scale of the helix, drive a variety of modes determined by the Fourier transform of the charge coupling parameters along the helix. Here we have

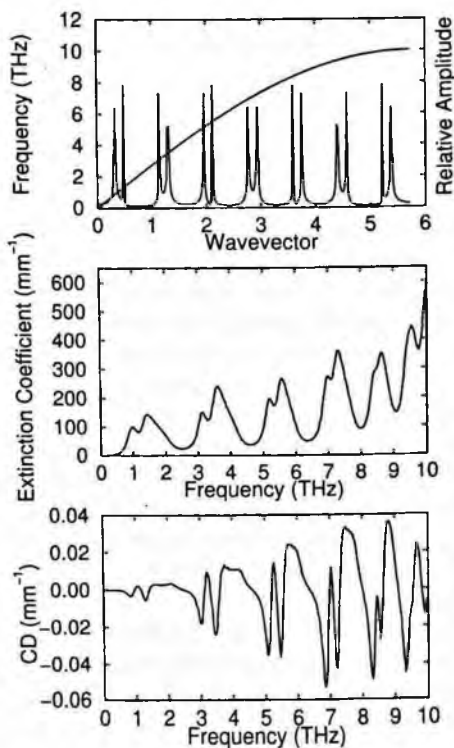


Figure 3. Simulated torsional acoustic modes (Top), THz Absorption (Middle) and THz circular dichroism (Bottom).

simply lumped a charge every 14 sites to simulate a very non-uniform excitation of the helix. The consequences of this particular distribution of charge are shown in Figure 2. The electromagnetic field at a given frequency will resonate with a torsional acoustic mode at a wavevector determined by the dispersion relation, but produces amplitude exhibiting resonant behavior at acoustic wavelengths or wavevectors determined by the spatial variation of the charge in the helix. We can conclude that a particular helix will have a signature determined by among other things the distributions of mass and charge.

Several key features emerge. As expected, the helical model absorbs strongly at the frequencies corresponding to the resonant excitation shown (Figure 3, Middle). For the model density, the absorption is quite strong. Indeed, it is well known that glasses with reasonable charge density, such as sodium silicate glasses, exhibit strong terahertz absorption that strengthens with increasing frequency. The terahertz circular dichroism spectral features consist of a series of zero crossings (Figure 3, Bottom). In the simple model the resonant features are simply determined by the distribution of charges. We may expect that any particular helical unit exhibits a unique, broad band “signature” pattern of such zero crossings.

Circular dichroism is always a small fraction of the direct absorption. Using visible, UV or infrared ratios as a guide we would expect that the ratio will be $\sim a/\lambda$ THz, where a is of the order of the atom or ion spacing. This is a very small number at terahertz frequencies. But the molecular unit, “terahertz chromophore” that is responsible for the terahertz absorption involves many atoms and ions²². For a helix 7 periods in length, the long wavelength acoustic mode in our simple model is $\sim 140 a$, and we expect to recover “circular dichroism/absorption” ratios that are reasonable relative to those of electronic excitations or near infrared vibrations. For this particular, rather simple model we find a ratio of $\sim 10^{-4}$, suggesting that the circular dichroism signal at terahertz frequencies may be at least as intense as that observed at optical wavelengths.

3.2. Terahertz Circular Dichroism Spectrometer

A reflective circular polarizer is used to construct the terahertz circular dichroism spectrometer (Error! Reference source not found.). The spectrometer explores terahertz circular dichroism by modulating the incident circular polarizations between left and right and measuring differential absorption of these two circular polarizations.

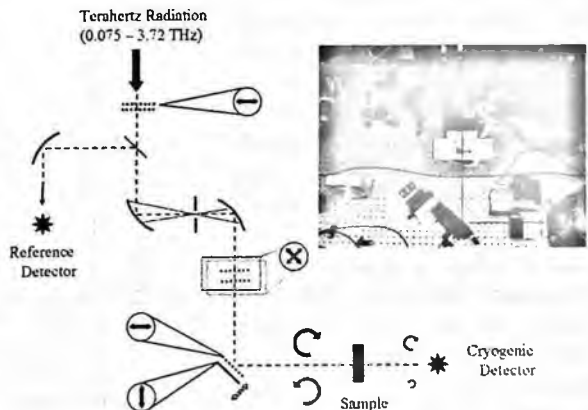


Figure 4. Schematic, photo (inset), of the terahertz circular dichroism spectrometer. A reflective circular polarizer (mirror and wire grid polarizer) is employed to produce left and right circular polarization.

The reflective circular polarizer consists of a free-standing wire-grid polarizer parallel to and in front of a flat mirror. Half of the radiation is immediately reflected by the grid, while the other half, polarized in the orthogonal direction, is reflected by the flat mirror and delayed by $\lambda/4$ according to the spacing between the wire-grid polarizer and the mirror. The resulting circular polarization is stable over time and not affected by thermal variation.

The quality of the right or left circular polarization depends on the relative amplitudes of the two phases and is readily adjusted by controlling the linear polarization of incident radiation ($\sim 45^\circ$ with respect to the polarizer in the circular polarizer reflector) and the aforementioned phase shift in the reflective circular polarizer.

The purity of circular polarization is stable and $\sim 85\%$ single-handed circular polarization in intensity (Figure 5). Since the terahertz circular dichroism is measured as a function of path length through the water/protein solution the offset presented by the circular polarization reflector is not a problem provided that it is stable, unchanging.

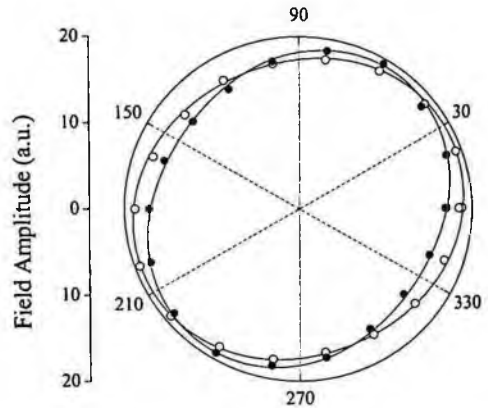


Figure 5. We obtain approximately 85% single handedness in intensity for both left and right circular polarizations.

3.3. Magnetic Terahertz Circular Dichroism in Semiconductors

To calibrate and test the terahertz circular dichroism spectrometer, we successfully measured magnetic terahertz circular dichroism in a semiconductor. A Helmholtz coil was fabricated that produced an axial magnetic field, homogeneous in the transverse direction. A doped, 500 μm thick, indium arsenide (InAs) wafer is placed in this magnetic field. Cyclotron motion of the carriers produces circular dichroism that is proportional to the applied magnetic field²⁴. By adjusting the strength of the magnetic field, we control the differential absorption of the two circular polarizations (Figure 6). Measuring this magnetic field dependent differential absorption mitigates the effect of offsets in the instrument.

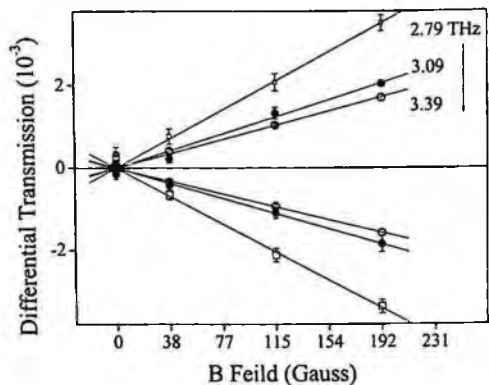


Figure 6. Frequency dependent magnetic circular dichroism in InAs measured by our terahertz circular dichroism spectrometer.

3.4. Terahertz Circular Dichroism in Solvated Biopolymers

To explore terahertz circular dichroism in biopolymers in water, we replace the semiconductor with the solution. Instead of a varying magnetic field, the path length of the solution is varied, as it was for the direct absorption. The measured terahertz circular dichroism is expressed as

$$\frac{T_{Left} - T_{Right}}{T_{Left} + T_{Right}} = 2 \cdot TCD \cdot \Delta l. \quad (1)$$

Concentrated solutions of proteins are contained in transparent polyethylene bags, and squeezed between rigid transparent polyethylene windows of the variable pathlength sample cell²⁵⁻²⁷. The differential transmission of left, right circularly polarizations is measured at various sample thickness, Δl , and the slope extracted to obtain a measure of the terahertz circular dichroism in the sample. This approach employs the readily measured transmitted powers, $T_{Left,Right}$, and change in pathlength, Δl . Offsets and artifacts are minimized as the only physical parameter that is changed is the pathlength through the solution.

We searched for terahertz circular dichroism signatures in BSA solutions (shown for 2.52 THz in Figure 7). Six distinct pathlengths were employed to cover a net traveling distance of 55 μm . At the current level of uncertainty, there is no dependence on differential transmission on pathlength, indicating that the terahertz circular dichroism signal is no larger than 0.1 cm^{-1} . But model estimates predict $\sim 0.01 \text{ cm}^{-1}$ and it is clear that the current instrument is not sensitive enough to recover terahertz circular dichroism in proteins in solution.

In the light of these results, previous early but encouraging reports by us on terahertz circular dichroism in biopolymers in water, are not valid and may be considered artifacts¹⁶.

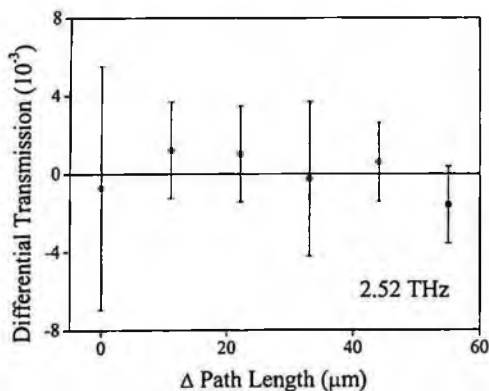


Figure 7. Upper limit of THz circular dichroism signature of BSA in water ($\sim 27\%$ w/w).

4. Conclusion and Direction

By direct absorption measurements, we have directly probed the low frequency collective motions of proteins in water by terahertz absorption spectroscopy. We have isolated, for the first time, the 0.075 - 3.72 THz absorption of bovine serum albumin¹⁴ and hen egg white lysozyme¹⁵ in water, and made direct comparison¹⁵ with existing

theoretical models¹⁰⁻¹³. We observe rough qualitative agreement between experiment and theory; important detailed differences are currently under further examination.

Terahertz circular dichroism provides a potential means for biochemical agent sensing while eliminating both interference artifacts and the strong background absorption of non-biological materials. However, the strength of the signals is very weak and will require at least an order of magnitude improvement in the sensitivity. In the light of these results, earlier preliminary reports by us of detection of terahertz circular dichroism by proteins in water may be viewed as suspect and experimental artifacts¹⁶. We have placed upper limits on terahertz circular dichroism of some biopolymers in water but they do not conflict with model estimates and more experimental work is called for.

Acknowledgement

We thank Munkyo Seo, Zach Griffith, and Mark Rodwell of UCSB for kindly providing us with harmonic multipliers; Jon Bjarnason, Tak Ling Wilkinson, and Elliot Brown of UCSB for kindly providing us with photomixer setup. This research is supported by Army Research Office through grants DAAD19-02-01-0080 and W911NF0510034; NASA through grant NAG5-12150. JX gratefully acknowledges the NASA graduate student research program (GSRP); a UCSB graduate opportunity fellowship, a NSF LEAPS (Let's explore applied physical science) Fellowship, a UCSB Mitsubishi chemical distinguished graduate fellowship, a UCSB science and engineering research grant, and a UCSB Broida-Hirschfelder fellowship.

References

1. M. F. Jeng, A. P. Campbell, T. Begley, A. Holmgren, D. A. Case, P. E. Wright and H. J. Dyson, High resolution solution structures of oxidized and reduced *Escherichia coli* thioredoxin, *Structure*, 2(9), 853-868 (1994).
2. R. Kitahara, S. Sareth and H. Yamada, High pressure NMR reveals active-site hinge motion of folate-bound *Escherichia coli* dihydrofolate reductase, *Biochemistry*, 39(42), 12789-12795 (2000).
3. B. Brooks and M. Karplus, Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor, *Proc. Natl. Acad. Sci. USA*, 80(21), 6571-6575 (1983).
4. B. Brooks and M. Karplus, Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme, *Proc. Natl. Acad. Sci. USA*, 82(15), 4995-4999 (1985).
5. Y. A. Chen and E. W. Prohofsky, Sequence and temperature dependence of the interbase hydrogen-bond breathing modes in B-DNA polymers: comparison with low-frequency Raman peaks and their role in helix melting, *Biopolymers*, 35(6), 573-582 (1995).

6. P. Dauber-Osguthorpe, D. J. Osguthorpe, P. S. Stern and J. Moulton, Low frequency motion in proteins: comparison of normal mode and molecular dynamics of streptomyces griseus protease A, *J. Comput. Phys.*, 151(1), 169-189 (1999).
7. N. Go, T. Noguti and T. Nishikawa, Dynamics of a small globular protein in terms of low-frequency vibrational modes, *Proc. Natl. Acad. Sci. USA*, 80(12), 3696-3700 (1983).
8. K. Hinsen, Analysis of domain motions by approximate normal mode calculations, *Prot. Struct. Func. Gen.*, 33(3), 417-429 (1998).
9. F. Tama, F. X. Gadea, O. Marques and Y.-H. Sanejouand, Building-block approach for determining low-frequency normal modes of macromolecules, *Prot. Struct. Func. Gen.*, 41(1), 1-7 (2000).
10. H. W. T. van Vlijmen and M. Karplus, Analysis of calculated normal modes of a set of native and partially unfolded proteins, *J. Phys. Chem. B*, 103(15), 3009-3021 (1999).
11. M. Levitt, C. Sander and P. S. Stern, Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme, *J. Mol. Biol.*, 181(3), 423 - 447 (1985).
12. A. Markelz, S. Whitmire, J. Hillebrecht and R. Birge, THz time domain spectroscopy of biomolecular conformational modes, *Phys. Med. Biol.* 47(21), 3797-3805 (2002).
13. C. B. Post, B. R. Brooks, M. Karplus, C. M. Dobson, P. J. Artymiuk, J. C. Cheetham and D. C. Phillips, Molecular dynamics simulations of native and substrate-bound lysozyme. A study of the average structures and atomic fluctuations, *J. Mol. Biol.*, 190(3), 455-479 (1986).
14. J. Xu, K. W. Plaxco and S. J. Allen, Probing the Collective Vibrational Dynamics of a Protein in Liquid Water by Terahertz Absorption Spectroscopy. *Protein Sci.*, 15(5), 1175-1181 (2006).
15. J. Xu, K. W. Plaxco and S. J. Allen, Collective dynamics of lysozyme in water: terahertz absorption spectroscopy and comparison with theory, submitted.
16. J. Xu, J. F. Galan, G. J. Ramian, P. G. Savvidis, A. M. Scopatz, R. R. Birge, S. J. Allen and K. W. Plaxco, Terahertz circular dichroism of biopolymers, *Proc. SPIE Int. Soc. Opt. Eng.*, 5268, 19-26 (2004).
17. J. Xu, G. J. Ramian, J. F. Galan, P. G. Savvidis, A. M. Scopatz, R. R. Birge, S. J. Allen and K. W. Plaxco, Terahertz circular dichroism spectroscopy: a potential approach to the *in situ* detection of life's metabolic and genetic machinery, *Astrobiology*, 3(3), 489-504 (2003).
18. D. S. Kliger, J. W. Lewis and C. D. Randall, Polarized light in optics and spectroscopy, Academic Press, San Diego (1990).
19. B. X. Huang, H.-Y. Kim and C. Dass, Probing three-dimensional structure of bovine serum albumin by chemical cross-linking and mass spectrometry, *J. Am. Soc. Mass. Spectrom.*, 15(8):1237-1247 (2004).
20. C. W. Deutsche and A. J. Moscowitz, Optical activity of vibrational origin. I. a model helical polymer, *J. Chem. Phys.*, 49(7), 3257-3272 (1968).
21. C. W. Deutsche and A. J. Moscowitz, Optical activity of vibrational origin. II. consequences of polymer conformation, *J. Chem. Phys.*, 53(7), 2630-2644 (1970).
22. G. Holzwarth and I. Chabay, Optical activity of vibrational transitions: a coupled oscillator model, *J. Chem. Phys.*, 57(4), 1632-1635 (1972).

23. J. A. Schellman, Vibrational optical activity, *J. Chem. Phys.*, 58(7), 2882-2886 (1973).
24. P. L. Richards and G. E. Smith, Far-infrared circular polarizer, *Rev. Sci. Instr.*, 35(11), 1535-1537 (1964).
25. A. J. Fitzgerald, E. Berry, N. N. Zinov'ev, S. Homer-Vanniasinkam, R. E. Miles, J. M. Chamberlain and M. A. Smith, Catalogue of human tissue optical properties at terahertz frequencies, *J. Biol. Phys.*, 29(2-3), 123-128 (2003).
26. J. T. Kindt and C. A. Schmittenmaer, Far-Infrared Dielectric Properties of Polar Liquids Probed by Femtosecond Terahertz Pulse Spectroscopy, *J. Phys. Chem.*, 100(24), 10373-10379 (1996).
27. J. Xu, K. W. Plaxco and S. J. Allen, Absorption spectra of liquid water and aqueous buffers between 0.3 - 3.72 terahertz, *J. Chem. Phys.*, 124, 036101, 1-3 (2006).

A SERS-BASED ANALYZER FOR POINT AND CONTINUOUS WATER MONITORING OF CHEMICAL AGENTS AND THEIR HYDROLYSIS PRODUCTS

STUART FARQUHARSON

FRANK E. INSCORE

*Real-Time Analyzers, 362 Industrial Park Road (#8),
Middletown, Connecticut 06457, United States of America¹
stu@rta_biz*

Protection of military personnel and civilians from water supplies poisoned by chemical warfare agents (CWAs) requires an analyzer that has sufficient sensitivity ($\mu\text{g/L}$, ppb), specificity (differentiate the CWA from its hydrolysis products), and speed (less than 10 minutes) to be of value. In an effort to meet these requirements, we have been investigating the ability of surface-enhanced Raman spectroscopy (SERS) to detect cyanide and sulfur mustard in water. In our work, we have developed a novel SERS-active material that consists of a porous glass with trapped metal particles. Previously, we coated the inside walls of glass vials and measured cyanide at 1 mg/L in water in as little as 1 minute. However, measurements of sulfur mustard have only been measured to 10 mg/L. Recently, we filled glass capillaries with the SERS-active material. Here we describe measurements of cyanide, sulfur mustard, and its hydrolysis product, thiodiglycol, using these capillaries and a portable Raman analyzer suitable for point and continuous water monitoring.

1. Introduction

The use of chemicals as weapons was introduced during World War I. It is estimated that chlorine, phosgene and sulfur-mustard (HD) resulted in an estimated death of 100,000 soldiers and 1 million injuries.¹ In more recent years, chemical agents have been used by terrorists. In Japan, the Aum Shinrikyo religious cult released Sarin (GB) within the Tokyo subway system in 1995.² The release of GB in this confined space had devastating effects resulting in 12 fatalities and hospitalization of thousands. Countering such attacks requires considering the potential deployment scenarios and having methods and/or devices to detect the chemical warfare agents rapidly, and at very low concentrations. Here we consider the deliberate poisoning of drinking water. This includes water supplies used in military operations and water delivered to major cities from reservoirs and through distribution systems.

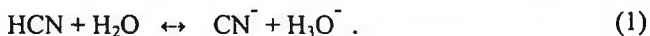
¹ Real-Time Analyzers, 362 Industrial Park Road (#8), Middletown, Connecticut, 06457, United States of America.

The required detection sensitivity for each agent depends on several factors, such as toxicity and hydrolysis properties (Table 1). In the case of cyanide, it is known that 4 milligrams per liter of water produces detectable changes in human blood chemistry and 8 mg/L causes severe, but reversible symptoms.³ The military has used this and other toxicological data to set a field drinking water standard (FDWS) for cyanide at 2 mg/L.⁴ The FDWS represents the maximum allowable concentration that is assumed safe when 15 L of water per day is consumed over 5 days (expected soldier intake in arid climates). Human toxicity data for the other chemical warfare agents in water have, in general, not been determined. The normal route of exposure for chemical warfare agents is inhalation, and most of the toxicity data is given as the LC_{50s},¹ the concentration that is lethal to 50% of an exposed population as a function of exposure time. In the case of mustard, animal studies along with the inhalation LC₅₀, the oral lethal dosage of 0.7 mg per kg of body mass (LD₅₀), and modeling studies,⁵ have been used to set the FDWS at 0.047 mg/L. The FDWS concentrations have also been used by the military to set the minimum detection requirement for poisons in water.

Table 1. Military field drinking water standard,⁴ lethal exposures and dosages,^{1,3,5} and water properties for selected chemical warfare agents.

Chemical	FDWS 5-day/15L (mg/L)	LC ₅₀ inhalation (mg-min m ⁻³)	LD ₅₀ oral (mg kg ⁻¹)	Water Solubility at 25°C	Hydrolysis Half-Life*	Hydrolysis Product
NaCN (CN)	2	2000	-	480 g/L	-	CN ⁻
Mustard (HD)	0.047	900	0.7	0.92 g/L	2-30 hours	TDG

In order to detect these poisons in water, their properties in water must also be considered, i.e. the solubility, rate of hydrolysis, and hydrolysis products formed. In the case of cyanide, as HCN, KCN, or NaCN, all of these chemicals are extremely soluble in water (completely miscible, 716, and 480 g/L, respectively).⁶ In solution the cyanide ion is formed in equilibrium with the conjugate acid, HCN Reaction (1), according to the K_a of 6.15x10⁻¹⁰.⁷ In the case of cyanide then it is important to know the pH, if one form of the chemical is to be detected versus the other. For example, if 2 mg/L of NaCN is added to water (the FDWS), then 1.25 mg/L of CN⁻ and 0.75 mg/L of HCN will be present.



(HD)

(TDG)

In the case of sulfur-mustard, the situation is somewhat more complex. It is marginally soluble in water tending to form droplets, and hydrolysis occurs at the droplet surface. This property has made measuring the hydrolysis rate constant difficult, and half-lives anywhere from 2 to 30 hours are reported.⁸ Chemically, the hydrolysis of HD involves the sequential replacement of the chlorine atoms by hydroxyl groups through cyclic sulfonium ion intermediates to form thiodiglycol (TDG), Reaction (2).⁹ If a median

hydrolysis rate is assumed, then early detection of poisoned water will require measuring HD, while post-attack or downstream monitoring will require measuring TDG.

The earliest technologies developed for CWA detection were based on electrochemical, ionization, or colorimetric analysis. Examples of the latter include phosgene, M8 and M9 tape, which change color from contact with a sample just like pH paper does. Although these tapes are easy to use, they are not generally agent specific and suffer from a high percentage of false-positives.¹⁰ For example, M8 changes color when in contact with common solvents such as acetonitrile, ethanol, methanol, or common petroleum products such as brake fluid, lighter fluid, or WD-40.¹¹ More rigorous laboratory methods have been successfully developed to detect chemical agents with minimum false-positive responses. More than a decade ago, Black et al. demonstrated the ability of combining gas chromatography with mass spectrometry detection (GC/MS) to measure mustard.¹² A comprehensive study was performed by Creasy et al. in analyzing chemical weapon decontamination waste from the Johnston Atoll.^{13, 14} These researchers used liquid chromatography with MS detection and achieved a detection limit of 0.140 mg/L for mustard. Although close to the FDWS, these researchers concede that chromatography coupled MS methods require extraction, derivatization, and repeated column calibration, making them labor intensive, time consuming (typically 30 to 60 minutes), and less than desirable for field use. Another variant of these separation/mass detection technologies is ion mobility spectrometry (IMS).¹⁵ Steiner et al. coupled electrospray ionization to the sample entry point of an IMS and a time-of-flight MS was added as an orthogonal detector.¹⁶ Water samples spiked with 10 mg/L (10 ppm) thiodiglycol could be measured in 1-min, once sample pretreatment was accomplished. It is worth noting that with proper care these MS-based technologies are likely to detect chemical agents with virtually no false-positives, but detection limits are still insufficient by 1 to 2 orders of magnitude.

More rapid analysis of agents in the solid, liquid and gas phase has been demonstrated by vibrational spectroscopy.^{17, 20, 21} Hoffland et al. reported infrared absorbance spectra and absolute Raman cross sections for several chemical agents,¹⁷ while Christesen measured Raman cross sections for mustard gas.¹⁸ Again, however these technologies also have limitations. Raman spectroscopy is simply not a very sensitive technique, and detection limits are typically 0.1% (1000 ppm). And infrared spectroscopy would have limited value in analyzing poisoned water, since the very strong infrared absorption of water would obscure most other chemicals present. Nevertheless, efforts to overcome these limitations have been demonstrated by using attenuated total reflectance in the sample cell.¹⁹

Enormous improvements in sensitivity for Raman spectroscopy can be achieved through surface-enhancement.²² The interaction of surface plasmon modes of metal particles with target analytes can increase scattering efficiency by as much 14-orders of magnitude, although 6-orders of magnitude are more common. Spencer, et al. used SERS to measure

cyanide and mustard on electrochemically roughed gold or silver foils.^{23, 24, 25} However, the sample needed to be dried on the substrates to obtain the best sensitivity. More recently, Tessier et al. obtained SERS of 0.04 mg/L cyanide in a stream flowing over a substrate formed by a templated self-assembly of gold nanoparticles.²⁶ In this case, optimum sensitivity required introduction of an acid wash and the measurements were irreversible.

In the past few years, we have also been investigating the ability of SERS to measure chemical agents at 0.001 mg/L in water and with sufficient spectral uniqueness to distinguish the agent and its hydrolysis products.^{27, 28, 29, 30, 31, 32} In our work, we have developed silver-doped sol-gels as the SERS-active medium. These sol-gels can be coated on the inside walls of glass vials, such that water samples can be added to perform point-analysis, or they can be incorporated into glass capillaries, such that flowing measurements can be performed.³³ Here, both sampling devices were used to measure and compare SER spectra of cyanide, sulfur mustard, and thiodiglycol. In addition, a field-usable Raman analyzer was used to measure 0.001 mg/L cyanide flowing in water with a detection time of less than 1-min.

2. Experimental

Sodium cyanide was purchased from Sigma-Aldrich (St. Louis, MO) and thiodiglycol (TDG, bis(2-hydroxyethyl)sulfide) was purchased from Cerilliant (Round Rock, TX). Both chemicals were used as received. Highly distilled sulfur mustard (HD, bis(2-chloroethyl)sulfide) was obtained at the U.S. Army's Edgewood Chemical Biological Center (Aberdeen, MD) and measured on-site. All samples were initially prepared in a chemical hood as 1000 parts-per-million (1 g/L or 0.1% by volume, Environmental Protection Agency definition) in HPLC grade water (Fischer Scientific, Fair Lawn, NJ) or in some cases methanol or ethanol (Sigma-Aldrich) to minimize hydrolysis.

Once prepared, the samples were transferred into 2-ml glass vials internally coated with a silver-doped sol-gel (*Simple SERS Sample Vials*, Real-Time Analyzers, Middletown, CT) or drawn by syringe or pump into 1-mm diameter glass capillaries filled with the same SERS-active material.^{34, 35, 36} In the case of flow measurements, a peristaltic pump (variable flow mini-pump, Control Co., Friendswood, TX) was used to flow the various cyanide solutions through a SERS-active capillary at 1 mL min⁻¹. The vials or capillaries were placed on aluminum plates machined to hold the vials or capillaries on a standard XY positioning stage (Conix Research, Springfield, OR), such that the focal point of an f/0.7 aspheric lens was positioned just inside the glass wall. The probe optics and fiber optic interface have been described previously.²⁹ SER spectra were collected using a Fourier transform Raman spectrometer equipped with a 785 nm diode laser and a silicon photo-avalanche detector (*IRA-785*, Real-Time Analyzers). All spectra were nominally collected using 100 mW, 8 cm⁻¹ resolution, and 1-min acquisition time, unless otherwise noted. Complete experimental details can be found in Reference. 37.³⁷ For added safety,

all samples were measured in a chemical hood. In the case of actual agents measured at Edgewood, the FT-Raman instrument was placed outside the laboratory and 30 foot fiber optic and electrical cables were used to allow remote SERS measurements and plate manipulation.

3. Results and Discussion

3.1. Cyanide

The surface-enhanced and normal Raman spectra of sodium cyanide are shown in Figure 1A. Sodium cyanide completely dissolves in water forming the ions in equilibrium with the conjugate acid, HCN as described above. Concentrations of 1.0, 0.1, and 0.01 mg/L result in CN^- concentrations of 0.52, 0.016, and 0.00021 mg/L as the corresponding pH decreases from just above the pKa of 9.21 at 9.24 to 8.48 and 7.54. This is significant in that only CN^- appears to interact sufficiently with silver to produce a SER spectrum, and no spectral signal is observed below pH 7 (except on electrodes at specific potential conditions).³⁸ The SER spectra of cyanide are dominated by an intense, broad peak at 2100 cm^{-1} attributed to the $\text{C}\equiv\text{N}$ stretch (Figure 1). This mode occurs at 2080 cm^{-1} in Raman spectra of solids, and the frequency shift in SER spectra is attributed to a strong surface interaction, which is supported by the appearance of a low frequency peak at 135 cm^{-1} due to a Ag-CN stretch (not shown). It is also observed that as the concentration decreases, the CN peak shifts to 2140 cm^{-1} . This shift has been attributed to the formation of a tetrahedral $\text{Ag}(\text{CN})_3^{2-}$ surface structure,³⁹ as well as to CN adsorbed to two different surface sites.⁴⁰ Alternatively, it has also been suggested that at concentrations near and above monolayer coverage, the CN^- species is forced to adsorb end-on due to crowding, and at lower concentrations the molecule can reorient to lie flat. This suggests that the 2100 and 2140 cm^{-1} peaks correspond to the end-on and flat orientations, respectively. However, a previous concentration study of cyanide on a silver electrode observed the reverse trend, i.e. greater intensity was observed for the 2100 cm^{-1} peak at low concentration.³⁹

Repeated measurements of cyanide in the SERS-active vials consistently allowed measuring 1 mg/L (1 ppm), but rarely below this concentration (Figure 1B). Nevertheless, this sensitivity is in general sufficient for point sampling of water supplies. In the case of continuous monitoring of water, the capillaries are a more appropriate sampling format, and they also allowed routine measurements at 0.01 mg/L and repeatable measurements at 0.001 mg/L (1 ppb, Figure 1C). Employing this format, a 50 mL volume of 0.01 mg/L cyanide solution was flowed at 2.5 mL min^{-1} through a SERS-active capillary, and spectra were recorded every 20 seconds. As Figure 2 shows, the cyanide peak was easily discerned as soon as the solution entered the capillary and remained relatively stable over the course of the experiment. It is worth noting, as indicated above, that the SERS peak in Figure 2 is in fact due to 210 ng/L!

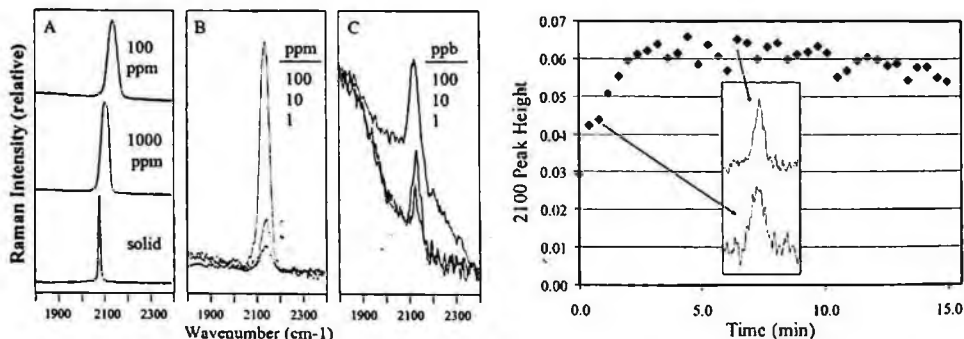


Fig. 1. Surface enhanced Raman spectra of CN in water in silver-doped sol-gel A) coated glass vials and B) filled glass capillaries. All spectra were recorded using 100 mW of 785 nm in 1-min and at a resolution of 8 cm^{-1} .

Fig. 2. 2100 cm^{-1} peak height measured during continuous flow of a 0.01 mg/L (10 ppb) cyanide in water. Surface-enhanced Raman spectra are shown for 1 and 6 min after sample introduction. A 2.5 mL min^{-1} flow rate was used and spectra were recorded every 20 sec using 100 mW of 785 nm.

3.2. HD and TDG

The surface-enhanced and normal Raman spectra of HD are shown in Figure 3. The SER spectrum of HD is dominated by a peak at 630 cm^{-1} with an extended high frequency shoulder composed of at least two peaks evident at 695 and 830 cm^{-1} , as well as a moderately intense peak at 1045 cm^{-1} (Figure 3A). The latter peak is assigned to a CC stretching mode, based on the assignment for a peak at 1040 cm^{-1} in the Raman spectrum of HD.⁴¹ The assignment of the 630 cm^{-1} peak is less straightforward, since the Raman spectrum of HD contains five peaks in this region at 640 , 655 , 700 , 740 , and 760 cm^{-1} .^{29, 41} Theoretical calculations for the Raman spectrum of HD indicate that the first three peaks are due to CCl stretching modes, and the latter two peaks to CS stretching modes.⁴² Based on these calculations, and the expected interaction between the chlorine atoms and the silver surface, it is reasonable to assign the 630 cm^{-1} SERS peak to a CCl mode.²⁹ However, recent SERS measurements of diethyl sulfide produced a very simple spectrum with an intense peak at 630 cm^{-1} ,^{43, 44} strongly suggesting CS or CSC stretching modes as the appropriate assignment for this peak.⁴⁵ The authors of the theoretical treatment concede that the CCl and CS assignments could be reversed.⁴² The CS assignment also indicates that HD interacts with the silver surface through the sulfur electron lone pairs. But, interaction between chlorine and silver is still possible and may be responsible for the 695 cm^{-1} peak. The 830 cm^{-1} peak may be due to a deformation mode (e.g. CCH).

The surface-enhanced and normal Raman spectra of TDG are shown in Figure 4. The SER spectrum of TDG is also dominated by a peak at 630 cm^{-1} with minor peaks at 820 , 930 , 1210 , and 1275 cm^{-1} (Figure 4A). Again, the 630 cm^{-1} peak is preferably assigned to a CSC stretching mode versus a CCl mode, especially since the chlorines have been replaced by hydroxyl groups. Furthermore, the lack of a 695 cm^{-1} peak in the TDG

spectrum supports the assignment of this peak in the HD spectrum to a CCl mode. The 930, 1210 and 1275 cm^{-1} SERS peaks are assigned to a CC stretch with CO contribution, and two CH_2 deformation modes (twist, scissors, or wag) based on the assignments for the corresponding peaks at 940, 1230 and 1290 cm^{-1} in the Raman spectrum of TDG.^{41, 43} It is worth noting that irradiation at high laser powers or for extended periods produces peaks at 715 and 1010 cm^{-1} , which are attributed to a degradation product, such as 2-hydroxy ethanethiol.⁴³

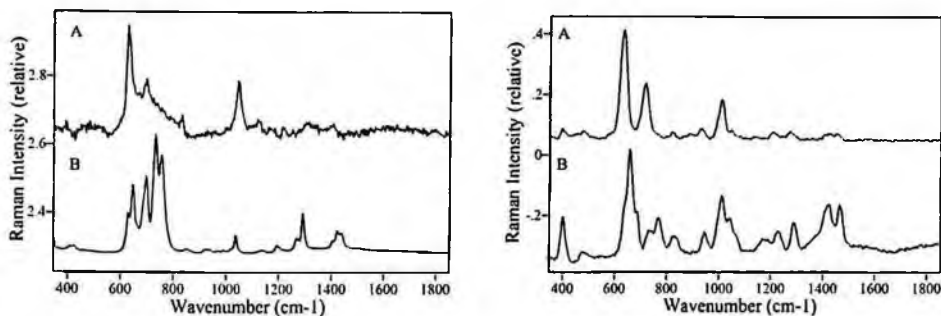


Fig. 3. A) SERS and B) RS of HD. A) 0.1% v/v (1000 ppm) in MeOH in a SERS-active vial, 100 mW of 785 nm, 1-min, B) neat sol. in glass container, 300 mW of 785 nm, 5-min.

Fig. 4. A) SERS and B) RS of TDG. A) 0.1% v/v in MeOH in SERS-active capillary, 100 mW of 785 nm, 1-min, B) neat sol. in glass capillary, 300 mW of 785 nm, 5-min.

Only a limited number of measurements of HD were performed to evaluate sensitivity, due to the safety requirements. HD was repeatedly observed at 1000 ppm and usually observed at 100 ppm in the SERS-active vials.²⁹ But even at the latter concentration, substantial improvements in sensitivity are required to approach the required 0.05 mg/L (50 ppb) sensitivity. More extensive experiments were performed on HD's hydrolysis product, TDG since this chemical is safely handled in a regular chemical lab.

Repeated measurements of TDG in the SERS-active vials consistently allowed measuring 10 ppm, but repeated measurements of 1 ppm did not yield any discernable peaks (lowest trace in Figure 5). Flowing TDG through SERS-active capillaries allowed repeatable measurements at 10 ppm, and routine measurements at 1 ppm as shown in Figure 6.⁴⁴ There is an important difference between the TDG spectra recorded for static and flowing samples, namely that the 715 cm^{-1} peak is noticeably more intense in the static sample. This suggests that it may represent a photo-degradation product, which was recently verified.⁴⁶

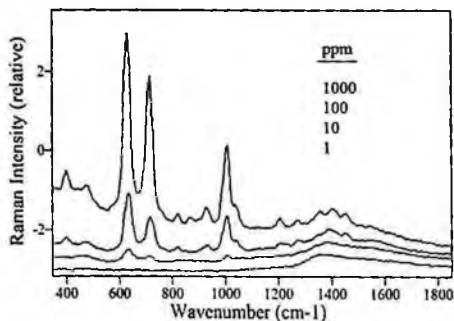


Fig. 5. SERS of 1000, 100, 10 and 1 ppm TDG in water (top to bottom). All in SERS-active capillaries, 100 mW of 785 nm, 1-min.

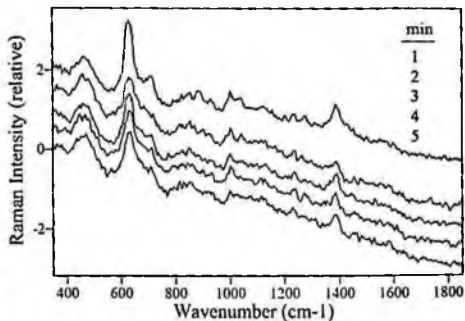


Fig. 6. SERS of 1 ppm TDG in water flowing through a SERS-active capillary at 1, 2, 3, 4, and 5 min. (top to bottom), 100 mW of 785 nm, 1-min each.

4. Conclusions

The ability to obtain surface-enhanced Raman spectra of two chemical agents, cyanide and mustard, and their hydrolysis products has been demonstrated using silver-doped sol-gels. Two sampling devices, SERS-active vials and capillaries, provided a simple means to measure water samples containing chemical agents. No sample pretreatment was required and all spectra were obtained in 1 minute. The SERS-active vials and capillaries provided sufficient sensitivity to measure cyanide below the required 2 mg/L sensitivity either as a point measurement or as a continuous flowing stream measurement. Measurements of TDG suggest that the sensitivity requirements for it and HD may be attainable with modest improvements.

5. Acknowledgements

The authors are grateful for the support of the U.S. Army (DAAD13-02-C-0015, Joint Service Agent Water Monitor program) and the Environmental Protection Agency (EP-D-05-034). The authors would also like to thank Mr. Chetan Shende for sol-gel chemistry development.

References

1. S. L. Hoenig, *Handbook of Chemical Warfare and Terrorism* (Greenwood Press, London, 2002)
2. H. Nozaki and N. Aikawa, Sarin poisoning in Tokyo subway, *Lancet* **345**, 1446 (1995).
3. Committee on Toxicology, *Review of Acute Human-Toxicity Estimates for Selected Chemical-Warfare Agents* (Nat Acad Press, Washington, DC, 1997).
4. Committee on Toxicology, *Guidelines for Chemical Warfare Agents in Military Field Drinking Water* (Nat Acad Press, Washington, DC, 1995).
5. T. C. Marrs, R. L. Maynard and F. R. Sidell, *Chemical Warfare Agents: Toxicology and Treatment* (John Wiley and Sons, 1996).
6. Material Safety Data Sheets, available at www.msds.com.

7. D. R. Lide, Ed, *Handbook of Chemistry and Physics* (CRC Press, Boca Raton, 1997).
8. N. B. Munro, S. S. Talmage, G. D. Griffin, L. C. Waters, A. P. Watson, J. F. King and V. Hauschild, The Sources, Fate, and Toxicity of Chemical Warfare Agent Degradation Products, *Environ. Health Perspect.* **107**, 933 (1999).
9. A. G. Ogsten, E. R. Holiday, J. St. L. Philpot and L. A. Stocken, The replacement reactions of b,b'-dichlorodiethyl sulphide and of some analogues in aqueous solution: the isolation of b-chloro-b'-hydroxydiethyl disulphide, *Trans. Faraday Soc.* **44**, 45 (1948).
10. B. Erickson, The Chemical Weapons Convention Redefines Analytical Challenge, *Anal. Chem. News & Features* June 1: 397A (1998).
11. Product literature at <http://www.wmdetect.com/Library/M8/M8%20Paper.htm>.
12. R. M. Black, R. J. Clarke, R. W. Read and M. T. Reid, Application of gas chromatography-mass spectrometry and gas chromatography-tandem mass spectrometry to the analysis of chemical warfare samples, found to contain residues of the nerve agent sarin, sulphur mustard and their degradation products, *J. Chromat.* **662**, 301 (1994).
13. W. Creasy, M. Brickhouse, K. Morrissey, J. Stuff, R. Cheicante, J. Ruth, J. Mays, B. Williams, R. O'Connor and H. Durst, Analysis of chemical weapons decontamination waste from old ton containers from Johnson atoll using multiple analytical methods, *Environ. Sci. Technol.* **33**, 2157 (1999).
14. W. R. Creasy, Postcolumn Derivatization Liquid Chromatography/Mass Spectrometry for Detection of Chemical-Weapons-Related Compounds, *Am. Soc. Mass. Spectrom.* **10**, 440 (1999).
15. G. A. Eiceman and Z. Caras, *Ion Mobility Spectrometry* (CRC Press, Boca Raton, 1994).
16. W. E. Steiner, B. H. Clowers, L. M. Matz, W. F. Siems and H. H. Hill Jr., Rapid screening of aqueous chemical warfare agent degradation products: ambient pressure ion mobility mass spectrometry, *Anal. Chem.* **74**, 4343 (2002).
17. L. D. Hoffland, R. J. Piffath and J. B. Bouck, Spectral signatures of chemical agents and simulants, *Opt. Eng.* **24**, 982 (1985).
18. S. D. Christesen, Raman cross sections of chemical agents and simulants, *Appl. Spectrosc.* **42**, 318 (1988).
19. E. H. J. Braue and M. G. Pannella, Circle Cell FT-IR Analysis of Chemical Warfare Agents in Aqueous Solutions, *Appl. Spectrosc.* **44**, 1513 (1990).
20. C-H. Tseng, C. K. Mann and T. J. Vickers, Determination of Organics on Metal Surfaces by Raman Spectroscopy, *Appl. Spectrosc.* **47**, 1767 (1993).
21. S. Kanan and C. Tripp, An infrared study of adsorbed organophosphonates on silica: a prefiltering strategy for the detection of nerve agents on metal oxide sensors, *Langmuir* **17**, 2213 (2001).
22. D. L. Jeanmaire and R. P. Van Duyne, Surface Raman Spectroelectrochemistry. *J. Electroanalytical Chem.* **84**, 1 (1977).
23. K. M. Spencer, J. Sylvia, S. Clauson and J. Janni, Surface Enhanced Raman as a Water Monitor for Warfare Agents in Water, *Proc. SPIE* **4577**, 158 (2001).
24. S. D. Christesen, M. J. Lochner, M. Ellzy, K. M. Spencer, J. Sylvia and S. Clauson, Surface Enhanced Raman Detection and Identification of Chemical Agents in Water, 23rd Army Science Conference (2002).
25. S. D. Christesen, K. M. Spencer, S. Farquharson, F. E. Inscore, K. Gosner and J. Guicheteau, Surface Enhanced Raman Detection of Chemical Agents in Water, In: S. Farquharson, Ed. *Applications of Surface-Enhanced Raman Spectroscopy* (CRC Press, Boca Raton, in preparation).
26. P. Tessier, S. Christesen, K. Ong, E. Clemente, A. Lenhoff, E. Kaler and O. Velev, On-line spectroscopic characterization of sodium cyanide with nanostructured gold surface-enhanced Raman spectroscopy substrates, *Appl. Spectrosc.* **56**, 1524 (2002).

27. Y. Lee and S. Farquharson, Rapid chemical agent identification by SERS. *Proc. SPIE* **4378**, 21 (2001).
28. S. Farquharson, P. Maksymiuk, K. Ong and S. Christesen, Chemical agent identification by surface-enhanced Raman spectroscopy, *Proc. SPIE* **4577**, 166 (2001).
29. S. Farquharson, A. Gift, P. Maksymiuk, F. Inscore, W. Smith, K. Morrissey and S. Christesen, Chemical agent detection by surface-enhanced Raman spectroscopy, *Proc. SPIE* **5269**, 16 (2004).
30. S. Farquharson, A. Gift, P. Maksymiuk, F. Inscore and W. Smith, pH dependence of methyl phosphonic acid, dipicolinic acid, and cyanide by surface-enhanced Raman spectroscopy, *Proc. SPIE* **5269**, 117 (2004).
31. F. Inscore, A. Gift, P. Maksymiuk and S. Farquharson, Characterization of chemical warfare G-agent hydrolysis products by surface-enhanced Raman spectroscopy, *Proc. SPIE* **5585**, 46 (2004).
32. S. Farquharson, A. Gift, P. Maksymiuk and F. Inscore, Surface-enhanced Raman spectra of VX and its hydrolysis products, *Appl. Spectrosc.* **59**, 654 (2005).
33. S. Farquharson and P. Maksymiuk, Simultaneous chemical separation and surface-enhancement Raman spectral detection using silver-doped sol-gels, *Appl. Spectrosc.* **57**, 479 (2003).
34. S. Farquharson, Y. H. Lee and C. Nelson, Material for surface-enhanced Raman spectroscopy and SER sensors, and method for preparing same, U.S. Patent Number 6,623,977 (2003).
35. S. Farquharson and P. Maksymiuk, Simultaneous chemical separation and surface-enhanced Raman spectral detection using metal-doped sol-gels, U.S. Patent Number 6,943,031. Separation and Plural-point surface-enhanced Raman spectral detection using metal-doped sol-gels, U.S. Patent Number 6,943,032 (2005).
36. S. Farquharson, A. Gift, P. Maksymiuk and F. Inscore, Rapid dipicolinic acid extraction from *Bacillus* spores detected by surface-enhanced Raman spectroscopy, *Appl. Spectrosc.* **58**, 351 (2004).
37. F. Inscore, A. Gift, P. Maksymiuk, J. Sperry and S. Farquharson, In: S. Farquharson, Ed. *Applications of Surface-Enhanced Raman Spectroscopy* (CRC press, Boca Raton, in preparation).
38. D. Kellogg and J. Pemberton, Effects of solution conditions on the surface-enhanced Raman scattering of cyanide species at silver electrodes, *J. Phys. Chem.* **91**, 1120 (1987).
39. J. Billmann, G. Kovacs and A. Otto, Enhanced Raman effect from cyanide adsorbed on a silver electrode. *Surf. Sci.* **92**, 153 (1980).
40. C. A. Murray and S. Bodoff, Depolarization effects in Raman scattering from cyanide on silver island films, *Phys. Rev. B* **32**, 671 (1985).
41. S. D. Christesen, Vibrational spectra and assignments of diethyl sulfide, 2-chlorodiethyl sulfide and 2,2'-dichlorodiethyl sulfide, *J. Raman Spectrosc.* **22**, 459 (1991).
42. C. Sosa, R. J. Bartlett, K. KuBulat and W. B. Person, A theoretical study of harmonic vibrational frequencies and infrared intensities of XCH₂CH₂SCH₂CH₂X and XCH₂CH₂SH (= H, Cl), *J. Phys. Chem.* **93**, 577 (1993).
43. F. Inscore and S. Farquharson, Surface-enhanced Raman Spectroscopic characterization of sulfur mustard, half-mustard, their hydrolysis products and related mono-sulfides, *J. Raman Spectros.* (in preparation).
44. F. Inscore and S. Farquharson, Detecting hydrolysis products of blister agents in water by surface-enhanced Raman spectroscopy, *Proc. SPIE*, **5993**, 19-23 (2005).
45. T. Joo, K. Kim and M. Kim, Surface-enhanced Raman study of organic sulfides adsorbed on silver, *J. Molec. Struct.* **16**, 191 (1987)
46. S. Farquharson, F. E. Inscore and S. Christesen, In K. Kneipp, M. Moskovits and H. Kneipp, Eds., *Surface-Enhanced Raman Scattering – Physics and Applications* (Springer-Verlag, Berlin, in press, 2006).

REAGENTLESS BIO-SAMPLING METHODS FOR IR DETECTION

LUKE D. DOUCETTE*, HE LI[§], BRIAN J. NINNESS*, CARL P. TRIPP*[§]

*Orono Spectral Solutions, 20 Godfrey Drive, Orono, ME 04473

[§]Laboratory for Surface Science and Technology, University of Maine, Orono ME 04469

Currently there exists a critical need within the military and homeland defense for highly sophisticated yet, small, lightweight portable sensors and detection systems for identifying and quantifying biological and biowarfare agents (BWA) in both liquid and aerosolized form. Our proposed BWA detection system is based upon Fourier Transform Infrared Spectroscopy (FTIR), where the main advantages of this approach are that it is reagentless, operates in heterogeneous aqueous environments, and provides fast detection and high sensitivity/selectivity to bacterial spores with minimal false alarms.

The key enabler to using FTIR for BWA detection is to develop selective and robust sampling protocols coupled to a miniaturized, portable FTIR unit. To that end, we have developed front-end liquid flow cells which incorporate electric field (E-Field) concentration methods for spores onto the surface of an Attenuated Total Reflection (ATR) IR crystal. IR spectra are presented which show collection and detection results with BG spores in water. The approaches we have developed take advantage of the fact that all spores are negatively charged in neutral pH solutions. Therefore, E-Field concentration of spores directly onto an ATR sampling element enables low level concentration measurements to be possible.

keywords: infrared spectroscopy, spore, detection, zeta potential, principal component analysis

1. Background

Bacterial spores are dormant cells which are produced by a variety of bacteria including *Bacillus subtilis* and *Bacillus anthracis*. These spores are encased within a complex multi-layered protein structure, whose primary function is to help the spore endure a wide range of harsh environmental stresses. Once conditions are favorable for germination, these spores retain the ability to return to a vegetative state. Irrespective of the particular strain (*B. anthracis*, *B. subtilis*, *B. cereus*), the spore dimensions are typically confined to a narrow range of approximately 1.2 μm in length by 0.75 μm wide.¹ The surface chemistry of the outer protein coat determines the extent of charge acquired at various solution conditions (pH, ionic strength).² This is primarily due to the presence of ionizable groups such as amine, carboxylate, and phosphate functionalities. The zero point of charge for typical spores is in the range of pH 4 – 5, indicating that negatively charged sites on the spore surface dominate at pH levels under normal physiological conditions.³ Therefore a straightforward reagentless approach to concentrate and collect bacterial spores is to use an oppositely charged surface as the collection medium.

The concept of using an electric field to concentrate bacteria and spores on an oppositely charged surface has been demonstrated in the literature.^{4,5,6} In particular, liquid flow through cells have been designed to concentrate spores onto surfaces by using electric fields coupled with detection methods that are often based upon the measurement of a single parameter (e.g. the change in refractive index of the underlying surface). Although the state of knowledge in this area clearly shows that concentration of spores is accomplished using an applied electrical field, there is a need for spore identification in order to reduce false alarms. Relying on physical changes (e.g. refractive index) does not constitute a specific biological detection and identification scheme.

On the other hand, successful identification of bacterial spores has been demonstrated by using Fourier transform infrared photoacoustic^{7,8} and transmission spectroscopy^{9,10,11} in conjunction with principal component analysis (PCA) statistical methods. In general, PCA methods are used to reduce and decompose the spectral data into orthogonal components, or factors, which represent the most common variations in all the data. As such, each spectrum in a reference library has an associated "score" for each factor. These scores can then be used to show clustering of spectra that have common variations, thus forming a basis for group member classification and identification.

In this article we present zeta potential measurements which show the spore surface charge dependence on pH for a variety of spore types. These results help provide the motivation for new sampling approaches that combine E-Field sampling with IR detection. Specifically, we present two reagentless methods for water-based collection and detection which are based upon E-Field concentration of spores onto the surface of an Attenuated Total reflection (ATR) IR crystal. An ATR based approach is used because there is a finite penetration of the IR beam into the aqueous medium. This overcomes the strong absorption of infrared radiation by water and thus enables continuous monitoring. The penetration depth of the evanescent wave in the ATR approach is 1- 5 μm which is a good match for "monolayer" or "bilayer" coverage of micron sized bacteria. The collection cells are designed to be easily interfaced with conventional pumps and IR detection equipment to enable rapid online monitoring of water-based bacterial spores. Finally, to demonstrate the identification capabilities of IR, results from a PCA discriminant analysis of transmission spectra for different spores are discussed.

2. Spore Surface Charge Characteristics

The extent of spore concentration using an E-Field is largely determined by the amount and type of charge associated with a given spore type in solution. To characterize this charge behavior, electrophoretic measurements using a Malvern Zetasizer were conducted to determine the zeta potential of four different bacillus spore types used in our studies (*B. Thuringiensis* (BT), *B. Atrophaeus* (BA), *B. Subtilis*, (BS), *B. Globigii Dugway* (BG)). Figure 1 shows the titration curves for these measurements, and indicates that the spores examined for this study have isoelectric points ranging

between 2.2 to 4.5, and are fully deprotonated (i.e. have a maximum net negative charge) at $\text{pH} > 6$. Also shown in Figure 1 is a zeta potential measurement obtained for alumina powder, which is seen to have an isoelectric point of approximately 9.5. In particular, over a normal pH range (6 – 8), alumina is positively charged while most types of spores are negatively charged. Therefore, anchoring an oppositely charged alumina layer to the surface of an ATR crystal could then be an effective means to concentrate spores for measurement and detection.

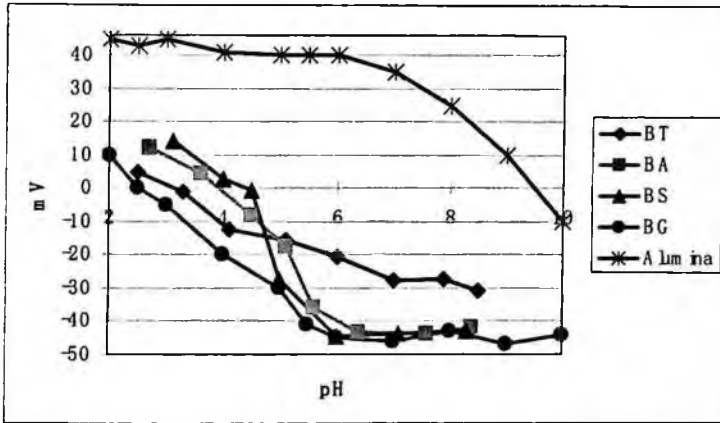


Figure 1: Zeta potential measurements showing surface charge dependence with pH for different spores and alumina

3. E-Field Spore Collection Methods for IR Detection

3.1 Method 1: Polymer + metal oxide coatings on ATR crystals

Electrophoretic measurements of alumina (Al_2O_3) and BG spores showed that, over a normal pH range (6 – 8), alumina is positively charged while BG spores are negatively charged (see Figure 1). To test the feasibility of using charged alumina films to concentrate oppositely charged spores from solution, a mixture of polyethylene (PE) and alumina powder was deposited onto a zinc selenide (ZnSe) ATR crystal (50 mm x 10 mm x 2 mm parallelogram), and a 10^8 spore/mL solution in deionized water (DIW) was flowed over the surface at approximately 1 mL/min. In this application, PE is used as a binder to anchor the alumina to the crystal surface and provides a robust film that remains stable for IR studies. The results of the ATR-FTIR measurements for this experiment are shown in Figure 2a, and demonstrate that the PE/alumina films concentrate the spores from solution when compared to a spiked reference spectrum of BG (Figure 2b). Using the reference as a calibrant, spore collection on the crystal surface saturated at approximately 5×10^6 total spores after approximately 1 hour collection time.

Although the PE + alumina coated ATR crystal provides one possible path for online monitoring of bacterial spores in solution, the time for adsorption and detection is

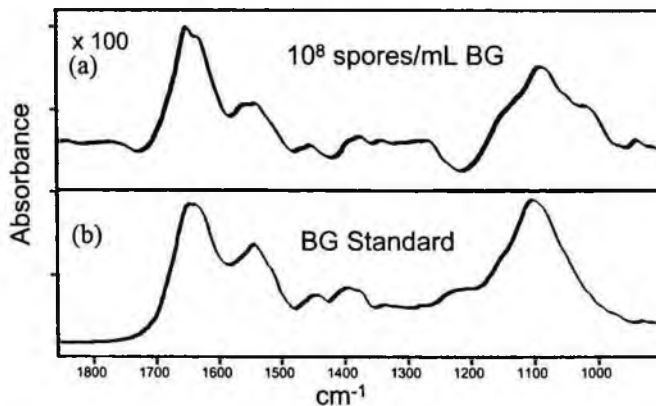


Figure 2 (a) BG spores adsorbed onto a Al_2O_3 +PE coating on a ZnSe ATR crystal b) BG Standard on an ATR crystal

ultimately limited by diffusion of spores in solution to the crystal surface (e.g. 1 hour collection/detection time). Furthermore, in uncontrolled environments, biofouling of the crystal is to be expected unless cleaning reagents are used to regenerate the crystal surface for continued use.

3.2 Method 2: E-Field Assisted ATR-FTIR

To overcome the limitations of alumina coated ATR crystals, a second collection method has been developed which uses an external electric field to both “assist” spores onto the crystal surface for detection, then “clean” by field reversal. To incorporate such an E-Field scheme with FTIR-ATR, we developed a flow-through cell that utilizes an ultra-thin metal coating on the crystal surface that is parallel to a counter electrode within the cell (see Figure 3a). By depositing an electrode on the ATR crystal surface, an applied E-field across the electrodes concentrates charged spores directly onto the ATR sampling element, thus enabling low level concentration measurements. In most ATR-IR applications, metallizing the active surface of an ATR crystal significantly absorbs the IR evanescent wave, often resulting in complete attenuation of the internally reflected IR radiation. However, in the thin film regime (< 20 nm), metal films on ATR crystals have been shown to be partially transparent in the mid-IR^{12,13,14}, thus enabling their use.

A flow through cell incorporating these concepts was designed and constructed from Teflon (Figure 3b), which allows flowing water to contact both sides of a metallized $50 \times 10 \times 2$ mm parallelogram ATR crystal. The ATR flow cell contains o-ring seals and spring loaded pogo pins that enable electrical contact to both crystal surfaces at locations on the crystal that do not come into contact with water. Inner cathode plates are constructed from stainless steel, and contain external electrical connections. The cell is designed to be easily interfaced with a standard power supply and peristaltic pump while in operation.

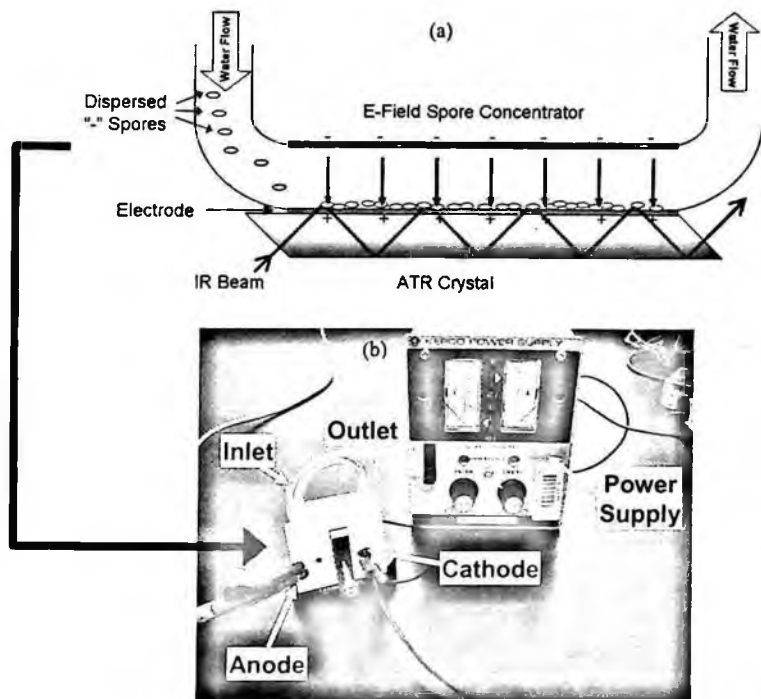


Figure 3: E-Field Assist ATR-FTIR

For the metal coated ATR crystal, a 5 nm palladium oxide (PdO) electrode film was deposited onto a germanium (Ge) crystal via magnetron sputtering at the Laboratory for Surface Science and Technology, University of Maine. Ge is used because of its relatively high index of refraction ($n = 4$), thereby retaining its IR waveguide characteristics when coated with a metallic film. PdO is used as the operating electrode material since pure Pd was observed to completely attenuate the IR signal; with PdO, however, the total IR signal was only attenuated by approximately $\frac{1}{2}$ the original intensity while still retaining its electrode behavior (resistivity of approximately $10^{-2} \Omega\text{cm}$).

To demonstrate the feasibility of the E-Field assist approach, a 10^8 BG spore/mL solution in DIW was flowed through the cell at approximately 1 mL/min. (same as previous test with the metal oxide coating). Figure 4a below shows the effects of concentration vs. voltage at 0 V, +2 V, and +5 V, and illustrates how the device increases the spore concentration when the voltage is applied. Spore saturation for both +2 V and +5 V measurements was observed to occur for collection times of approximately 1 minute. Although H_2 and O_2 gases can be produced by electrolysis at approximately 2.1 V, gas formation for our experiments was negligible due to the rapid collection time and relatively low operating voltages. For comparison, Figure 4b also shows the spectrum obtained at 5 V against that obtained with a standard (corresponding to 7×10^6 spores for the 5 V spectrum), and illustrates how the primary spectral features are consistent

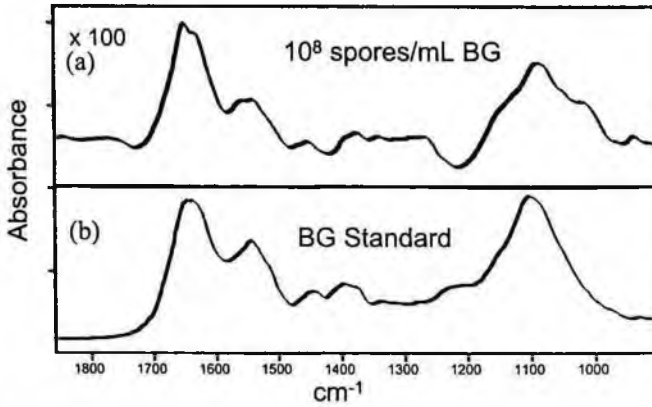


Figure 2 (a) BG spores adsorbed onto a Al_2O_3 +PE coating on a ZnSe ATR crystal b) BG Standard on an ATR crystal

ultimately limited by diffusion of spores in solution to the crystal surface (e.g. 1 hour collection/detection time). Furthermore, in uncontrolled environments, biofouling of the crystal is to be expected unless cleaning reagents are used to regenerate the crystal surface for continued use.

3.2 Method 2: *E-Field Assisted ATR-FTIR*

To overcome the limitations of alumina coated ATR crystals, a second collection method has been developed which uses an external electric field to both “assist” spores onto the crystal surface for detection, then “clean” by field reversal. To incorporate such an E-Field scheme with FTIR-ATR, we developed a flow-through cell that utilizes an ultra-thin metal coating on the crystal surface that is parallel to a counter electrode within the cell (see Figure 3a). By depositing an electrode on the ATR crystal surface, an applied E-field across the electrodes concentrates charged spores directly onto the ATR sampling element, thus enabling low level concentration measurements. In most ATR-IR applications, metallizing the active surface of an ATR crystal significantly absorbs the IR evanescent wave, often resulting in complete attenuation of the internally reflected IR radiation. However, in the thin film regime (< 20 nm), metal films on ATR crystals have been shown to be partially transparent in the mid-IR^{12,13,14}, thus enabling their use.

A flow through cell incorporating these concepts was designed and constructed from Teflon (Figure 3b), which allows flowing water to contact both sides of a metallized 50x10x2 mm parallelogram ATR crystal. The ATR flow cell contains o-ring seals and spring loaded pogo pins that enable electrical contact to both crystal surfaces at locations on the crystal that do not come into contact with water. Inner cathode plates are constructed from stainless steel, and contain external electrical connections. The cell is designed to be easily interfaced with a standard power supply and peristaltic pump while in operation.

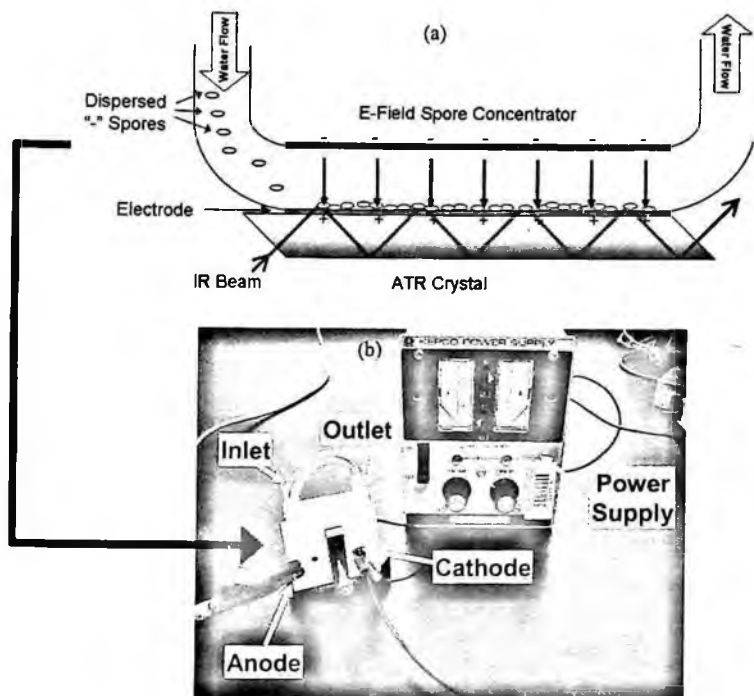


Figure 3: E-Field Assist ATR-FTIR

For the metal coated ATR crystal, a 5 nm palladium oxide (PdO) electrode film was deposited onto a germanium (Ge) crystal via magnetron sputtering at the Laboratory for Surface Science and Technology, University of Maine. Ge is used because of its relatively high index of refraction ($n = 4$), thereby retaining its IR waveguide characteristics when coated with a metallic film. PdO is used as the operating electrode material since pure Pd was observed to completely attenuate the IR signal; with PdO, however, the total IR signal was only attenuated by approximately $\frac{1}{2}$ the original intensity while still retaining its electrode behavior (resistivity of approximately $10^{-2} \Omega\text{cm}$).

To demonstrate the feasibility of the E-Field assist approach, a 10^8 BG spore/mL solution in DIW was flowed through the cell at approximately 1 mL/min. (same as previous test with the metal oxide coating). Figure 4a below shows the effects of concentration vs. voltage at 0 V, +2 V, and +5 V, and illustrates how the device increases the spore concentration when the voltage is applied. Spore saturation for both +2 V and +5 V measurements was observed to occur for collection times of approximately 1 minute. Although H_2 and O_2 gases can be produced by electrolysis at approximately 2.1 V, gas formation for our experiments was negligible due to the rapid collection time and relatively low operating voltages. For comparison, Figure 4b also shows the spectrum obtained at 5 V against that obtained with a standard (corresponding to 7×10^6 spores for the 5 V spectrum), and illustrates how the primary spectral features are consistent

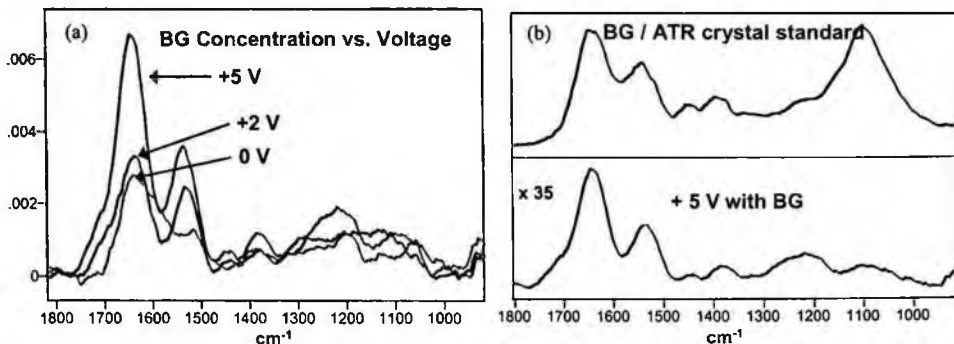


Figure 4

between the two spectra. The standard was obtained by concentrating spores into DIW, dispersing approximately 400 μg of spores onto both surfaces of a Ge ATR crystal, and allowed to dry before the spectrum was obtained. For the 5 V spectrum the relative peak heights at approximately 1250 and 1100 cm^{-1} are less intense than those measured with the standard. In general, these absorption bands correspond to the dense polysaccharide region within the spore cell walls.¹⁵ Therefore, the differences between the standard and 5 V spectra may be explained by the fact that the 5 V spectrum was obtained *in-situ* where a certain amount of cell wall deconstruction might be expected within an aqueous medium under applied voltage.

To “clean” the surface of the ATR crystal, the field was reversed (-2 V) and the spores were ejected as per the observed reduction in the large amide bands between 1700 – 1500 cm^{-1} . However, the ATR crystal degraded with repeated use, and reversing the field did not entirely regenerate the crystal to its initial state of performance. This reduction in performance lifetime is likely due to electrochemical changes that occurred within the electrode layer brought upon by repeated E-Field application. This suggests that a more robust film (i.e. platinum) may be needed for long term use. However, pure metal electrodes would likely require a shorter ATR crystal/flow cell in order to reduce the number of internal reflections since it was observed that such a film i.e. Pd completely attenuates the IR beam with the existing cell geometry.

4. Spore Identification via PCA Classification Methods

To demonstrate the robustness of Principal Component Analysis (PCA) classification and identification methods when applied to the IR spectra of different spore types, transmission FTIR spectral libraries were generated for three different spore strains (BG, BA, and BS) and classification models were developed based upon Mahalanobis Distance by PCA with Residuals (MD/PCA/R) statistical methods using PLSplus IQ[®] (Thermo Electron) chemometric software. Figure 5 shows a representative IR spectrum of a BG spore sample recorded in transmission, which in many respects is representative of all spore types due to their nearly identical compositions at the molecular level.

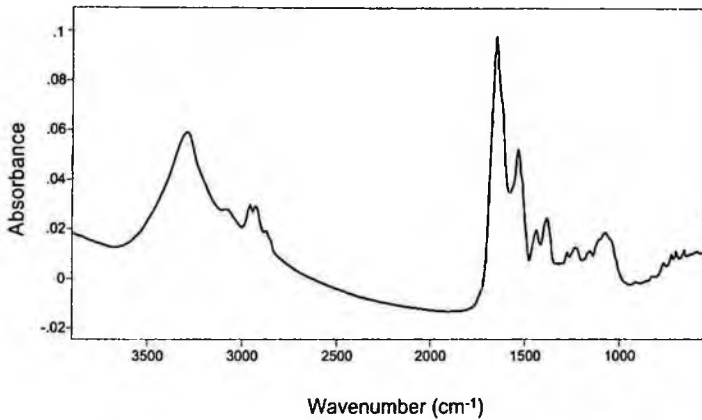


Figure 5. IR spectrum of BG spores in transmission

However, despite these spectral similarities, Figure 6 shows the PCA scores plot for the first three principal components of the reference library containing all three spore types, and clearly demonstrates how the principal component scores cluster into three easily distinguished groups which corresponded to the three different spore types. Prior to MD/PCA/R calculations, all spectra in the library were mean centered and normalized by setting the largest amide band ($\sim 1650\text{ cm}^{-1}$, see Figure 5) to 1 absorbance and the trough preceding it to zero. To increase classification sensitivity, cross validation was performed on the second derivatives of the spectra when the model was calibrated.

A truth table in Table I below summarizes classification results for blind samples (which were not part of the libraries) that were tested against separate

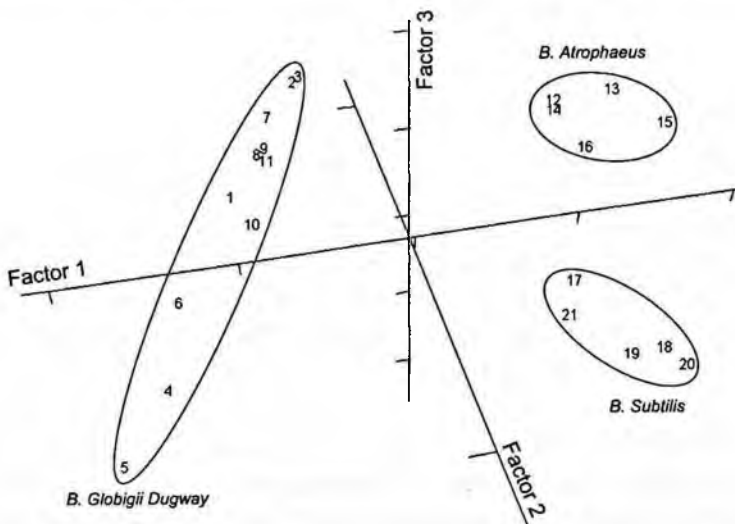


Figure 6: PCA scores plot for *B. Globigii Dugway*, *B. Subtilis*, and *B. Atrophaeus*

classification libraries for *B. Globigii Dugway* (Library 1), *B. Atrophaeus* (Library 2), and *B. Subtilis* (Library 3). Separate libraries for each group were generated since this provided further sensitivity for discriminant testing. Null hypothesis testing was used as the metric for discriminant analysis, where a blind sample having a P value < 0.001 was classified as a nonmember (null hypothesis is rejected), and a blind sample having a P value \geq 0.001 was classified as a member (null hypothesis is accepted). The results in Table 1 show how the three spore types were correctly classified as belonging to members of their respective libraries while testing as nonmembers for the other two libraries (i.e. no false positives). Furthermore, transmission FTIR spectra for two additional bacillus analogs (*B. Subtilis* subsp. *subtilis*, and *B. Thuringiensis* Serovar *kurstaki*) were tested, and both were correctly classified as nonmembers for all three libraries. These results, as well as the work of others, demonstrate the spore identification capabilities that can be achieved with an IR-based detection system.

Table 1

Test Sample	Classification Library P _{values} ^a		
	Library 1 (<i>B. Globigii Dug.</i>)	Library 2 (<i>B. Atrophaeus</i>)	Library 3 (<i>B. Subtilis</i>)
<i>B. Globigii Dugway</i>	0.506	< 0.001	< 0.001
<i>B. Atrophaeus</i>	< 0.001	0.019	< 0.001
<i>B. Subtilis</i>	< 0.001	< 0.001	0.365
<i>B. Subtilis sub. Subtilis</i>	< 0.001	< 0.001	< 0.001
<i>B. Thuringiensis</i>	< 0.001	< 0.001	< 0.001

^a P_{values} < (\geq) 0.001 are classified as a nonmember (member) of the library

5. Conclusions

New reagentless, E-Field based sampling approaches have been developed and integrated with conventional ATR-FTIR flow cells for the detection and identification of bacterial spores in water. Our results show how E-field concentration of spores onto the surface of an ATR crystal surface can enable low concentrations of spores to be measured. The collection cells are designed for easy interface with virtually any IR spectrometer or detection system, and can prove useful in a variety of military, homeland defense, and technological applications.

6. Acknowledgements

We gratefully acknowledge the assistance and guidance from Janet and Jim Jensen of the Edgewood Chemical and Biological Center, Aberdeen Proving Grounds, MD. This work was funded by the Department of Defense – Army, Contract #'s W911SR-04-C-0029 and W911SR-05-P-0041.

7. References

- ¹ V. G. R. Chada, E. A. Sanstad, R. Wang, A. Driks, Morphogenesis of Bacillus Spore Surfaces, *J. of Bacteriology* **185**, 6255-6261 (2003).
- ² S. Ross, I. D. Morrison, *Colloidal Systems and Interfaces* (Wiley, New York, 1988).
- ³ A. C. C. Plette, W. H. Van Riemsdijk, M. F. Benedetti, A. van der Wal, pH Dependent Charging Behavior of Isolated Cell Walls of a Gram-Positive Soil Bacterium, *Journal of Colloid and Interface Science* **173**, 354-363 (1995).
- ⁴ M. A. Brusatori, P. R. Van Tassel, Biosensing Under an Applied Voltage Using Optical Waveguide Lightmode Spectroscopy, *Biosensors and Bioelectronics* **18**, 1269-1277 (2003).
- ⁵ Y. Huang, K. L. Ewalt, M. Tirado, R. Haigis, A. Forster, D. Ackley, M. J. Heller, J. P. O'Connell, M. Krihak, Electric Manipulation of Bioparticles and Macromolecules on Microfabricated Electrodes, *Analytical Chemistr.* **73**, 1549-1559 (2001).
- ⁶ A. T. Poortinga, R. Bos, H. J. Busscher, Reversibility of Bacterial Adhesion at an Electrode Surface, *Langmuir* **17**, 2851-2856 (2001).
- ⁷ S. E. Thompson, N. S. Foster, T. J. Johnson, N. B. Valentine, J. E. Amonette, Identification of Bacterial Spores Using Statistical Analysis of Fourier Transform Infrared Photoacoustic Spectroscopy Data, *Appl. Spectrosc.* **57**, 893-899 (2003).
- ⁸ N. S. Foster, N. B. Valentine, Sandra E. Thompson, T. J. Johnson, and J. E. Amonette, FTIR Transmission and Photoacoustic Spectroscopy for the Statistical Identification of Bacteria, *Proc. SPIE* **5269**, 172-182 (2004).
- ⁹ N. S. Foster, S. E. Thompson, N. B. Valentine, J. E. Amonette, T. J. Johnson, Identification of Sporulated and Vegetative Bacteria Using Statistical Analysis of Fourier Transform Mid-infrared Transmission Data, *Appl. Spectrosc.* **58**, 203-211 (2004).
- ¹⁰ M. M. Mossoba, F. M. Khambaty, and F. S. Fry, Novel Application of a Disposable Optical Film to the Analysis of Bacterial Strains: A Chemometric Classification of Mid-Infrared Spectra, *Appl. Spectrosc.* **56**, 32-736 (2002).
- ¹¹ L. E. Rodriguez-Saona, F. M. Khambaty, F. S. Fry and E. M. Calvey, A Novel Approach for the Rapid Discrimination of Bacterial Strains by Fourier-Transform Near-Infrared Spectroscopy, *Proc. SPIE* **4206**, 22-31 (2001).
- ¹² T. Burgi, ATR-IR Spectroscopy at the Metal-Liquid Interface: Influence of Film Properties on Anomalous Band-Shape, *Phys. Chem. Chem. Phys.* **3**, 2124-2130 (2001).
- ¹³ T. Bugi, R. Wirz, and A. Baiker, In Situ Attenuated Total Reflection Infrared Spectroscopy: A Sensitive Tool for the Investigation of Reduction-Oxidation Processes on Heterogeneous Pd Metal Catalysts, *J. Phys. Chem. B* **107**, 6774-6781 (2003).
- ¹⁴ E. Zippel, M. W. Breiter, and R. Kellner, Fourier-transform Infrared, Attenuated Total Reflection Spectroscopy; A Complementary Tool for the Investigation of the Adsorption of CO on Thin Metal Layers, *J. Chem. Soc. Faraday Trans.* **87**, 637-642 (1991).
- ¹⁵ H. H. Mantsch, D. Chapman, *Infrared Spectroscopy of Biomolecules*, (Wiley, New York, 1996).

PRECONCENTRATION OF MICROORGANISMS INTO A TINY VOLUME OF LIQUID FOR ENHANCED SPECTRAL DETECTION

SOLOMON ZAROMB

Zaromb Research Corp., 9S 706 William Drive, Burr Ridge, Illinois 60527, United States

DENNIS MARTELL

Zaromb Research Corp., 9S 706 William Drive, Burr Ridge, Illinois 60527, United States

NATHAN SCHATTKE

Schattke Chemical Consulting, 308 Colton Street, Yorkville, Illinois 60560

GARY HANKINS

Hankins Manufacturing, 11S 360 Carpenter Street, Lemont, IL 60439

Concentration of 1-micron-size micro-organisms from about 100 ml of liquid, whether drawn from a bio-aerosol collector or from an environmental water source, into a volume of 1 to 2 ml. is achieved by a liquid flow system including a reversible filter through which filtered liquid can be recirculated or disposed of and from which a concentrated sample is recovered by opening a solenoid valve leading to a detector or to a collection container and reversing the pump for a short time. The reversing action flushes the collected particles off the filter and into the detector or container.

The effectiveness of this approach was demonstrated by measuring the cumulative fluorescence of 1-micron-size blue fluorescent microspheres versus cumulative volume withdrawn from our WEP collector after capture from an aerosol suspension of about 140 particles/ml drawn from a test chamber over a 5-minute period at a sampling rate of 500 liters/minute. As the liquid was being withdrawn from the WEP collector, it was filtered at a rate of 1 ml/second and the filter back-flushed with small portions of filtrate at 1-minute intervals. The relative cumulative concentration of captured particles in the first back-flushed 1-ml fraction was around 56 as compared with a value of 1.4 in the first 60-ml filtered fraction, which constitutes a liquid-to-liquid concentration enhancement by a factor of 40 and an air-to-liquid concentration factor of $>1.25 \times 10^6$.

Keywords: Preconcentration of microorganisms¹; enhanced spectral detection²

1. Introduction

Since optical devices can only detect those particles which fall within their viewing volume, a capability to concentrate hazardous biological particles into such a volume should greatly enhance the detection sensitivity of such devices. This work aims at the concentration of 1-micron-size micro-organisms from about 100 ml of liquid, whether drawn from a bio-aerosol collector or from an environmental source, into a volume of 1 to 2 ml. Extension of the same approach to the concentration of smaller (0.1-micron-size) viral particles is also under way.

2. Background

As part of a proprietary bio-aerosol collector based on Wet Electrostatic Precipitation [WEP] (1), shown in Fig. 1, we provide a liquid flow system including a reversible filter through which filtered liquid can be re-circulated or disposed of and from which a concentrated sample is recovered by opening a solenoid valve leading to a detector or to a collection container and reversing the pump for a short time. The reversing action flushes the collected particles off the filter and into the detector or container. The components of this water management system are exhibited in Fig. 2 and packaged in the self-contained unit of Fig. 3.

3. Results

Using the system of Fig. 2, we obtained the results of Figs. 4-7, which illustrate the effectiveness of our approach. Fig. 4 shows the cumulative fluorescence of 1-micron-size blue fluorescent micro-spheres versus cumulative volume withdrawn from our WEP collector after capture from an aerosol suspension of about 140 particles/ml drawn from a test chamber over a 5-minute period at a sampling rate of 500 liters/minute. As the liquid was being withdrawn from the WEP collector, it was filtered at a rate of 1 ml/second and the filter back-flushed with small portions of filtrate at 1-minute intervals. Figs. 5-7 show



Fig. 1. View of the WEP Bio-Aerosol Collector in Operation [Carrying Size: 20.7" x 17.2" x 8.6"]

the concentrations of the captured particles in successive fractions of the withdrawn liquid. As can be seen from Figs. 6 and 7, the relative cumulative concentration of captured particles in the first back-flushed 1-ml fraction was around 56 as compared with a value of 1.4 in the first 60-ml of to be filtered fraction, which constitutes a concentration enhancement by a factor of 40.

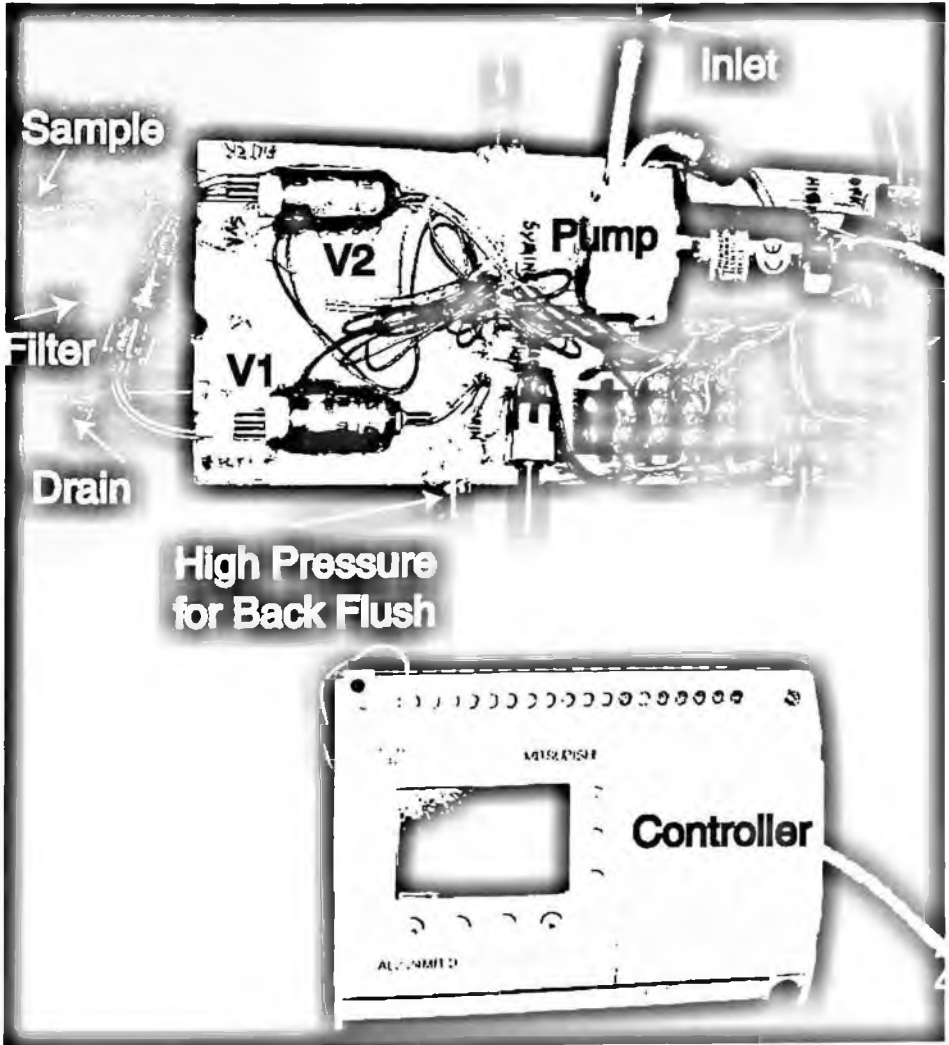


Fig. 2. Assembled Components for Water Management and Sample Collection

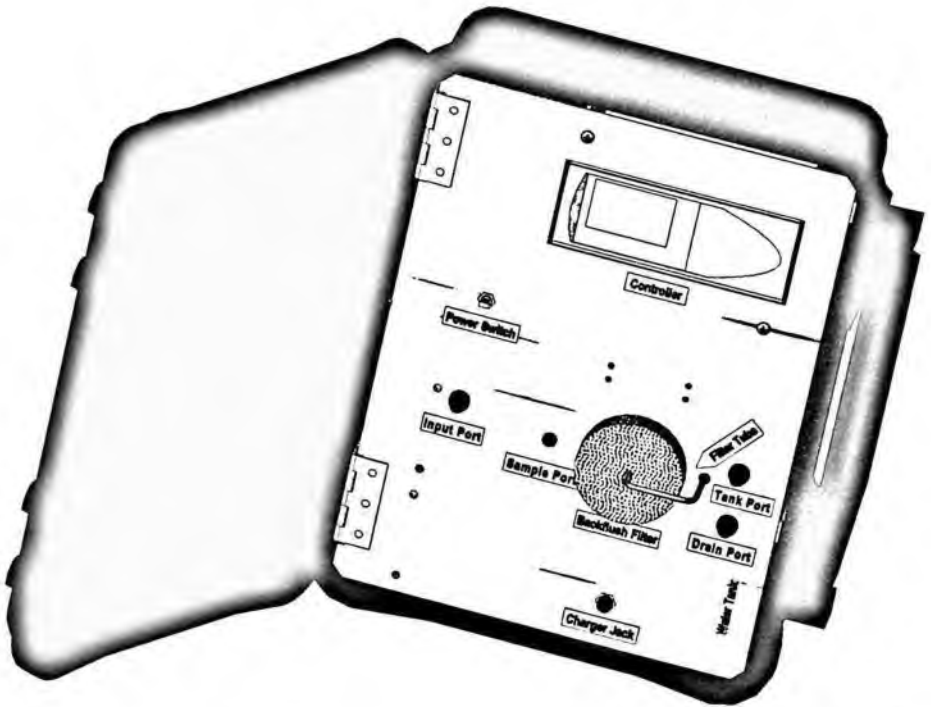


Fig. 3. Fully Self-Contained Battery-Powered Unit for Concentrating Micro-Organisms from 100 ml of Liquid into a 1.5-ml Volume [Carrying Size: 13.4" x 11.6" x 6"]

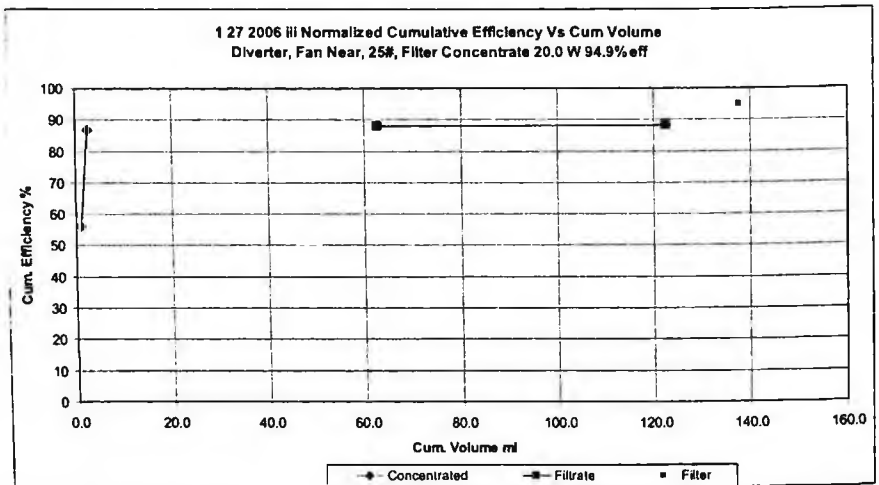


Fig. 4. Normalized Cumulative Collection Efficiency vs. Cumulative Volume

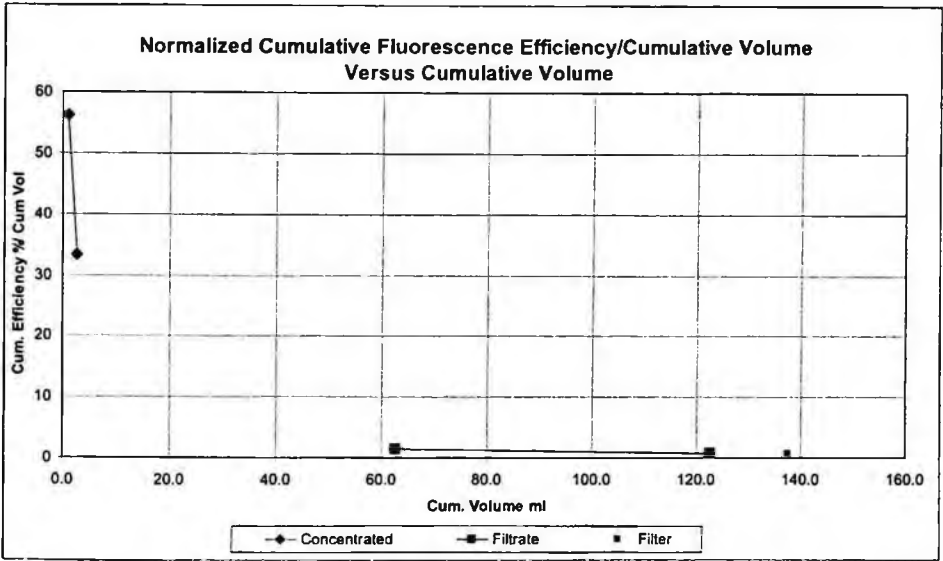


Fig. 5. Normalized Cumulative Collection Efficiency Divided by Cumulative Volume versus Cumulative Volume

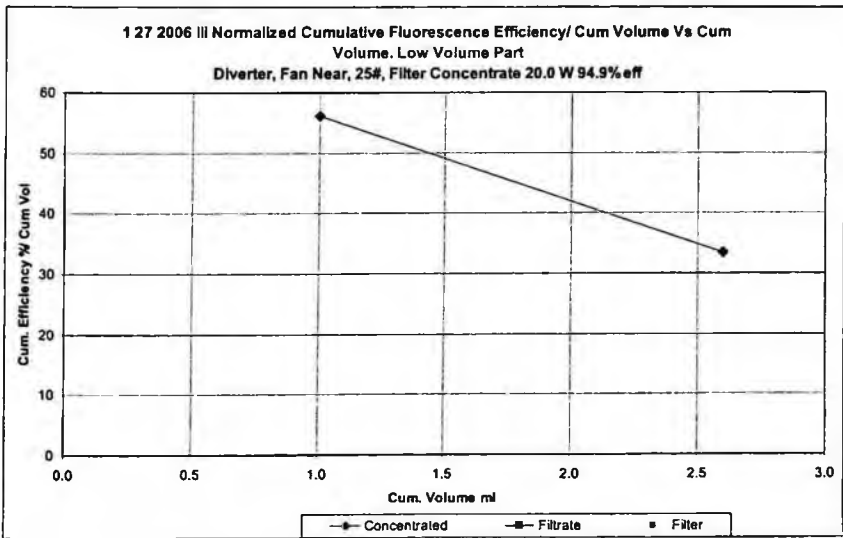


Fig. 6. Same as Fig. 5, but With an Expanded Low-Volume Scale

Normalized Cumulative Fluorescence Efficiency/Cumulative Volume
Versus Cumulative Volume – High Volume Part

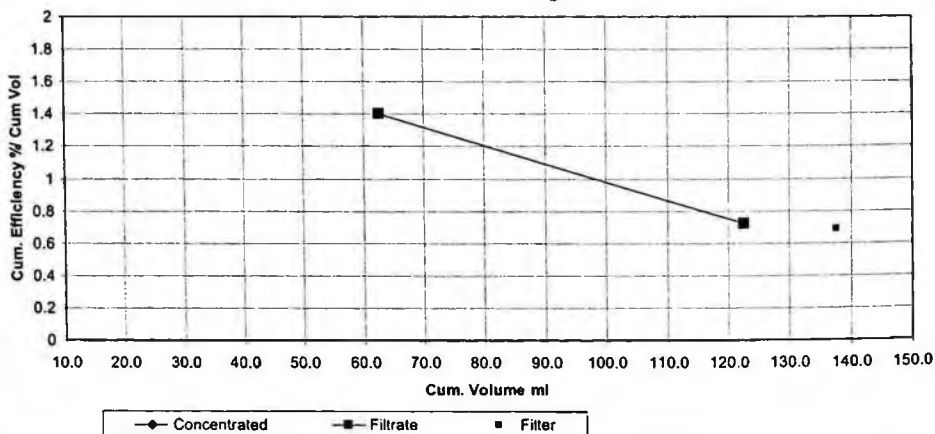


Fig. 7. Same as Fig. 5, but With an Expanded High-Volume Scale

It can also be seen from Fig. 4 that >50% of all the captured fluorescent beads was recovered in the first back-flushed 1-ml fraction from the first 60-ml filtered fraction and >80% was recovered within a total volume of 2.6 ml in the first two back-flushed fractions (see Figs. 4 and 6). Since the collection into the first back flushed 1-ml sample amounted to >50% of the sampled aerosol, the resulting concentration factor is >50% \times 500,000 [ml/min] \times 5 min/1 ml = $>1.25 \times 10^6$.

3.1 Back Flushing with BG Spores

Some confidence in the applicability of our measurements to that of concentrating bacterial cells may be gained from the following measurements with BG spores. Although our previous attempts to measure concentrations of suspended spores by a "live-dead" procedure gave spurious results, we were able to measure total BG spore concentration by transmission photometry. To obtain valid data, we adjusted the starting spore concentrations to fall within a narrow range in which the light absorption was neither too low nor too high to fall within the easily readable scale. To measure more concentrated samples, successive dilutions were performed until the concentration was readily measurable. We could thus test our back flushing approach using the same filter (25-mm Nalgene 190-2545 0.45 μ m Cellulose Acetate) as the one that yielded acceptable results with the fluorescent 1-micron microspheres that were used throughout our project. After passing through that filter 100 ml of a 10^{-3} weight-% BG suspension in distilled water, we collected successive back flushed samples and measured their BG

concentrations as shown in Figs. 8 and 9. The retrieval of 40% of the BG contents of the 100-ml starting suspension into a volume of only 1.75 ml is tantamount to a concentration of the original suspension by a factor of $40\% \times 100 \text{ ml} / 1.75 \text{ ml} = 23$.

4. CONCLUSIONS

The apparatus of Figs. 2 and 3 effectively concentrates micron-size micro-organisms in water samples from the WEP or other bio-aerosol collectors or environmental sources. Current efforts are directed at adapting it to the concentration of toxin virus particles.

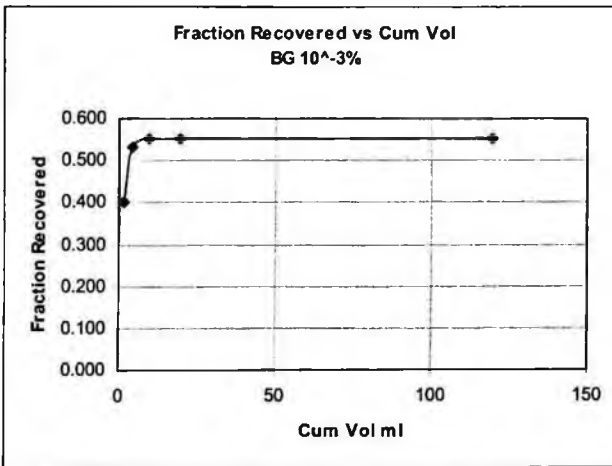


Fig. 8. Fraction of BG Retrieved from 100-ml Suspension

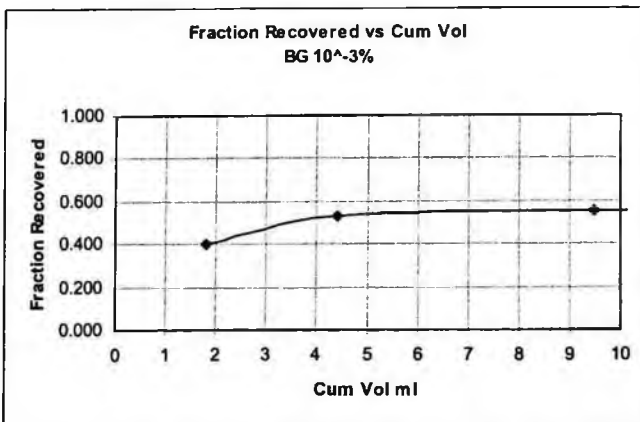


Fig. 9. Same as Fig. 8 with an Expanded Low-Volume Scale

5. REFERENCES

1. S. Zaromb, D. Martell, N. Schatke, and G. Hankins, Electrostatic-Precipitation-Based Aerosol Collector, *2006 Scientific Conference on Obscuration and Aerosol Research* (Battelle Eastern Science & Technology Center, Aberdeen, Maryland, 2006).

ISSR 2006

FRONTIER SESSION

PROGRAMMABLE ADAPTIVE SPECTRAL IMAGERS FOR MISSION-SPECIFIC APPLICATION IN CHEMICAL/BIOLOGICAL SENSING

NEIL GOLDSTEIN, PAJO VUJKOVIC-CVIJIN, MARSHA FOX, STEVEN ADLER-GOLDEN,
JASON CLINE, BRIAN GREGOR, AND JAMINE LEE

*Spectral Sciences, Inc. 4 4th Avenue
Burlington, MA 01803, USA
neil@spectral.com*

ALAN C. SAMUELS

*Edgewood Chemical Biological Center
AMSRD-ECP-RT-DP(Building E3330/56)
5183 Blackhawk Road
Edgewood Area, Aberdeen Proving Ground, MD 21010-5424, USA*

SHAWN D. HIGBEE

*Air Force Research Laboratory
29 Randolph Road
Hanscom AFB, MA 01730, USA*

LATIKA S. BECKER

*USA SMDC, P.O. Box 1500,
Huntsville AL 36907, USA*

TENG OOI

*US Army Aviation and Missile Research, Development, and Engineering Center,
Redstone Arsenal, AL 35898, USA
and Department of Aeronautics and Astronautics, Stanford University,
Stanford, CA 9430, USA*

An innovative passive standoff system for the detection of chemical/biological agents is described. The spectral, temporal and spatial resolution of the data collected are all adjustable in real time, making it possible to keep the tradeoff between the sensor operating parameters at optimum at all times. The instrument contains no macro-scale moving parts and is therefore an excellent candidate for the development of a robust, compact, lightweight and low-power-consumption sensor. The design can also serve as a basis for a wide variety of spectral instruments operating in the visible, NIR, MWIR, and LWIR to be used for surveillance, process control, and biomedical applications.

Keywords: chemical/biological; spectral imaging; remote sensing.

1. Introduction

There is a long-standing need for standoff sensors capable to detect, identify and quantify chemical and biological warfare agents and other hazardous species. Currently dominant Fourier transform infrared spectrometers (FTIR) are based on interferometry, and are therefore mechanically complex, expensive, and susceptible to mechanical disturbances. Adaptive multispectral sensors, based on mathematical transforms other than Fourier's which do not require macroscopic moving parts to provide the same spectral multiplexing capability (Felgett's advantage) as FTIR instruments is described. The technology combines a unique on-the-fly spectral adaptability with data acquisition and processing speeds high above those of its FTIR counterparts. The transition from a traditional nonimaging sensor to a multispectral imaging spectrometer is conceptually and technologically simple with adaptive multispectral sensors.

The adaptive spectrometer described in this paper produces either non-imaging, one-dimensional or two-dimensional multispectral radiance datasets ("datacube" in the case of two-dimensional spectral mapping) for gas or aerosol discrimination and classification. The spectral, temporal and spatial resolution of the data collected by the instrument are adjustable in real time, making it possible to keep the tradeoff between sensor parameters at optimum at all times. The instrument contains no macro-scale moving parts making it an excellent candidate for the development of a robust, compact, lightweight and low-power-consumption device suitable for field operation.

Potential uses of the AS include early standoff detection of chemical/biological agents. Extensive research into both passive remote chemical/biological sensors and active laser-based (LIDAR) systems has been carried out over the past several decades in government, academic and commercial research institutions.^{1,2} Challenges in building a lightweight, compact and rugged standoff sensing system include the complexity of the optical system, alignment sensitivity when moving parts are involved (e.g. interferometer of a FTIR instrument) and complexity of data interpretation for the passive system, and laser tuning, energy and pulse rate requirements of the laser source in active systems.

The key innovation in passive dispersive (as opposed to interferometric) spectrometers introduced recently is the use of a programmable digital spatial light modulator (SLM) to encode spectral information.^{3,4,5,6} The most advanced version of SLM currently is a reflective digital micromirror array (DMA).^{7,8} The feasibility of DMA-based multiplexed spectrometers has been experimentally demonstrated recently.^{9,10,11}

2. Instrument Concept

The principles of operation of the Adaptive Spectral Imager are illustrated on Fig. 1.^{11,12,13} The instrument uses two polychromators (or a single polychromator in two passes) and a spatial light modulator (SLM) placed in the intermediate focal plane

between them. The first polychromator disperses the image into spectral components and images them onto the SLM. The SLM spectrally encodes the image by modulating the intensity of selected spectral bands in the dispersed image. The encoded image is then spectrally recombined by the second polychromator and imaged onto a photodetector (Fig. 1).

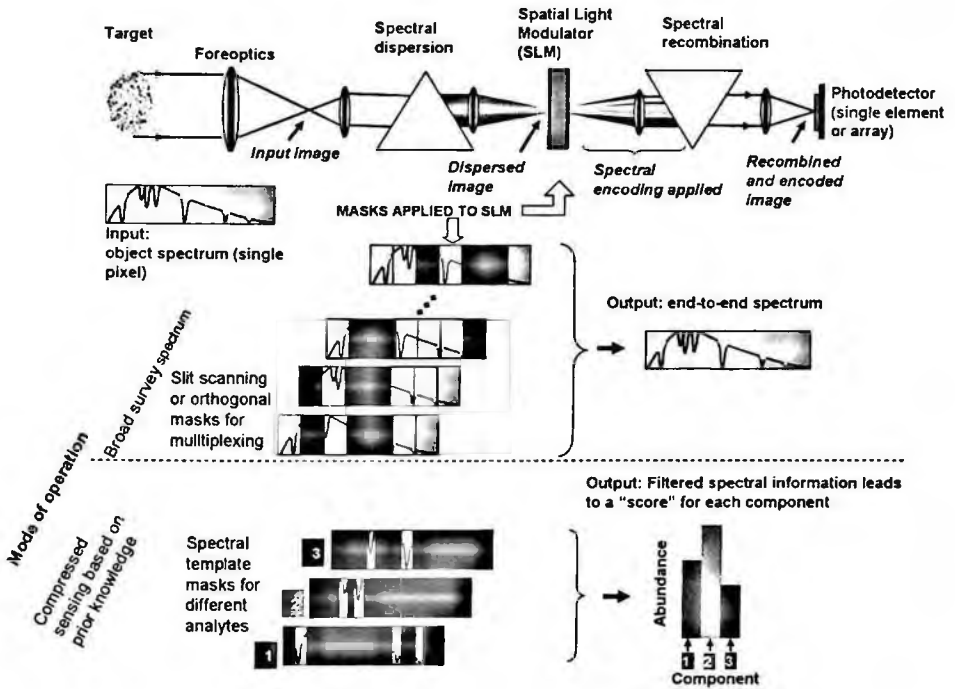


Fig. 1. Adaptive spectrograph concept: light from a standoff spectral scene is dispersed, dynamically encoded with a series of masks, and recombined on the detector.

Spectral encoding is achieved by transmission, rejection or intensity modulation of particular bands. This operation can be easily accomplished with a programmable digitally controlled two-dimensional SLM. Spectral encoding of this sort can be seen as equivalent to applying a certain mathematical transform to the spectrally resolved image. The transform can be the extraction of a value within a fixed window (spectral filter), a single sliding window (a scanning slit), a number of simultaneously applied slits conforming to the Hadamard transform in order to achieve multiplexed channel acquisition, an application-specific matched filter based on the prior knowledge about the analyte and the background, etc. In any case, the two-dimensional SLM whose pixels are individually addressable (and small relative to the resolution element) is programmed to generate a dynamic series of spatial patterns (masks) that implement the transform functions. Certain transform functions, such as matched filters, principal component

analysis, and similar, can simplify data processing. Rather than indiscriminately collecting raw spectral information for post-processing, these techniques use programmable hardware to select distinctive meaningful components of the total available spectral signal and reject the spurious or information-lacking components. In this sense, the technique described performs in-hardware compressed spectral sensing. The consequences of the approach include better signal-to-noise ratio (SNR) and better economy of data acquisition, storing and processing. The latter is of no small importance for hyperspectral and multispectral imaging.

The adaptive spectrometer described here uses a DMA to spectrally encode the signal. For chemical/biological sensing, the spectrometer operates in the long wavelength infrared (LWIR) spectral range (8-12 μm). A single-element detector, a linear array or a two-dimensional focal plane array (FPA) may be used to acquire the signal, leading to different functionalities described below.

In the case of an unknown analyte and an unknown background, survey “end-to-end” spectra need to be acquired. This mode of operation is efficiently implemented with adaptive spectral imagers by using an orthogonal set of DMA masks that perform Hadamard transform of the spectrum.¹⁴ For Hadamard transform, the DMA is programmed with a series of binary masks based on Simplex (S) matrices, where elements 0 and 1 translate into on-off modulation of individual micromirror pixels of the DMA.^{3,4} Recording a N -channel spectrum in this way requires N masks based on a set of N matrices resulting in $N^{1/2}/2$ multiplex (Felgett’s) advantage in the signal to noise ratio (SNR) relative to the simple slit-scanning approach. A Hadamard-transform spectrometer collects the light contained in approximately half of all the wavelength channels simultaneously at any point in time, thereby providing the much needed photon collection efficiency in the LWIR. The recorded images are decoded by the inverse Hadamard transform to provide the spectrum for each individual pixel (Fig. 1).

For chemical/biological detection it is common to have prior knowledge of the spectral features to be detected, at least on the part of the analyte. An adaptive spectrometer can use this information to achieve fast and sensitive detection by applying target-specific and background-specific matched-filter masks to the DMA spectral modulator. Such masks can provide analyte-specific spectral selectivity for a single column within a two-dimensional image. An adaptive spectrometer can be programmed with matched filters directly on the DMA to detect substances on which prior knowledge of spectral signatures exists. Furthermore, the spectrometer can be programmed so that successive DMA masks match different analytes to be monitored virtually simultaneously. Due to its high modulation bandwidth, the spectrometer can easily cycle through thousands of masks per second and provide rapid multi-analyte screening. Detection and identification is achieved when the integrated intensity of the signal through a given mask exceeds a threshold. The use of a single matched filter per analyte in the receiver

hardware rather than time-consuming end-to-end spectrum acquisition also reduces the signal susceptibility to artifacts caused by source fluctuation or motion and vibration of the sensor platform. The data collection time per image frame is pushed to the millisecond and sub-millisecond scale with the DMA frame rate in the 10kHz-40kHz range, currently available.¹⁵ Cycling through DMA masks at such high speed and using compressed sensing provides a critical advantage for the Adaptive Spectral Imager relative to other imaging spectrometers for applications aboard moving instrument platforms or against moving targets.

3. Instrument Design for a Passive Sensor Based on Adaptive Spectral Imager with a Two-Dimensional Focal Plane Array

Fig. 2 shows a schematic representation of an optical system for the Adaptive Imaging Spectrometer containing a core spectrograph and foreoptics. The foreoptics, which may be a standard IR camera lens or telescope, forms an image of the scene at the entrance of the core spectrograph. The spectrometer uses a novel dual-pass spectrometer design with a single concave grating to both disperse and recombine the light. We have found that we can achieve near-diffraction limited image quality with this arrangement.¹⁶ The image delivered to the input of the spectrograph is dispersed and reimaged onto the DMA spectral modulator. The DMA consists of an array (typically 1024 x 768)¹⁵ of addressable micromirrors. The grating subsequently also spectrally recombines the light and re-images it on a two-dimensional FPA detector.

The spatial light modulator used in our system is a modified DMA chip by Texas Instruments (DMD 0.7 XGA 12° DDR).¹⁵ The array consists of 1024 x 768 aluminum coated individual micromirrors of 14 μ m x 14 μ m each, capable of $\pm 12^\circ$ tilt relative to the diagonal. Mirrors are individually addressable through a driver circuit available from the manufacturer. For IR applications, the original glass window needs to be replaced with an IR-transparent window. We have developed the technology to mount coated ZnSe windows on Texas Instruments DMA chips and preserve long lifetime of the chip.

Our preliminary analysis indicates that we can build an adaptive imaging spectrometer with superb imaging properties and high sensitivity towards chemical/biological agents with 180mm diameter of the primary mirror in the Offner design, with an effective F-number of F/2.8. The input field of this staring spectrometer is 8 x 8 mm in the object plane, with diffraction limited performance achieved over the entire field. This results in approximately 100 spectral resolution units and 10,000 spatial resolution units. Assuming that the image generated on the FPA fills a detector pixel, and a typical LWIR FPA sensitivity, radiometric calculations for our LWIR adaptive spectral imager suggest the signal-to-noise ratio (SNR) around 25dB per collected spectrum, high enough to characterize a distant chemical/biological agent cloud.

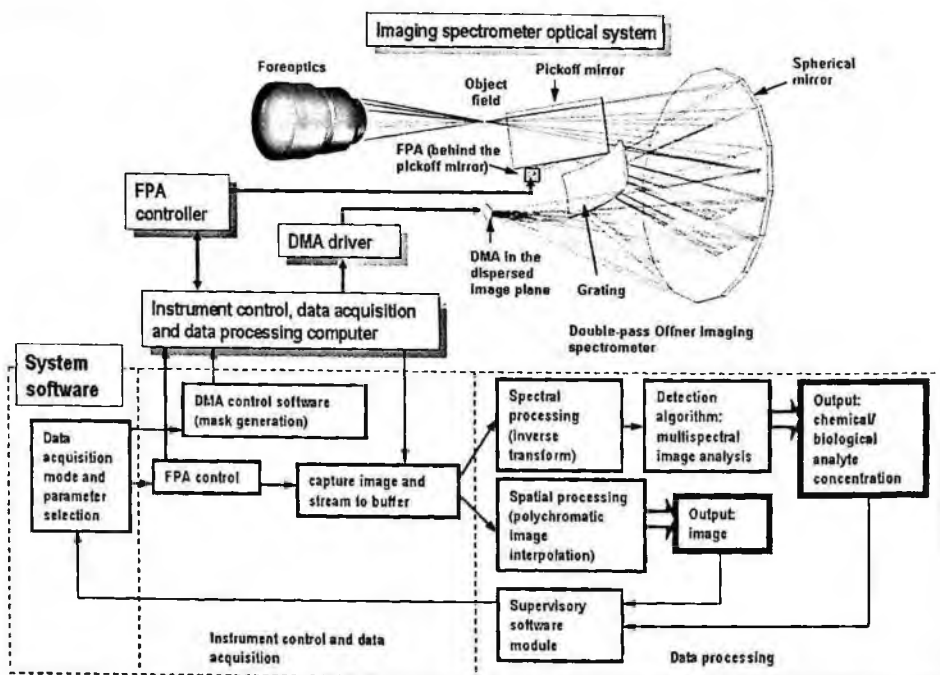


Fig. 2. System design.

The conceptual block diagram of the software package that controls the instrument, collects data, and delivers spatial and spectral information to higher-level external programs is presented on Fig. 2. The data stream captured by the FPA is routed to parallel spatial and spectral processing. The spatial image is based on the integrated intensity of all open spectral channels (about half of the total for the Hadamard transform). In order to make this image close as much as possible to the panchromatic image, additional simple processing can be applied in the form of interpolation and fitting pixel intensities. The output of this procedure is a high quality real time two-dimensional imagery.

In the case of collecting wide end-to-end spectra with the Hadamard transform-based modulation, the spectral processing software starts with pixel-by-pixel processing of the data. The reduction of Hadamard encodegrams into spectra is accomplished by applying the inverse Hadamard transform.¹³ While this operation in a general case could be rather time-consuming due to the large number of matrix inversions involved, the inversion of binary square cyclic S -matrices is extremely efficient and fast (actually, orders of magnitude faster than Fast Fourier Transform (FFT) calculations). The resulting datacubes are arranged in a band-sequential format, standard for processing multispectral and hyperspectral imagery.

4. Experimental Verification

A feasibility prototype of the adaptive spectral imager has been demonstrated in the visible/near infrared 600nm – 1,050nm spectral region, and in the shortwave infrared region 1,300-1,600nm. No experimental demonstration of this instrument has yet been done in the LWIR. For the experiments described here a silicon CCD camera and a single element InGaAs detector were used, respectively. An optical system based on refractive and transmissive components was assembled for the purpose. Data acquisition and processing software was developed for the control of the DMA, capturing data from the photodetectors, and for data reduction. Spectral calibration with known input wavelength(s) produced the instrument calibration factor of 0.41nm per micromirror of the DMA array.

Fig. 3 shows a simple optical setup used to test the collection of Hadamard spectra and spectrally filtered radiation in a manner analogous to that used in a chemical-biological detection system. Note that it uses only a single detector, and relies on the programmable DMA to provide both spectral and spatial resolution. Each column of the DMA corresponds to a wavelength band, and each row of the DMA corresponds to a one-dimensional spatial element. A set of optical fibers is placed at the input to provide three unique spectral signals, which can be differentiated by selecting rows of the DMA. Fig. 4 shows the spectra of each of the fibers. The top fiber contains narrow radiation obtained from a laser diode, the bottom fiber contains light from a hot black-body lamp which is passed through a water absorption cell. The middle fiber contains lamp radiation and laser diode radiation.

Fig. 5 compares spectra and the standard deviation of the spectra (STD) obtained by Hadamard modulation spectroscopy with that obtained by scanning a slit. The spectra are identical, but the noise level is reduced in the Hadamard spectrum due to the higher light collection efficiency (multiplex advantage). In each case, the data is obtained by scanning a slit pattern on the DMA through 127 positions. In the slit-scan case, each exposure contains light from a single spectral band, while in the case of the Hadamard spectra, each exposure contains light collected from 64 spectral bands. After decoding, the Hadamard spectrum shows the expected signal to noise improvement of $(m/2)^{1/2}=6$, where $m=127$ is the number of resolution units in the spectrum.¹³

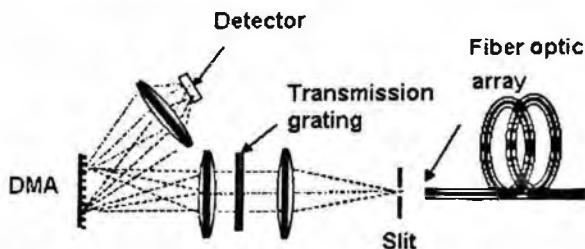


Fig. 3. Experimental setup for single-detector experiments.

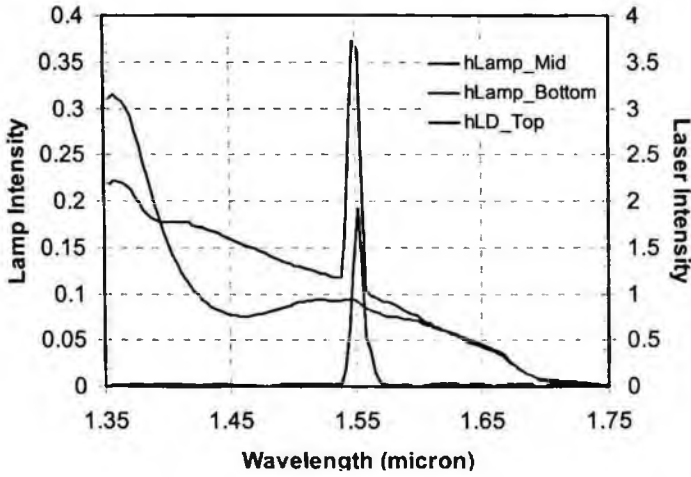


Fig. 4. Hadamard spectra of three spatial regions: top, (dark blue), middle, (red), and bottom, (light blue).

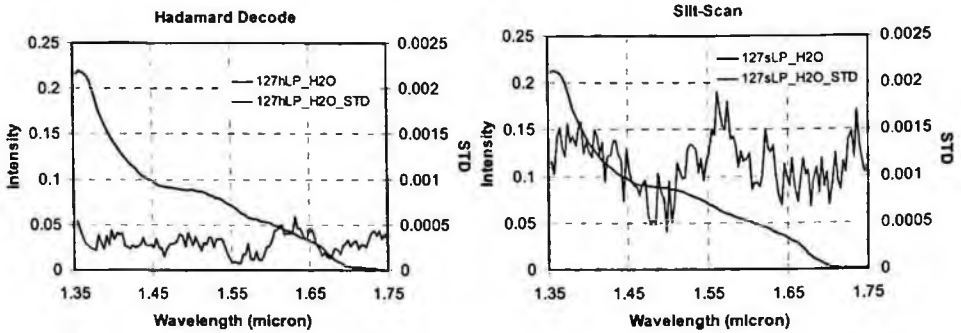


Fig. 5. Comparison of slit-scan and Hadamard spectra.

The same system was used to record temporal data from a H_2 /air flame using spectral band-pass filters.¹⁷ Fig. 6 shows water flame spectra obtained under two different flame conditions: a near stoichiometric flame and a lean flame. The spectral data match the known theoretical spectra of water. The width of the spectral band is an indicator of water temperature. Thus, we can determine the density and temperature of the water vapor within the flame using a set of three spectral bands corresponding to the center and wing of the water band, and a null band outside of the water radiation band.

Fig. 7 shows how the detected intensity in each band changes with the condition of the flame. The intensity of both the center and wing band follow the water density and temperature. The ratio of the two bands follows the water temperature. Comparing the two bands yields the column density and temperature of water as a function of fuel flow rate.

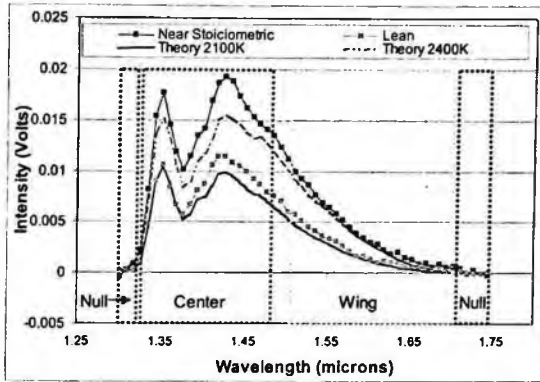


Fig. 6. Spectral of H^2 /air flame showing spectral bands used for spectrally selected monitoring.

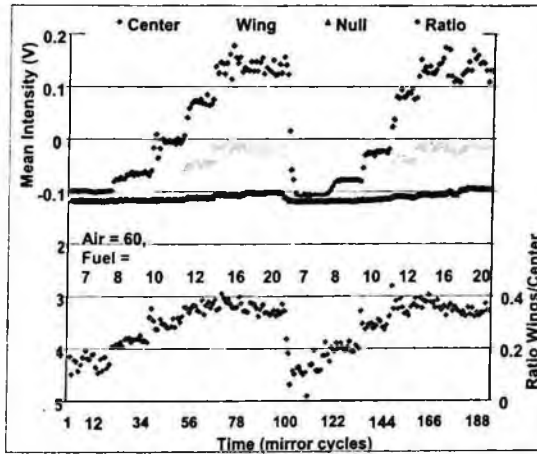


Fig. 7. Time resolved data for each spectral band.

A similar setup was used to demonstrate full two-dimensional imaging. The single detector was replaced with a second spectrograph and a CCD camera, which recorded data in the spectral range of 600-1050 nm. The imaging spectral capability of the prototype was investigated by using three blackbody objects at different temperatures placed into the instrument's field of view (Fig. 8). The blackbodies were represented by electrically heated resistors kept at constant temperature. One of the sources used (Object 1, Fig. 8) was a calibrated 1250K blackbody source. A seven-element Hadamard mask was used to produce seven-band spectra of each object with adequate spectral resolution for temperature measurement at a fast update rate. Fig. 8 shows a dynamic sequence of data collected with our adaptive spectral imager. The left-hand side shows the time sequence of spatial (spectrally combined) images, as generated after each Hadamard

transform mask was applied to the DMA. Since half of all the spectral channels are open at any moment in time, and since the FPA readout is updated after every mask, the spectrally recombined image on the FPA shows a near-panchromatic view of the objects. The right-hand side of Fig. 8 shows the raw (instrument function included) Hadamard-decoded spectra of three objects, updated partially after each mask, and updated fully after each complete cycle of 7 masks. Data are scaled by an approximate instrument function, to compensate for the detector falloff at wavelengths greater than 900 nm. Fig. 9 shows the same data along with a fit to the Planck function that reveals the approximate temperature of the object. Note that the vertical (signal intensity) scale does not represent the radiance of the object, since the solid angle has not been calibrated. Based on this approach we found the approximate temperatures of the three objects to be 1400K, 1600K and 1900K (Fig. 9).

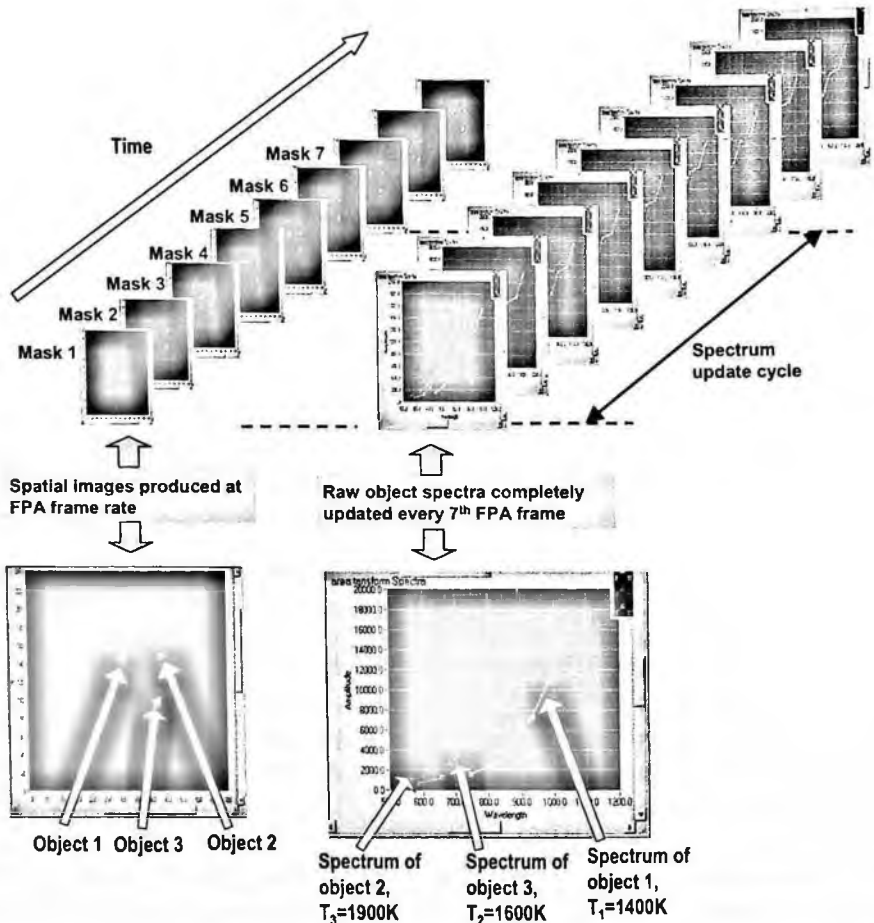


Fig. 8. Data acquisition dynamics for adaptive imaging spectrometer as demonstrated by spectral imaging objects at different temperatures within the field of view of the instrument.

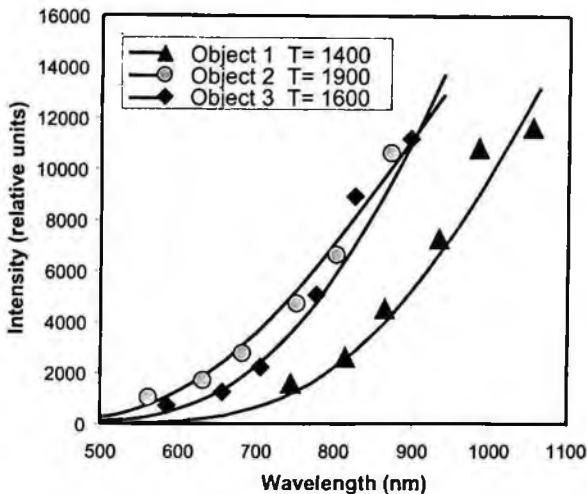


Fig. 9. The spectra of three objects in the field of view of the imaging spectrometer, together with blackbody curves fitted to the temperatures indicated.

5. Conclusions

A class of adaptive spectrometers or multispectral imagers capable of performing a variety of spectral transforms with real-time adjustment of data acquisition parameters is described. Fast adaptive spectral imagers based on this approach can greatly improve the performance of standoff chemical/biological sensors by allowing real-time optimization based on prior information, the amount and type of the information required, and the data rate. The instrument's hardware architecture appears superior to FTIR technology, since no macroscopic moving parts are needed. The technology is also applicable to spectrometers and spectral imagers in diverse areas such as surveillance, process control, and biomedical imaging, and can be adapted for use in any spectral domain from the ultraviolet (UV) to the LWIR region.

6. References

1. T. G. Quinn, R. Gross, J. T. Ditillo, W. M. Lagna, Improved Lightweight Standoff Chemical Agent Detection Sensor, www.blockeng.com/paper1.htm, *Proc. SPIE Electrooptical Technology for Remote Chemical Detection and Identification*, (1997).
2. J. T. Ditillo, R. L. Gross, M. L. Althouse, W. M. Lagna, W. R. Loerop, P. J. Deluca, T. G. Quinn, and L. B. Grim, "Lightweight standoff chemical agent detector," *Proceedings of SPIE* Volume: 2366, p. 166, (1995).
3. M. K. Bellamy, A. N. Mortensen, R. M. Hammaker, and W. G. Fateley, Chemical Mapping in the Mid- and Near-IR Spectral Regions by Hadamard Transform/FT-IR Spectrometry, *Appl. Spectrosc.*, 51, 477-486 (1997).
4. R. A. DeVerse, R. M. Hammaker, and W.G. Fateley, An Improved Hadamard Encoding Mask for Multiplexed Raman Imaging Using Single Channel Detection, *J. Mol. Struct.*, 521, 77-88 (2000).

5. R. Riesenberg, and U. Dillner, HADAMARD Imaging Spectrometers, *SPIE's 44th Annual Meeting & Exhibition, Optical Science, Engineering and Instrumentation*, Denver, USA, Proceedings of SPIE, Vol. 3753, pp. 203-213, (18-23 July 1999).
6. A. Wuttig, and R. Riesenberg, Sensitive Hadamard Transform Imaging Spectrometer with a simple MEMS, *Proc. SPIE*, **4881**, pp. 167-178, (2003).
7. L. J. Hornbeck, Deformable Mirror Spatial Light Modulator, *Proc. SPIE*, Vol. 1150, p.86 (1989).
8. L. J. Hornbeck, Digital Light Processing™ for High-Brightness, High-Resolution Applications, 21st Archives, VXM Network, www.vxm.com (1998).
9. C.M. Wehlburg, J. C. Wehlburg, S. M. Gentry, and J. L Smith, Optimization and characterization of an imaging Hadamard spectrometer, *Proc. SPIE*, **4381**, p. 506-515, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII, Sylvia S. Shen, Michael R. Descour, Eds., (2001).
10. P. Vujkovic-Cvijin, Infrared Derivative Spectroscopy For Open Path Sensing, TR-467, Phase I Final Report, Spectral Sciences contract 2923.1 (2006).
11. N. Goldstein and P. Vujkovic-Cvijin, Adaptive Spectral Imager, TR-447 Phase I Final Report, Spectral Sciences contract 2923.1 (2005).
12. P. Vujkovic-Cvijin, N. Goldstein, , M. J. Fox, S. D. Higbee, S. Latika L. C. Becker, K. Teng and T.K Ooi., Adaptive Spectral Imager for Space-Based Sensing, *Proc. SPIE* Vol. 6206, paper 6206-33 (2006).
13. N. Goldstein, "Chemical Biological Agent Standoff Detection" TR-476 Phase I Final Report, Spectral Sciences contract 2911.1 (2006)
14. M. Harwit and N. J. A. Sloane, *Hadamard Transform Optics*, Academic Press, NY, 1979.
15. Texas Instruments "DMD Discovery Chip Set" Product Brochure, <http://www.tyrexsales.com/dlp/>, (2005).
16. N. Goldstein, P. Vujkovic-Cvijin, M. J. Fox, S. Adler-Golden, J. Lee, J. A. Cline, B. Gregor, Spectral Encoder, US Patent Application, (2006).
17. J. Cline, FASPEC – Fast Adaptive Spectrometer Programmable for the Evaluation of Combustion, TR-463 Phase I Final Report, Spectral Sciences contract 2883.1 (2005)

DESIGN & OPTIMIZATION OF AN I-RTD HYBRID THz OSCILLATOR BASED UPON $\text{In}_{1-x}\text{Ga}_x\text{As} / \text{GaSb}_y\text{As}_{1-y}$ HETEROSTRUCTURE SYSTEMS

WEIDONG ZHANG

*Dept. of ECE, NC State University
Raleigh, NC 27695
wzhang5@ncsu.edu*

DWIGHT WOOLARD

*U.S. Army Research office
RTP, NC 27709 USA
dwight.woolard@us.army.mil*

BORIS GELMONT

*University of Virginia
Charlottesville VA 22903, USA
gb7k@virginia.edu*

ELLIOTT BROWN

*University of California, Santa Barbara
Santa Barbara, CA 93106, USA
erbrown@ece.ucsb.edu*

ROBERT TREW

*North Carolina State University
Raleigh, NC 27695, USA
trew@ncsu.edu*

An $\text{In}_{1-x}\text{Ga}_x\text{As} / \text{GaSb}_y\text{As}_{1-y}$ hetero-system with staggered band-lineups as solid-state platform for design of an interband resonant double-barrier tunneling diode (I-RTD) based optically-pulsed (OT) hybrid device for generating THz oscillations is theoretically investigated. It is demonstrated that this optical I-RTD hybrid is compatible with the robust state-of-art 1.55 micron laser technology. Multi-band wave equations in the framework of six-band Kane's model are applied for understanding the carrier dynamics when strain-induced effects are present. Simulation results for practical circuit implementations clearly show the superiority of this new oscillator concept.

Keywords: Type-II, Interband Tunneling, Optical Triggering, Terahertz oscillator.

1. Introduction

In recent years, the potential of terahertz (THz) spectral sensing in an array of applications areas relevant to defense and security have been given increasing attention¹. A series of exploratory research investigations into THz phenomenology are being actively performed. A major hurdle that could prohibit future experimental studies and practical implementations of sensor systems is the general lack of a robust (i.e., adequate power and efficiency) source. In the electromagnetic spectrum, lower and higher frequency oscillations other than THz energy can be implemented well by utilizations of diode-transistor carrier transport and quantum optical transition, respectively, while effective solid-state approaches for THz sources are inadequate (although a few excellent low-power commercial products are being widely used such as diode multiplier and photo-mixing). The solid-state room-temperature THz oscillation source becomes a challenge from fundamental physics and new technological innovations are needed. In

the past, these facts have motivated the conceptualization and investigation of a novel type of Interband Resonant-Tunneling-Diode (I-RTD) THz oscillator by our research group²⁻³. Soon after an optical hybrid type of I-RTD approach was proposed.

Under certain conditions, in double-barrier resonant tunneling diode with properly-designed staggered-bandgap structure, a nanoscale local space-charge transfer from right barrier to spacer occurs due to interband tunneling arises. This charging of the right-barrier region in I-RTD electrostatically changes the spatial potential. The net diode current is modulated as a function of time correspondingly. The basic idea of the hybrid I-RTD oscillator is to combine the natural Zener tunneling with the artificial optical annihilation of VB well hole-charge by laser pulses to produce a two-phase charging-discharging transient that leads to relaxation oscillations. Optimally, the pulse repetition rates (PRF) needs to be equal to the THz oscillation.

In previous investigations⁴, the optical triggering (OT)-I-RTD oscillator based upon a specific type II semiconductor hetero-system AlGaSb/InAs/AlGaSb was quantitatively studied and its capability to produce significant values of THz power was demonstrated. However, the simulations indicate a mid-IR photonic-energy ($\sim 0.26\text{eV}$) with high intensity level is required for effective optically-triggering. It turns out that laser repetition rates at this long wavelength are much less than THz^{4,5}. To overcome this limitations, one turns to the more mature 1.55 micron laser where significantly larger PRFs are available owing to developments in telecommunication industry. The problem lies in that laser pulses of such short wavelength unavoidably generate electron-hole pairs in the narrow-band InAs collector (0.36eV), which interfere with the ultra-fast charging-discharging processes. Hence the goal of this work is to consider an alternative hetero-system (i.e., $\text{In}_{1-x}\text{Ga}_x\text{As} / \text{GaSb}_y\text{As}_{1-y}$) with staggered band-lineups and wider bandgaps ($> 0.8\text{ eV}$) that allow the $1.55\ \mu\text{m}$ operation. Prior experimental studies⁶⁻⁹ indicate that carefully selected compositions x, y provide degrees of freedoms on achieving specified characteristics such as the relative locations of band-edges, band gaps and offsets, thereby controllable quantum effects.

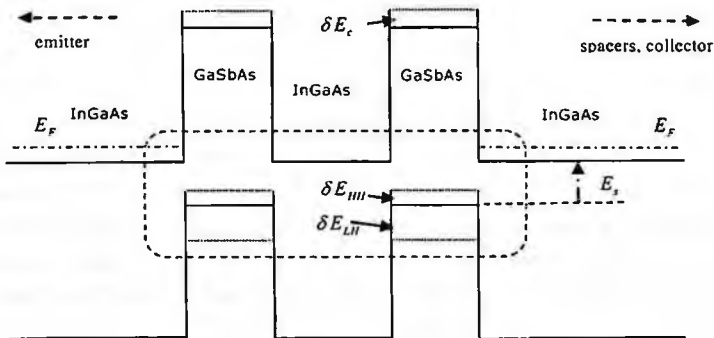


Fig. 1. The band diagram of $\text{In}_{1-x}\text{Ga}_x\text{As} / \text{GaSb}_y\text{As}_{1-y}$ interband resonant tunneling diode

2. The Hetero-structures

The band diagram for the device structure is given in Fig.1. On the left side is heavily-doped InGaAs emitter, while on the right side are undoped spacer (10nm), lightly doped spacer (50nm, $1 \times 10^{17}/\text{cm}^3$) and heavily-doped ($3 \times 10^{18}/\text{cm}^3$) collector, from left to right, respectively. An undoped InGaAs layer that is sandwiched between two GaSbAs layers (~2nm) acts as quantum well for electrons. The GaSbAs layers serve as barriers for electrons, wells for holes. Under an external bias voltage, there is conduction band (CB) electron resonant transport. When the ground heavy-hole level near the band-edge of right GaSbAs valence band (VB) well is raised above the Fermi level in the collector region, those VB electrons can tunnel through the forbidden-gap and produce trapped holes. The hole-charge can be discharged by external injection of ultra-short laser pulses. With sufficient optical intensity, at the end of each pulse/discharging phase, a large amount of holes are eliminated and the operation of device characterized by quantities such potential profile, current and charge storage is forced to return back to its initial state. Clearly the oscillation frequency is set by the PRFs, and provides frequency tuning.

In describing the semiconductor alloys InGaAs and GaSbAs combined together to form a hetero-interface, the locations of band-edges are labeled as E_{c1} , E_{v1} for InGaAs and E_{c2} , E_{v2} for GaSbAs. The dependencies of band-edges E_{c1} , E_{v1} , E_{c2} and E_{v2} on compositions x or y can be found in ¹⁰. For ternary compounds InGaAs/GaSbAs grown by the MBE method, the Vegard's law is obeyed, i.e. lattice constants for either semiconductor hosts depend linearly on compositions x or y ⁶. The lattice constants are denoted by symbols $a_1 = 6.0583 - 0.405x$ for InGaAs and $a_2 = 5.0532 + 0.4427y$ for GaSbAs¹⁰. The residual gap is $E_g = E_{c1} - E_{v2} = 0.68 + 0.31x - 2.02y + 0.43x^2 + 1.2y^2$. Simple calculations indicate that the lattice constant of GaSbAs, a_2 , need be larger than the lattice constant of InGaAs, a_1 , in order to obtain a small E_g (~0.1eV) ¹⁰. In this case, the in-plane strain in GaSbAs layers will be compressive (i.e. ϵ is negative).

Binary alloy InP (lattice constant 5.8687Å) is chosen to be the substrate. To ensure that the collector region is semi-transparent to 1.55 μm laser pulses, the composition x is chosen to be at least 0.52. For $x=0.52$, the bandgap of $\text{In}_{0.48}\text{Ga}_{0.52}\text{As}$ is $E_{g1} = 0.804$ eV at room temperature. The mismatch between $\text{In}_{0.48}\text{Ga}_{0.52}\text{As}$ ($a_1 = 5.8477\text{\AA}$) layers and the substrate InP is a negligible value 0.36%. For a composition $x=0.56$, the bandgap of $\text{In}_{0.44}\text{Ga}_{0.56}\text{As}$ is 0.847eV, the lattice constant is $a_1 = 5.8325\text{\AA}$ and the mismatch between $\text{In}_{0.44}\text{Ga}_{0.56}\text{As}$ layers and substrate is 0.6%. Hence we assume all emitter/well/spacer/collector/ InGaAs layers are unstrained in the modeling.

For this newly considered superlattice, the unusual feature lies in the fact that the valence-band in one host semiconductor GaSbAs (barriers) is actually located very close to the conduction-band bottom in the other host semiconductor InGaAs, thus a strong interaction exists between these bands ⁹ despite the fact that each individual compound has a relatively wide bandgap. Hence the description of quantum transport processes within this heterostructure is a multi-band problem. In studies presented here, multi-band

wave equations in the framework of six-band Kane model are applied for the understanding of carrier dynamics^{2,3}. Furthermore, strain-induced effects on individual band-edges and thereafter carrier transports are incorporated¹¹.

3. The Transport Physics

3.1. Six-band Pikus-Bir Hamiltonian

The derivations for the strained semiconductors follow closely the classical works of Pikus and Bir¹¹ (or refer to¹² for the valence-band). The position vector of a specific atom in the primitive cell in the undeformed crystal is described by $\vec{r} = (x, y, z)$. The position vector of the same atom in the same primitive cell in the deformed crystal is described by $\vec{r}' = (x', y', z')$. These two vectors are connected through the strain tensor $\vec{r}' = \vec{r} + \vec{\varepsilon} \cdot \vec{r}$, where both coordinates (x', y', z') and (x, y, z) use the original same Cartesian unit vectors \vec{e}_x , \vec{e}_y and \vec{e}_z . The momentum operator is associated by $\vec{p}' = \vec{p} - \vec{\varepsilon} \cdot \vec{p}$ with the momentum \vec{p}' for deformed crystal and the momentum \vec{p} for regular crystal. The wave vector is connected by $\vec{k}' = \vec{k} - \vec{\varepsilon} \cdot \vec{k}$ with the wave vector \vec{k}' for deformed crystal and the wave vector \vec{k} for regular crystal. For uniformly deformed crystal, the Schrödinger wave equation with the deformed coordinate \vec{r}' is,

$$\left[\frac{p'^2}{2m_0} + V(\vec{r}') \right] \psi_{n\vec{k}'}(\vec{r}') = E_n \psi_{n\vec{k}'}(\vec{r}'), \quad (1)$$

where m_0 is the mass of free electrons. The periodic property of crystal is well maintained under a small stress despite that the strain reduces the symmetry of crystal. The Bloch theorem is still applicable $\psi_{n\vec{k}'}(\vec{r}') = \exp(i\vec{k}' \cdot \vec{r}') u_{n\vec{k}'}(\vec{r}')$. Eq.(1) can be converted back to the equations of old system with undeformed crystal coordinate \vec{r} through coordinate and momentum connections listed above¹¹,

$$\left[\frac{p^2}{2m_0} + \frac{\hbar}{m_0} \vec{k} \cdot \hat{p} + \frac{\hbar^2 k^2}{2m_0} + V(\vec{r}) + \sum_{\alpha, \beta} D_{\varepsilon}^{\alpha\beta} \varepsilon_{\alpha\beta} - 2 \frac{\hbar}{m_0} \sum_{\alpha, \beta} k_{\alpha} \varepsilon_{\alpha\beta} p_{\beta} \right] u_{n\vec{k}}^{(s)}(\vec{r}) = E_n u_{n\vec{k}}^{(s)}(\vec{r}). \quad (2)$$

where $D_{\varepsilon}^{\alpha\beta} = V_{\alpha\beta} - p_{\alpha} p_{\beta} / m_0$ and $u_{n\vec{k}}^{(s)}(\vec{r})$ is the abbreviation of $u_{n\vec{k}}(\vec{r} + \vec{\varepsilon} \cdot \vec{r})$. Here and following, the indexes α , β run over subscripts x , y and z . $V_{\alpha\beta}$ is the expansion coefficient of the potential $V(\vec{r}') = V(\vec{r} + \vec{\varepsilon} \cdot \vec{r})$ at the neighborhood of old undeformed coordinate \vec{r} in a series in terms of $\vec{\varepsilon}$ ¹¹. $u_{n\vec{k}}^{(s)}(\vec{r})$ at $\vec{k} = \vec{k}_0 + \Delta\vec{k}$ can always be written as an expansions of the complete set of *unstrained* Bloch functions at $\vec{k} = \vec{k}_0 = 0$,

$$u_{n\vec{k}}^{(s)}(\vec{r}) = \sum_m a_{m,\vec{k}}^{(s)} u_{m,\vec{k}_0=0}(\vec{r}). \quad (3)$$

The Pikus-Bir Hamiltonian matrix elements among doubly degenerate CB Bloch functions and four-fold degenerate VB Bloch functions (for crystals of zinc blende symmetry) are of interest. The basis functions are chosen to be ($\vec{k} = 0$),

$$\begin{aligned}
u_1 &= \left| \frac{1}{2}, \frac{1}{2} \right\rangle = |S \uparrow\rangle, & u_2 &= \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = |S \downarrow\rangle, \\
u_3 &= \left| \frac{3}{2}, \frac{3}{2} \right\rangle = \frac{1}{\sqrt{2}} |(X+iY) \uparrow\rangle, & u_4 &= \left| \frac{3}{2}, \frac{1}{2} \right\rangle = -i\sqrt{\frac{2}{3}} |Z \uparrow\rangle + i\frac{1}{\sqrt{6}} |(X+iY) \downarrow\rangle, \\
u_5 &= \left| \frac{3}{2}, -\frac{1}{2} \right\rangle = \sqrt{\frac{2}{3}} |Z \downarrow\rangle + \frac{1}{\sqrt{6}} |(X-iY) \uparrow\rangle, & u_6 &= \left| \frac{3}{2}, -\frac{3}{2} \right\rangle = i\frac{1}{\sqrt{2}} |(X-iY) \downarrow\rangle.
\end{aligned} \quad (4)$$

where S, X, Y, Z denote the symmetry properties (transform like atomic functions $s, p_x, p_y,$ and p_z) of the solutions of the equation $H_0\phi_i = (\hat{p}^2/2m_0 + V)\phi_i = E_i\phi_i$, H_0 is the Hamiltonian of the electron in a periodic potential. The system of equations become,

$$\sum_j (H_{jj} - E\delta_{jj})f_j = 0, \quad \{j, j' = 1, 2, \dots, 6\}. \quad (5)$$

For matrix elements between CB and VB Bloch functions (i.e. $j = 1, 2$ and $j' = 3, 4, 5, 6$), following Kane's work, the mutual interaction between conduction and valence bands (the $\bar{k} \cdot \hat{p}$ interaction) are treated exactly¹³,

$$H_{jj'} = \langle u_j | \frac{\hbar}{m_0} \bar{k} \cdot \hat{p} | u_{j'} \rangle - 2 \frac{\hbar}{m_0} \sum_{\alpha, \beta} \varepsilon_{\alpha\beta} \langle u_j | k_\alpha \hat{p}_\beta | u_{j'} \rangle, \quad (6)$$

where the term $\sum_{\alpha, \beta} \langle u_j | D_{\alpha\beta}^{\alpha\beta} \varepsilon_{\alpha\beta} | u_{j'} \rangle = 0$ is dropped off because of parity. The Kane parameter P is defined as,

$$P = -\frac{i\hbar}{m_0} \langle S | \hat{p}_x | X \rangle = -\frac{i\hbar}{m_0} \langle S | \hat{p}_y | Y \rangle = -\frac{i\hbar}{m_0} \langle S | \hat{p}_z | Z \rangle, \quad (7)$$

Hence readily, as an example, the matrix element H_{13} is obtained from Eq.(6) and (7),

$$\begin{aligned}
H_{13} &= \frac{1}{\sqrt{2}} \frac{\hbar}{m_0} \left(k_x \langle S | \hat{p}_x | X \rangle + ik_y \langle S | \hat{p}_y | Y \rangle \right) - \frac{\sqrt{2}\hbar}{m_0} \sum_{\alpha} (\varepsilon_{\alpha x} k_\alpha \langle S | \hat{p}_x | X \rangle + i\varepsilon_{\alpha y} k_\alpha \langle S | \hat{p}_y | Y \rangle), \\
H_{13} &= i\frac{1}{\sqrt{2}} P \left[k_x - 2 \sum_{\alpha} (\varepsilon_{\alpha x} + i\varepsilon_{\alpha y}) k_\alpha \right].
\end{aligned} \quad (8)$$

Similarly for other CB-VB matrix elements,

$$\begin{aligned}
H_{14} &= \sqrt{\frac{2}{3}} P (k_x - 2 \sum_{\alpha} \varepsilon_{\alpha x} k_\alpha), & H_{15} &= i\frac{1}{\sqrt{6}} P \left[k_x - 2 \sum_{\alpha} (\varepsilon_{\alpha x} - i\varepsilon_{\alpha y}) k_\alpha \right], \\
H_{24} &= -\frac{1}{\sqrt{6}} P \left[k_x - 2 \sum_{\alpha} (\varepsilon_{\alpha x} + i\varepsilon_{\alpha y}) k_\alpha \right], & H_{25} &= i\sqrt{\frac{2}{3}} P \left(k_x - 2 \sum_{\alpha} \varepsilon_{\alpha x} k_\alpha \right), \\
H_{26} &= -\frac{1}{\sqrt{2}} P \left[k_x - 2 \sum_{\alpha} (\varepsilon_{\alpha x} - i\varepsilon_{\alpha y}) k_\alpha \right], & H_{16} &= H_{25} = 0.
\end{aligned} \quad (9)$$

where the complex wave vector number k_x is $k_x = k_x \pm ik_y$. As for matrix elements for individual degenerate bands (CB, VB), the effects from higher conduction bands and lower valence bands can be treated as perturbations by applying Lödwin's theorem¹⁴. According to the procedure, the states of the Hamiltonian system are divided into two classes (A) and (B). Here class (A) includes the conduction and valence bands while class (B) contains other remote (unperturbed) bands (states) with less importance. As for conduction band, following Lödwin's perturbation method,

$$H_{11} = E_c + \frac{\hbar^2 k^2}{2m_0} - 2\frac{\hbar}{m_0} \sum_{\alpha,\beta} \varepsilon_{\alpha\beta} \langle c | k_\alpha \hat{p}_\beta | c \rangle + \frac{\hbar^2}{m_0^2} \sum_{\alpha,\beta} k_\alpha k_\beta \sum_r^{(B)} \frac{\langle c | \hat{p}_\alpha | r \rangle \langle r | \hat{p}_\beta | c \rangle}{E_c - E_r} + \sum_{\alpha,\beta} \langle c | D_c^{\alpha\beta} \varepsilon_{\alpha\beta} | c \rangle, \quad (10)$$

where index r wanders through all bands in class (B), never equaling 1,2,...,6. $|c\rangle$ denotes the conduction band (u_1, u_2) and $|r\rangle$ denotes bands in class (B). E_c is the CB band edge. Because of the S symmetry properties of the conduction band, i.e. $|c\rangle = |S\rangle$,

$$H_{11} = E_c + \frac{\hbar^2 k^2}{2m_c} + a_c (\varepsilon_x + \varepsilon_y + \varepsilon_z), \quad (11)$$

with the mass m_c and deformation potential a_c defined as,

$$\frac{1}{2m_c} = \frac{1}{2m_0} + \frac{1}{m_0^2} \sum_r^{(B)} \frac{|\langle S | \hat{p}_x | r \rangle|^2}{E_c - E_r}, \quad a_c = \langle S | D_c^x | S \rangle. \quad (12)$$

As for matrix elements among degenerate VB Bloch functions ($j, j' = 3,4,5,6$),

$$\begin{aligned} H_{jj'} = & \langle u_j | H_0 | u_{j'} \rangle + \langle u_j | \frac{\hbar}{m_0} \bar{k} \cdot \hat{p} | u_{j'} \rangle - 2\frac{\hbar}{m_0} \sum_{\alpha,\beta} \varepsilon_{\alpha\beta} \langle u_j | k_\alpha \hat{p}_\beta | u_{j'} \rangle \\ & + \sum_{\alpha,\beta} \langle u_j | D_c^{\alpha\beta} \varepsilon_{\alpha\beta} | u_{j'} \rangle + \frac{\hbar^2 k^2}{2m_0} \delta_{jj'} + \frac{\hbar^2}{m_0^2} \sum_r^{(B)} \frac{\langle u_j | \bar{k} \cdot \hat{p} | r \rangle \langle r | \bar{k} \cdot \hat{p} | u_{j'} \rangle}{E_v - E_r}, \\ & + \sum_r^{(B)} \frac{\langle u_j | \frac{\hbar}{m_0} \bar{k} \cdot \hat{p} | r \rangle \langle r | \sum_{\alpha,\beta} D_c^{\alpha\beta} \varepsilon_{\alpha\beta} | u_{j'} \rangle + \langle u_j | \sum_{\alpha,\beta} D_c^{\alpha\beta} \varepsilon_{\alpha\beta} | r \rangle \langle r | \frac{\hbar}{m_0} \bar{k} \cdot \hat{p} | u_{j'} \rangle}{E_v - E_r} \end{aligned}, \quad (13)$$

where higher order terms other than ε , $\varepsilon \cdot k$ and k^2 are considered as third-order perturbations and ignored here. The second and third terms $\langle u_j | \frac{\hbar}{m_0} \bar{k} \cdot \hat{p} | u_{j'} \rangle = 0$ and

$-2\frac{\hbar}{m_0} \sum_{\alpha,\beta} \varepsilon_{\alpha\beta} \langle u_j | k_\alpha \hat{p}_\beta | u_{j'} \rangle = 0$ can be simply dropped off. The last term of summation

vanishes for functions of like-parity $\langle u_j | \frac{\hbar}{m_0} \bar{k} \cdot \hat{p} | r \rangle = 0$ and for functions of unlike-parity

$\langle r | \sum_{\alpha,\beta} D_c^{\alpha\beta} \varepsilon_{\alpha\beta} | u_{j'} \rangle = 0$ ¹¹. The summation in the sixth term includes the second-order

perturbing effects on valence bands from higher or lower bands in class (B) while excluding the conduction band, i.e. $r \neq c$. The equation above can be rewritten as,

$$H_{jj'} = E_v \delta_{jj'} + \langle u_j | \hat{F} | u_{j'} \rangle + \langle u_j | \hat{H}^c | u_{j'} \rangle, \quad (14)$$

where $\hat{H}^c = \sum_{\alpha,\beta} D_c^{\alpha\beta} \varepsilon_{\alpha\beta}$, and the operator \hat{F} is written as $\hat{F} = \hat{H}^{LK} + \hat{H}^c$ with,

$$\hat{H}^{LK} = \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2}{m_0^2} \sum_{\alpha,\beta} k_\alpha k_\beta \sum_{r,c}^{(B)} \frac{\hat{p}_\alpha | r \rangle \langle r | \hat{p}_\beta}{E_v - E_r}, \quad \hat{H}^c = \frac{\hbar^2}{m_0^2} \sum_{\alpha,\beta} k_\alpha k_\beta \frac{\hat{p}_\alpha | c \rangle \langle c | \hat{p}_\beta}{E_c}. \quad (15)$$

E_v is the VB band edge. $E_g = E_c - E_v$ is simply the bandgap. Note that the index r, c in the second term of \hat{H}^{LK} indicates that the summation now includes both bands in class (B) and the conduction band. \hat{H}^c represents that the interaction from the conduction

band is subtracted off to keep \hat{F} unchanged. \hat{H}^{LK} is the Luttinger-Kohn Hamiltonian. When basis functions are chosen to be $|X\rangle$, $|Y\rangle$ and $|Z\rangle$, \hat{H}^{LK} has the matrix elements (for example, see ¹⁵),

$$(\hat{H}^{LK})_{l,xzy} = \begin{pmatrix} Lk_x^2 + M(k_y^2 + k_z^2) & Nk_x k_y & Nk_x k_z \\ Nk_x k_y & Lk_y^2 + M(k_x^2 + k_z^2) & Nk_y k_z \\ Nk_x k_z & Nk_y k_z & Lk_z^2 + M(k_x^2 + k_y^2) \end{pmatrix}, \quad (16)$$

where three constants are,

$$L = \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0} \sum_{r,c}^{(B)} \frac{|\langle X|\hat{p}_x|r\rangle|^2}{E_r - E_c}, \quad M = \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0} \sum_{r,c}^{(B)} \frac{|\langle X|\hat{p}_y|r\rangle|^2}{E_r - E_c}, \quad (17)$$

$$N = \frac{\hbar^2}{m_0^2} \sum_{r,c}^{(B)} \frac{\langle X|\hat{p}_x|r\rangle\langle r|\hat{p}_y|Y\rangle + \langle X|\hat{p}_y|r\rangle\langle r|\hat{p}_x|Y\rangle}{E_r - E_c}, \quad (18)$$

In the same basis vectors \hat{H}^c becomes,

$$(\hat{H}^c)_{l,xzy} = \begin{pmatrix} P^2 & k_x k_y & k_y k_z \\ k_y k_x & k_y^2 & k_y k_z \\ k_x k_x & k_y k_y & k_z^2 \end{pmatrix}, \quad (19)$$

Hence the matrix elements of \hat{F} with respect to $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ are,

$$(\hat{F})_{l,xzy} = \begin{pmatrix} L'k_x^2 + M'(k_y^2 + k_z^2) & N'k_x k_y & N'k_x k_z \\ N'k_y k_x & L'k_y^2 + M'(k_x^2 + k_z^2) & N'k_y k_z \\ N'k_x k_x & N'k_y k_y & L'k_z^2 + M'(k_x^2 + k_y^2) \end{pmatrix}, \quad (20)$$

where,

$$L' = L + P^2/E_x, \quad M' = M, \quad N' = N + P^2/E_x. \quad (21)$$

These parameters L' , M' and N' can be replaced by modified Luttinger constants,

$$L' = -\frac{\hbar^2}{2m_0}(\gamma_1 + 4\gamma_2), \quad M' = -\frac{\hbar^2}{2m_0}(\gamma_1 - 2\gamma_2), \quad N' = -\frac{3\hbar^2}{m_0}\gamma_3. \quad (22)$$

The strained-related term \hat{H}^c with respect to basis functions $|X\rangle$, $|Y\rangle$ and $|Z\rangle$, according to arguments in ¹¹, for most of III-V compounds is

$$(\hat{H}^c)_{l,xzy} = \begin{pmatrix} l\varepsilon_x + m(\varepsilon_y + \varepsilon_z) & n\varepsilon_y & n\varepsilon_z \\ n\varepsilon_x & l\varepsilon_y + m(\varepsilon_x + \varepsilon_z) & n\varepsilon_z \\ n\varepsilon_x & n\varepsilon_y & l\varepsilon_z + m(\varepsilon_x + \varepsilon_y) \end{pmatrix}, \quad (23)$$

The coefficients l , m and n are related to deformation potentials by

$$l = a_v + 2b, \quad m = a_v - b, \quad n = \sqrt{3}d. \quad (24)$$

where a_v is the valence-band hydrostatic deformation potential. b and d are the shear deformation potentials.

As an example, the matrix element H_{33} is,

$$H_{33} = E_v + \frac{1}{2}(F_{xx} - iF_{yx} + iF_{xy} + F_{yy}) + \frac{1}{2}(H_{xx}^e - iH_{yx}^e + iH_{xy}^e + H_{yy}^e), \quad (25)$$

where $F_{xx} = \langle X | \hat{F} | X \rangle$, $H_{xx}^e = \langle X | \hat{H}^e | X \rangle$, etc. are denoted as matrix elements relative to the basis $|X\rangle$, $|Y\rangle$ and $|Z\rangle$. According to Eq.(14), (20), (22), (23) and (24),

$$H_{33} = E_v + \frac{1}{2}[Lk_x^2 + M(k_y^2 + k_z^2) + Lk_y^2 + M(k_x^2 + k_z^2)] + \frac{1}{2}[l\varepsilon_x + m(\varepsilon_{yy} + \varepsilon_{zz}) + l\varepsilon_{yy} + m(\varepsilon_{xx} + \varepsilon_{zz})],$$

$$H_{33} = E_v + \frac{\hbar^2}{2m_0}[-(\gamma_1' + \gamma_2')k^2 + 3\gamma_2'k_z^2] + a_v(\varepsilon_x + \varepsilon_{yy} + \varepsilon_{zz}) + \frac{b}{2}(\varepsilon_x + \varepsilon_{yy} - 2\varepsilon_{zz}). \quad (26)$$

similarly other matrix elements are,

$$H_{66} = H_{33}, \quad H_{34} = i\frac{\sqrt{3}\hbar^2}{m_0}\gamma_3'k_xk_z - id(\varepsilon_{xz} - i\varepsilon_{yz}),$$

$$H_{35} = \frac{\sqrt{3}\hbar^2}{m_0}\left[-\frac{\gamma_2'}{2}k_x^2 + i(\gamma_3' - \gamma_2')k_xk_y\right] + \frac{\sqrt{3}}{2}b(\varepsilon_{xx} - \varepsilon_{yy}) - id\varepsilon_{xy},$$

$$H_{44} = E_v - \frac{\hbar^2}{2m_0}[3\gamma_2'k_z^2 - (\gamma_1' + \gamma_2')k^2] - \frac{\hbar^2\gamma_1'}{m_0}k^2 + a_v(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}) - \frac{b}{2}(\varepsilon_{xx} + \varepsilon_{yy} - 2\varepsilon_{zz}),$$

$$H_{46} = H_{35}, \quad H_{56} = -i\frac{\sqrt{3}\hbar^2}{m_0}\gamma_3'k_xk_z + id(\varepsilon_{xz} - i\varepsilon_{yz}), \quad H_{45} = H_{36} = 0. \quad (27)$$

The rest of matrix elements can be obtained from their conjugates. The special case of biaxial strain is considered here, i.e. InGaAs layers are pseudomorphically grown on (001) InP substrate. All InGaAs layers are strain-free. As for GaSbAs layers, the in-plane strain and the strain along perpendicular direction are

$$\varepsilon = \varepsilon_{xx} = \varepsilon_{yy} = (a_1 - a_2)/a_2, \quad \varepsilon_{zz} = -2c_{12}\varepsilon/c_{11}. \quad (28)$$

respectively, where c_{11} and c_{12} are elastic stiffness constants of GaSbAs. Other strain tensor components, $\varepsilon_{xy} = \varepsilon_{yz} = \varepsilon_{zx} = 0$. The Pikus-Bir Hamiltonian can be simplified,

$$H = \begin{bmatrix} E_v + P_c & 0 & i\frac{1}{\sqrt{2}}(1-2\varepsilon)Pk & \sqrt{\frac{2}{3}}(1-2\varepsilon_z)Pk & i\frac{1}{\sqrt{6}}(1-2\varepsilon)Pk & 0 \\ 0 & E_v + P_c & 0 & -\frac{1}{\sqrt{6}}(1-2\varepsilon)Pk & i\sqrt{\frac{2}{3}}(1-2\varepsilon_z)Pk & -\frac{1}{\sqrt{2}}(1-2\varepsilon)Pk \\ -i\frac{1}{\sqrt{2}}(1-2\varepsilon)Pk & 0 & E_v - P_c - Q_c + K_{hh} & 0 & 0 & 0 \\ \sqrt{\frac{2}{3}}(1-2\varepsilon_z)Pk & -\frac{1}{\sqrt{6}}(1-2\varepsilon)Pk & 0 & E_v - P_c + Q_c + K_{hh} & 0 & 0 \\ -i\frac{1}{\sqrt{6}}(1-2\varepsilon)Pk & -i\sqrt{\frac{2}{3}}(1-2\varepsilon_z)Pk & 0 & 0 & E_v - P_c + Q_c + K_{hh} & 0 \\ 0 & -\frac{1}{\sqrt{2}}(1-2\varepsilon)Pk & 0 & 0 & 0 & E_v - P_c - Q_c + K_{hh} \end{bmatrix} \quad (29)$$

where the quadratic k terms are denoted as,

$$K_{hh} = \frac{\hbar^2}{2m_0}[-(\gamma_1' + \gamma_2')k^2 + 3\gamma_2'k_z^2], \quad K_{hh} = -\frac{\hbar^2}{2m_0}[3\gamma_2'k_z^2 - (\gamma_1' + \gamma_2')k^2] - \frac{\hbar^2\gamma_1'}{m_0}k^2, \quad (30)$$

and the notations ¹² for energy band-edge shifts,

$$P_c = a_c(\varepsilon_x + \varepsilon_y + \varepsilon_z), \quad P_v = -a_v(\varepsilon_x + \varepsilon_y + \varepsilon_z), \quad Q_c = -\frac{b}{2}(\varepsilon_x + \varepsilon_y - 2\varepsilon_z), \quad (31)$$

Here all the conduction band-edge E_c and valence-edges E_v are put into an absolute energy scale. It is clearly shown in the matrix elements of (28) that $|3/2, \pm 3/2\rangle$ heavy-hole (HH) band-edges are shifted by $\delta E_{HH} = -P_c - Q_c$ and $|3/2, \pm 1/2\rangle$ light-hole (LH) band-edges are shifted by $\delta E_{LH} = -P_c + Q_c$, and conduction band-edges are shifted by $\delta E_c = P_c$ from their previous unstrained positions. The corrections to the Kane interband matrix element P on both in-plane and perpendicular directions are trivial.

Under conditions of compressive strain $\varepsilon < 0$, for the GaSbAs layers, the CB-edge is raised as $\delta E_c > 0$, the heavy-hole VB-edge is raised as $\delta E_{HH} > 0$, and the light-hole energy VB-edge is lowered as $\delta E_{LH} < 0$, from their unstrained positions. For unstrained InGaAs layers, the band-edges stay stationary. Hence compressive strain actually reduces the residual gap E_s to $E_s - \delta E_{HH}$ at the hetero-interfaces. Both the CB and VB band offsets are effectively increased, i.e., the CB offset and VB offset are $(\delta E_c - \delta E_{HH})$ and δE_{HH} greater than before strain, respectively. These results show the strain-induced effects provide extra flexibility for I-RTD design similar to quantum-well lasers. Immediately with Eq.(29), Eq.(5) can be written explicitly for each projection f_j $\{j=1, \dots, 6\}$. k_x is written as $k_x = k_y e^{i\phi}$ and $k_y = k_x e^{-i\phi}$ with the angle $\phi = \tan^{-1}(k_y/k_x)$ and the in-plane momentum wave vector k_x . The linear transformations,

$$a_{c_x} = \frac{1}{\sqrt{2}}(ie^{i\phi} f_1 \pm f_2), \quad a_{v_x} = \frac{1}{\sqrt{2}}(e^{i\phi} f_4 \pm f_5), \quad a_{h_x} = \frac{1}{\sqrt{2}}(\pm e^{2i\phi} f_3 + e^{-i\phi} f_6), \quad (32)$$

together with simple algebraic operations, the decoupled Hamiltonian equations can be obtained,

$$(E_c + P_c + U - E)a_{c_x} + \sqrt{\frac{2}{3}}P \left[i(1 - 2\varepsilon_x)k_x \mp \frac{1}{2}(1 - 2\varepsilon)k_y \right] a_{v_x} \mp \frac{1}{\sqrt{2}}(1 - 2\varepsilon)Pk_y a_{h_x} = 0, \quad (33)$$

$$-\sqrt{\frac{2}{3}}P \left[i(1 - 2\varepsilon_x)k_x \pm \frac{1}{2}(1 - 2\varepsilon)k_y \right] a_{c_x} + (E_v - P_c + Q_c + K_{hh} + U - E)a_{v_x} = 0, \quad (34)$$

$$\mp \frac{1}{\sqrt{2}}(1 - 2\varepsilon)Pk_y a_{c_x} + (E_v - P_c - Q_c + K_{hh} + U - E)a_{h_x} = 0, \quad (35)$$

3.2. Equations for Strained & Unstrained Layers

For the strained GaSbAs layers, from Eqs.(33)–(35),

$$(E_{c_2} + P_c + U - E)a_{c_x} + \sqrt{\frac{2}{3}}P \left[i(1 - 2\varepsilon_x)k_x \mp \frac{1}{2}(1 - 2\varepsilon)k_y \right] a_{v_x} \mp \frac{1}{\sqrt{2}}P(1 - 2\varepsilon)k_y a_{h_x} = 0, \quad (36)$$

$$-\sqrt{\frac{2}{3}}P \left[i(1 - 2\varepsilon_x)k_x \pm \frac{1}{2}(1 - 2\varepsilon)k_y \right] a_{c_x} + (E_{v_2} - P_c + Q_c + K_{hh} + U - E)a_{v_x} = 0, \quad (37)$$

$$\mp \frac{1}{\sqrt{2}} P(1-2\varepsilon) k_i a_{c\pm} + (E_{v2} - P_c - Q_c + K_{hh} + U - E) a_{h\pm} = 0, \quad (38)$$

For the unstrained InGaAs layers, by allowing $\varepsilon = \varepsilon_{\pm} = 0$ in Eqs. (36)-(37),

$$(E_{c1} + U - E) a_{c\pm} + \sqrt{\frac{2}{3}} P \left(i k_z \mp \frac{1}{2} k_i \right) a_{i\pm} \mp \frac{1}{\sqrt{2}} P k_i a_{h\pm} = 0, \quad (39)$$

$$-\sqrt{\frac{2}{3}} P \left(i k_z \pm \frac{1}{2} k_i \right) a_{c\pm} + (E_{v1} + K_{hh} + U - E) a_{i\pm} = 0, \quad (40)$$

$$\mp \frac{1}{\sqrt{2}} P k_i a_{c\pm} + (E_{v1} + K_{hh} + U - E) a_{h\pm} = 0, \quad (41)$$

3.3. Conduction Band Transport

At first, the scenario of flat band is considered. The InGaAs conduction band edge E_{c1} is taken as reference energy level, i.e. $E - E_{c1} \rightarrow E$. Since all the band-edges are put in the absolute scale, simple algebraic subtractions on these band-edges are allowed and convenient. Therefore, substituting the following energy relationships (ΔU_C , the CB offset, ΔU_V , the VB offset, E_{gb} , the GaSbAs bandgap, E_{gw} , the InGaAs bandgap),

$$\begin{aligned} E_{c2} - E_{c1} &= \Delta U_C, E_{c2} - E_{v2} = E_{gb}, E_{c1} - E_{v1} = E_{gw}, \\ E_{v2} - E_{v1} &= \Delta U_V, E_{v2} - E_{c1} = -E_{gb} + \Delta U_C. \end{aligned} \quad (42)$$

into Eqs.(36)-(37) yields

$$(\Delta U_C + P_c + U - E) a_{c\pm} + \sqrt{\frac{2}{3}} P \left[i(1-2\varepsilon_{\pm}) k_z \mp \frac{1}{2} (1-2\varepsilon) k_i \right] a_{i\pm} \mp \frac{1}{\sqrt{2}} P(1-2\varepsilon) k_i a_{h\pm} = 0, \quad (43)$$

$$-\sqrt{\frac{2}{3}} P \left[i(1-2\varepsilon_{\pm}) k_z \pm \frac{1}{2} (1-2\varepsilon) k_i \right] a_{c\pm} + (-E_{gb} + \Delta U_C - P_c + Q_c + U - E) a_{i\pm} = 0, \quad (44)$$

$$\mp \frac{1}{\sqrt{2}} P(1-2\varepsilon) k_i a_{c\pm} + (-E_{gb} + \Delta U_C - P_c - Q_c + U - E) a_{h\pm} = 0, \quad (45)$$

into Eqs.(39)-(41) yields,

$$(U - E) a_{c\pm} + \sqrt{\frac{2}{3}} P \left(i k_z \mp \frac{1}{2} k_i \right) a_{i\pm} \mp \frac{1}{\sqrt{2}} P k_i a_{h\pm} = 0, \quad (46)$$

$$-\sqrt{\frac{2}{3}} P \left(i k_z \pm \frac{1}{2} k_i \right) a_{c\pm} + (-E_{gw} + U - E) a_{i\pm} = 0, \quad (47)$$

$$\mp \frac{1}{\sqrt{2}} P k_i a_{c\pm} + (-E_{gw} + U - E) a_{h\pm} = 0, \quad (48)$$

For the conduction band transport, those terms proportional to $k^2 \cdot a_{i\pm}$, $k^2 \cdot a_{h\pm}$ (in K_{hh} and K_{hh} terms) can be ignored. When external bias voltage V_a is applied, the collector contact is set as potential zero and the energy E is now referred to the flat-band InGaAs CB band-edge at the emitter. The transforms $a_{c\pm} \rightarrow a_{c\pm}$, $ia_{i\pm} \rightarrow a_{i\pm}$, $-a_{h\pm} \rightarrow a_{h\pm}$ are substituted into linear equations above. Eqs.(43)-(44) for strained GaSbAs layers become,

$$(\Delta U_c + P_c + U - |E_{cr}|) a_{cz} + \sqrt{\frac{2}{3}} P \left[(1-2\varepsilon_{\pm}) k_z \pm \frac{1}{2} (1-2\varepsilon) ik_l \right] a_{tz} \pm \frac{1}{\sqrt{2}} P (1-2\varepsilon) k_l a_{hz} = 0, \quad (49)$$

$$- \sqrt{\frac{2}{3}} P \left[(1-2\varepsilon_{\pm}) k_z \mp \frac{1}{2} (1-2\varepsilon) ik_l \right] a_{cz} + \left[|E_{cr}| - (-E_{gb} + \Delta U_c - P_c + Q_c + U) \right] a_{tz} = 0, \quad (50)$$

$$\mp \frac{1}{\sqrt{2}} P (1-2\varepsilon) k_l a_{cz} + \left[|E_{cr}| - (-E_{gb} + \Delta U_c - P_c - Q_c + U) \right] a_{hz} = 0. \quad (51)$$

Eqs.(49)-(51) for unstrained InGaAs layers become,

$$(U - |E_{cr}|) a_{cz} + \sqrt{\frac{2}{3}} P \left(k_z \pm \frac{1}{2} ik_l \right) a_{tz} \pm \frac{1}{\sqrt{2}} P k_l a_{hz} = 0, \quad (52)$$

$$- \sqrt{\frac{2}{3}} P \left(k_z \mp \frac{1}{2} ik_l \right) a_{cz} + (|E_{cr}| + E_{gv} - U) a_{tz} = 0, \quad (53)$$

$$\mp \frac{1}{\sqrt{2}} P k_l a_{cz} + (|E_{cr}| + E_{gv} - U) a_{hz} = 0. \quad (54)$$

where $|E_{cr}| = E + eV_o$ is defined. Readily, the heavy-hole projection a_{hz} can be derived in term of a_{cz} from Eq. (51). Upon substituting into Eq.(49)-(50), a_{hz} is eliminated. When the physical variable k_z is replaced by the operator $\hat{k}_z \rightarrow -i\partial/\partial z$, the differential equations for CB and LH projections in GaSbAs layers are,

$$\sqrt{\frac{2}{3}} P \left[(1-2\varepsilon_{\pm}) \frac{\partial}{\partial z} \mp (1-2\varepsilon) \frac{k_l}{2} \right] a_{tz} = -i \left\{ P_c + U(z) + \Delta U_c - |E_{cr}| + \frac{(1-2\varepsilon)^2 P^2 k_l^2}{2[|E_{cr}| - U(z) - \Delta U_c - (-P_c - Q_c) + E_{gb}]} \right\} a_{cz}, \quad (55)$$

$$\sqrt{\frac{2}{3}} P \left[(1-2\varepsilon_{\pm}) \frac{\partial}{\partial z} \pm (1-2\varepsilon) \frac{k_l}{2} \right] a_{cz} = i \left[|E_{cr}| - U(z) - \Delta U_c - (-P_c + Q_c) + E_{gb} \right] a_{tz}. \quad (56)$$

Similarly from Eqs.(52)-(54) the differential equations in InGaAs layers are,

$$\sqrt{\frac{2}{3}} P \left(\frac{\partial}{\partial z} \mp \frac{k_l}{2} \right) a_{tz} = -i \left\{ U(z) - |E_{cr}| + \frac{P^2 k_l^2}{2[|E_{cr}| - U(z) + E_{gv}]} \right\} a_{cz}, \quad (57)$$

$$\sqrt{\frac{2}{3}} P \left(\frac{\partial}{\partial z} \pm \frac{k_l}{2} \right) a_{cz} = i \left[|E_{cr}| - U(z) + E_{gv} \right] a_{tz}, \quad (58)$$

It can be shown that the envelope functions a_{cz} and a_{tz} are continuous at the boundary between strained and unstrained layers. The formalisms of CB current and interband bands current (in next sub-section) are almost the same as in²⁻³.

3.4. Interband Tunneling

The deviations here are similar to previous section. The energy relationships in Eq. (42) are used. After the bias voltage is applied, the energy E is referred to the top of VB at the collector contact. According to Eqs. (36)-(41), for strained GaSbAs layers,

$$\sqrt{\frac{2}{3}} P \left[(1-2\varepsilon_{\pm}) \frac{\partial a_{tz}}{\partial z} \mp \frac{k_l}{2} (1-2\varepsilon) a_{tz} \right] = -(E_{gb} + U + \Delta U_v + P_c - E) a_{cz} \pm \frac{P}{\sqrt{2}} (1-2\varepsilon) k_l a_{hz}, \quad (59)$$

$$\sqrt{\frac{2}{3}}P \left[(1-2\varepsilon_z) \frac{\partial a_{c\pm}}{\partial z} \pm \frac{k_z}{2} (1-2\varepsilon) a_{c\pm} \right] = (U + \Delta U_V - P_c + Q_c - E) a_{iz}, \quad (60)$$

$$\frac{\hbar^2}{2m_h} \frac{d^2}{dz^2} a_{hz} - \left\{ E - [U(z) + \Delta U_V - P_c - Q_c] \right\} a_{hz} \mp \frac{P}{\sqrt{2}} k_z (1-2\varepsilon) a_{c\pm} - \frac{\hbar^2}{2m_0} (\gamma_1' + \gamma_2') k_z^2 a_{hz} = 0, \quad (61)$$

For unstrained InGaAs layers, the differential equations are,

$$\sqrt{\frac{2}{3}}P \left(\frac{\partial a_{iz}}{\partial z} \mp \frac{k_z}{2} a_{iz} \right) = -(E_{gv} + U - E) a_{c\pm} \pm \frac{P}{\sqrt{2}} k_z a_{hz}, \quad (62)$$

$$\sqrt{\frac{2}{3}}P \left(\frac{\partial a_{c\pm}}{\partial z} \pm \frac{k_z}{2} a_{c\pm} \right) = (U - E) a_{iz}, \quad (63)$$

$$\frac{\hbar^2}{2m_h} \frac{d^2}{dz^2} a_{hz} - [E - U(z)] a_{hz} \mp \frac{P}{\sqrt{2}} k_z a_{c\pm} - \frac{\hbar^2}{2m_0} (\gamma_1' + \gamma_2') k_z^2 a_{hz} = 0, \quad (64)$$

where $m_0/m_h = \gamma_1' - 2\gamma_2'$ is the effective mass of heavy hole. For the (heavy-hole) valence electron escaping problem, the quantity of a_{iz} is small when compared to a_{hz} . Hence those terms proportional to $k^2 \cdot a_{iz}$ (in K_n) are actually ignored.

3.5. Optical Discharging

During the optical discharging process, the rate equation for hole charge is $dQ/dt = J_R + J_o$, where J_R is the interband tunneling current and J_o is the optical discharging current. $J_o = -\gamma Q$ is proportional to the sheet hole charge density. According to Fermi's golden rule, the coefficient γ is,

$$\gamma = \frac{8\pi e^2 I}{n_r^2 c \hbar} \frac{k_M}{eV_a - E_\varepsilon} |M_{hf}|^2, \quad (65)$$

where the transition matrix is $M_{hf} = \int dz a_h^* z \psi_f$ for holes excited from the initial confined HH state a_h to the final free HH state ψ_f . The energy of escaped holes is $E_\varepsilon = E_h - \hbar\omega_p$ and $\hbar\omega_p$ is the photon energy. n_r is the index of refraction. I is the optical intensity. For free heavy holes, the Schrödinger equation in GaSbAs layers is,

$$\frac{\hbar^2}{2m_h} \frac{d^2}{dz^2} \psi_f - \{ E_\varepsilon - [U(z) + \Delta U_V - P_c - Q_c] \} \psi_f = 0, \quad (66)$$

in InGaAs layers is,

$$\frac{\hbar^2}{2m_h} \frac{d^2}{dz^2} \psi_f - [E_\varepsilon - U(z)] \psi_f = 0, \quad (67)$$

Hence the momentum is $\hbar k_{hf} = \sqrt{2m_h(eV_a - E_\varepsilon)}$.

4. Simulation Results

The oscillation characteristics are primarily determined by the interband tunneling rate J_R and the hole-charge storage capability in the right GaSbAs barrier region. The

interband current needs to be significant for high frequency oscillation so that a substantial amount of hole-charge is built-up over a very short time. A small residual gap E_r can be obtained with proper choices of compositions x , y while keeping the magnitude of lattice mismatch within about 2.5% to ensure the change from one host semiconductor to the other is instant¹⁰. The actual residual gap is further reduced because of the (compressive) strain. After the residual gap is determined, the hole generation process is set at an appropriate rate for the designate oscillation frequency by setting the bias voltage across the diode properly. Intuitively, for a large residual gap, a large dc bias voltage V_0 across the diode is required in order to induce a significant interband current; conversely, for a small residual gap, a small dc bias voltage is proper. The well width w is a tuning parameter for resonant conduction current since the width determines the relative position of quasi- energy level to the bottom of the conduction band.

The net conduction current for I-RTDs in circuits is written as, $I = A(J_c + J_r + J_d)$, where A is the device area, J_c is the conduction band current density, J_r is the interband tunneling current density and J_d is the displacement current density⁴. The device area is chosen to be $A = 100 \mu\text{m}^2$. The 1.55 micron laser pulses with pulse duration 0.3ps and optical intensity $\sim 3.5 \times 10^7 \text{W/cm}^2$ (without the optical loss) are set for optical discharging. Note the optical intensity has to be high enough to have sufficient amount of (not necessarily complete¹⁰) hole-charge discharged during the pulse duration to drive oscillation cycles, otherwise more and more hole-charge will accumulate in right VB well after first several cycles. With these parameters and device models, one can define explicitly an optical hybrid I-RTD diode simulator. When the diode is embedded into external circuits, a modified harmonic balance method is applied to analyze the device operation. The major power output comes from the fundamental harmonic for this laser-driven relaxation I-RTD oscillator. Hence in simulations, we assume that through the proper oscillator cavity design, the first-harmonic oscillation is matched by external circuits while embedding impedances for higher harmonics are short-circuited⁴. In the I-RTD diode simulator, it is possible to tune structure parameters (compositions x , y , the well width w), operation parameters (the dc voltage across the diode V_0 , the first harmonic amplitude V_{1a} and V_{1b}) such that the first harmonic current component is dominant. The desired matching impedance and first harmonic output power can be estimated from harmonic components of net current after time-domain simulations.

Design Example 1: $\text{In}_{0.48}\text{Ga}_{0.52}\text{As}/\text{GaSb}_{0.77}\text{As}_{0.23}$. The bandgaps, residual gap, strains and strain-induced band-edges shifts can be found in¹⁰. The residual gap is 0.11 eV. The InGaAs bandgap is 0.804 eV at room temperature. According to the analysis in¹⁰, the 1.55 micron optical laser does not introduce harmful electron-hole pair generation. Simulation results for oscillations over frequencies $\sim 300\text{-}600\text{GHz}$ are obtained as listed in Table.1. The output power is almost flat; the matching impedances for the first harmonic vary slightly. These results indicate that the OT-I-RTD can be designed to be a broadband oscillator with large values of output power at THz frequencies.

Table 1. For different frequencies: GaSb_{0.77}As_{0.23}/In_{0.48}Ga_{0.52}As

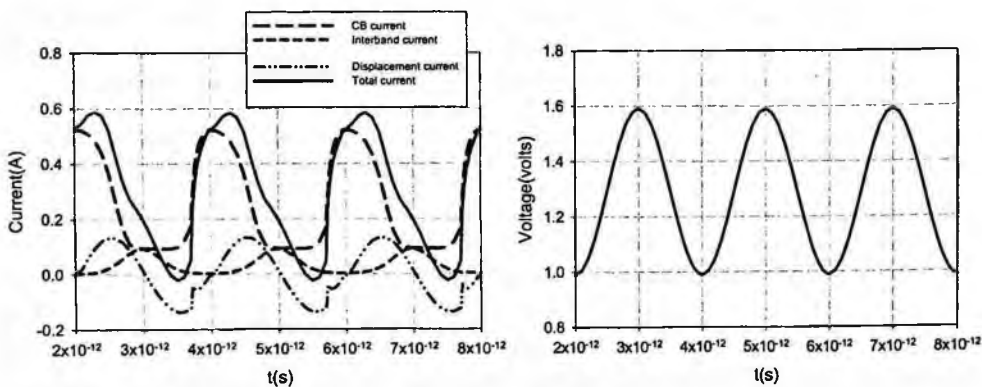
Frequency(GHz)	V ₀ (volt)	V ₁ (volt)	Output power(mW)	Matching impedance(ohm)
500	0.85	0.31	14.5	0.43+j1.12
417	0.83	0.31	14.2	0.47+j1.17
385	0.79	0.31	17.6	0.66+j1.17
357	0.79	0.31	17.5	0.67+j1.18
333	0.79	0.31	17.4	0.68+j1.19

Design Example 2:GaSb_{0.70}As_{0.30}/In_{0.44}Ga_{0.56}As (500 GHz) During device operation, the bandgap of semiconductor alloy actually decreases as the temperature rises due to the heat generation. This leads to concerns about the transparency to 1.55 μm laser. Hence it is prudent to make the bandgap of InGaAs considerably larger than 0.8 eV. This requires increasing the composition x over 0.52. However, the simple analysis shows that larger x leads to the increase of residual gap¹⁰. Hence, a compromise need be made in the design of device structure. In the next design example, the constituents are set to be x=0.56, y=0.70. The corresponding bandgap of InGaAs at room temperature is 0.847eV. It is reduced to 0.8 eV at 400K using the empirical formula in¹⁶. Values for band offsets and residual gaps are given in Table 2.

Table 2. In_{0.44}Ga_{0.56}As/GaSb_{0.70}As_{0.30}(unit of energy: eV)

E_{g1}	E_{g2}	E_s	ΔU_c	ΔU_v	ϵ	ϵ_{∞}	δE_c	δE_{HH}	δE_{LH}
0.847	0.678	0.162	0.515	0.685	-2.2%	2.0%	0.166	0.069	-0.092

The optical discharging condition is the same. The well width is 5.9 nm and the barrier width is kept at 2nm. The bias voltages are V₀=1.29, V_{1a}=-V₁=-0.30V, V_{1b}=0 volts. The output power of the first harmonic is 24.2 mW for diode area of 100 μm². The embedding impedance for the first harmonic is 0.57+j0.86Ω. Fig.2 shows simulation results. The residual gap is 0.162 eV, greater than in Example 1. Hence the dc bias voltage V₀ need be set at a higher value and the well width is assigned a smaller value.

Fig.2. Simulation results of oscillations at 500GHz for In_{0.44}Ga_{0.56}As/GaSb_{0.70}As_{0.30}

5. Conclusions

Practically, the state-of-the-art 1.55 μm pulsed laser possesses PRF $\sim 50\text{--}150$ GHz. Hence, with this hybrid innovation, the OT-I-RTD oscillator design demonstrates a capability for producing $\sim (100/500) \times (15\text{--}25) \text{ mW} = 3\text{--}5 \text{ mW}$ at 500 GHz at room temperature. Larger device area will definitely lead to more power output. The power capability seems to be limited only by thermal heating and laser PRF. It is of importance that the power producing capability will scale directly with advances in the repetition-rate and intensity of 1.55 micron laser technology, which makes this concept increasingly attractive.

References

1. D. Woolard, E. Brown, M. Pepper and M. Kemp, Terahertz frequency sensing and imaging: A Time of Reckoning Future Applications, *Proceedings of the IEEE*, **93**, 1722-1743 (2005).
2. B. Gelmont, D. Woolard, W. Zhang and T. Globus, Electron Transport within Resonant Tunneling Diodes with Staggered-Bandgap Heterostructures, *Solid State Electronics* **46**, 1513-1518 (2002).
3. D. Woolard, W. Zhang, and B. Gelmont, A Novel Interband-Resonant-Tunneling-Diode (I-RTD) Based High-Frequency Oscillator, *Solid State Electronics* **49**, 257-266 (2005).
4. W. Zhang, D. Woolard, E. Brown and B. Gelmont, A Novel I-RTD Based Optically-Pulsed Hybrid Device for Generating THz oscillations, *Pro.SPIE Symp. Optics East*, **5995**, 59950S1-S12 (2005).
5. see *U.S Army THz website*: <http://www.usarmythz.com/THz/R+D/THz-IRTD-Zhang.shtml>.
6. L. L. Chang and L. Esaki, Electronic properties of InAs-GaSb superlattices, *Surface Science* **98**, 70-89 (1980).
7. Chin-An Chang, R. Ludeke, L. L. Chang and L. Esaki, Molecular-beam epitaxy(MBE) of $\text{In}_{1-x}\text{Ga}_x\text{As}$ and $\text{GaSb}_{1-y}\text{As}_y$, *Applied Physics Letter*, **30**, 759-761 (1977).
8. H. Sakaki, L. L. Chang, R. Ludeke, Chin-An Chang, G. A. Sai-Halasz and L. Esaki, $\text{In}_{1-x}\text{Ga}_x\text{As-GaSb}_{1-y}\text{As}_y$ heterojunctions by molecular beam epitaxy, *Appl. Phys. Lett.* **31**, 211-213 (1977).
9. A. Sai-Halasz, R. Tsu. and L. Esaki, A new semiconductor superlattice, *Appl. Phys. Lett.* **130**, 651-653 (1977).
10. D. Woolard, W. Zhang, E. Brown, B. Gelmont and R. Trew, An optically triggered I-RTD hybrid device for continuous-wave generation of THz oscillations, *Pro. of SPIE Terahertz for Military and security Applications IV*, **6212** May 5, 2006.
11. G. L. Bir and G. E. Pikus, "Symmetry & Strain-induced Effects in Semiconductors" (Wiley, NY, 1974).
12. S. L. Chuang, "Physics of Optoelectronic Devices" (John Wiley & Sons, New York, 1995).
13. E. O. Kane, Band Structure of Indium Antimonide, *J.Phys.Chem.Solids*. **1**. 249-261 (1956), also, E. O. Kane, in *Narrow Gap Semiconductors. Physics and Applications, Lec.Notes Phys.* **133** (Springer Verlag, Berlin, 1980).
14. P. Lödwin, A Note on the Quantum-Mechanical Perturbation Theory, *The Journal of Chemical Physics*, **19**, 1396-1401 (1951).
15. J. M. Luttinger and W. Kohn, Motion of electrons and holes in perturbed periodic fields, *Physics Review*, **97**, 869-883 (1955).
16. S. Paul, J. B. Roy and P. K. Basu, Empirical expressions for the alloy composition and temperature dependence of the band gap and intrinsic carrier density in $\text{Ga}_x\text{In}_{1-x}\text{As}$, *J.Appl.Phys.* **69**, 827-829 (1991).

STUDY OF TRANSPORT AND DEVICES BASED ON THE PHOTO-EXCITED TWO-DIMENSIONAL ELECTRONIC SYSTEM

R. G. MANI*

*Gordon McKay Laboratory of Applied Science, Harvard University, 9 Oxford Street,
Cambridge, MA 02138 U.S.A.
mani.rg@gmail.com*

A high mobility two-dimensional electron system exhibits large changes in the resistance, and zero-resistance states, under microwave and Terahertz excitation. We describe associated experimental results and the possibility of using this system as a radiation detector.

Keywords: zero-resistance states, radiation-induced, 2DES, 2DEG, Terahertz, microwaves, devices.

1. Experimental Summary

A two-dimensional electron system (2DES) is realized when electrons are confined to move within a plane, which can be formed, for example, at the planar interface between two epitaxially grown semiconductors such as GaAs and AlGaAs, in the GaAs/AlGaAs heterostructure system. At high magnetic fields and liquid Helium temperatures, such a semiconductor system exhibits quantum Hall effects, where the Hall conductance is quantized in integral and fractional units of the quantum of electrical conductance, e^2/h .^{1,2} Our recent experiments at much lower magnetic fields have indicated that the excitation of high quality GaAs/AlGaAs specimens with microwaves *decreases* the electrical resistance, and further, that the dissipative electrical resistance vanishes, within experimental uncertainty, at a sufficiently large radiation intensity.³ This observation, which has been confirmed by a number of experimental groups,⁴⁻⁸ has produced considerable interest⁹⁻²³ within the semiconductor community because the absorption of microwaves by electrons is normally expected to heat up the semiconductor system, and *increase* the resistance, unlike what was reported by experiment.³⁻⁸ Here, we examine through experiment this novel zero-resistance effect and associated magnetoresistance oscillations in semiconductors, with a view of realizing new THz devices. Thus, we begin here by providing a summary of our experimental observations.

* Future address: Department of Physics and Astronomy, Georgia State University, Atlanta, GA 30303.

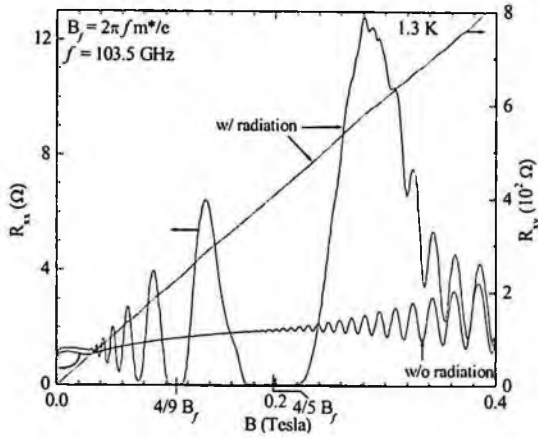


Fig. 1) Magnetoresistance R_{xx} over low B obtained both with (w/) and without (w/o) radiation at 103.5 GHz. Note the giant radiation-induced oscillations in the resistance, which lead into zero-resistance states in the vicinity of $(4/5) B_f$ and $(4/9) B_f$, and the concomitant linear Hall effect.

A photo-excited ultra high mobility 2DES exhibits giant magnetoresistance oscillations, which lead into novel zero-resistance states about magnetic fields $B = (4/5)B_f$ and $B = (4/9)B_f$ of the characteristic field $B_f = 2\pi fm^*/e$, where m^* is the electron mass in the semiconductor system, e is electron charge, and f is the radiation frequency.³ To observe this effect, Hall bar type specimens are fabricated from ultra high mobility GaAs/AlGaAs heterostructures exhibiting an electron density $n \cong 3 \times 10^{11} \text{ cm}^{-2}$ and a mobility $\mu \cong 1.5 \times 10^7 \text{ cm}^2/\text{Vs}$. A sample is then mounted inside a waveguide, immersed in a low temperature cryostat, and irradiated with electromagnetic waves in the microwave and Terahertz (THz) bands, at a power level of $\leq 1 \text{ mW}$, in the presence of a small perpendicular magnetic field.

Fig. 1 illustrates the dissipative (R_{xx}) and Hall (R_{xy}) resistance characteristics of a high mobility device under microwave excitation at 103.5 GHz. Here, the new radiation-induced effect occurs at relatively low magnetic fields, in this instance $B < 0.4$ Tesla, and the effect includes giant magnetoresistance oscillations. Remarkably, it turns out that, at the deepest minima, the resistance falls well below the R_{xx} measured without photoexcitation, as it vanishes over broad B -intervals around $(4/5)B_f$ and $(4/9)B_f$.¹ Here, note that the small magnetic fields involved in this low B -field study can be easily realized with small solenoids or commercially available permanent magnets.

As B_f is linearly related to f , i.e., $B_f = 2\pi fm^*/e$, the characteristic magnetic field, B_f , and the inverse-magnetic-field periodicity, B_f^{-1} , of these radiation - induced magnetoresistance oscillations can be easily changed by varying the radiation frequency, f . Figure 2(b) demonstrates this possibility for $42 \leq f \leq 107 \text{ GHz}$. Here, one observes that characteristic features shift to higher magnetic fields with increasing- f .

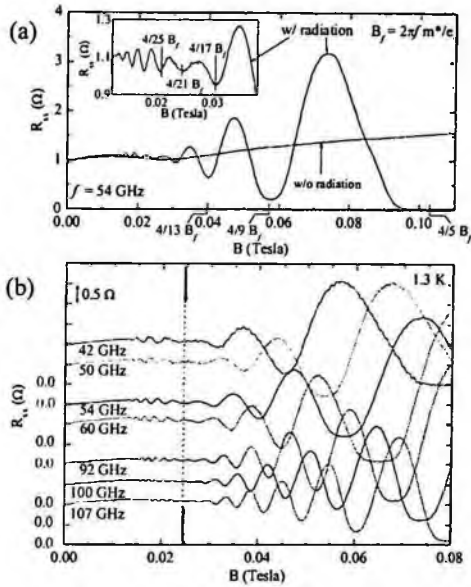


Fig. 2) (a) The magnetoresistance R_{xx} with (w/) and without (w/o) microwave excitation. With radiation, R_{xx} minima occur about $B = [4/(4j+1)] B_f$, and follow this rule through the beat, see inset. (b) R_{xx} oscillations are shown for a set of f .

These radiation-induced resistance oscillations can be empirically described as exponentially damped sinusoidal waves in R_{xx} , i.e., $R_{xx} \sim -A \exp(-\lambda/B) \sin(2\pi B_f/B)$, in the “small oscillation,” low magnetic field limit. Note that this sinusoidal form for the induced resistance oscillations implies that the photo-excited resistance equals the dark resistance at $B_f/B = j$, and resistance minima occur about $B_f/B = j + 1/4$, with $j = 1, 2, 3, \dots$

Although the data of Fig. 1 and Fig. 2 are mainly concerned with radiation frequencies below 107 GHz, experiment indicates that these oscillations and associated zero-resistance states persist to much higher radiation frequencies. Indeed, up to $f = 500$ GHz, which has been the highest accessible frequency, we have observed such magnetoresistance oscillations, without indication of loss of sensitivity. Thus, we present in Fig. 3, the radiation-induced magnetoresistance observed at 360 GHz. Here, if one ignores the “fast” Shubnikov-de Haas magnetoresistance oscillations, which are superimposed on the radiation induced resistance oscillations at $B > 0.2$ Tesla, the basic features are virtually identical to what is observed in Figs. 1 and 2, except that the B-field scale is now expanded, as expected by the scaling relation $B_f = 2\pi f m^*/e$.

A natural question that arises in the context of possible applications is: how much radiation power is necessary to produce such magneto-resistance oscillations? In order to answer this question, we exhibit the dependence of the radiation-induced magnetoresistance oscillations on the absolute radiation intensity in Fig. 4. Here, the quoted radiation intensities are the measured values at the microwave source, and they do not reflect losses that come about in the transmission of the radiation from the source to

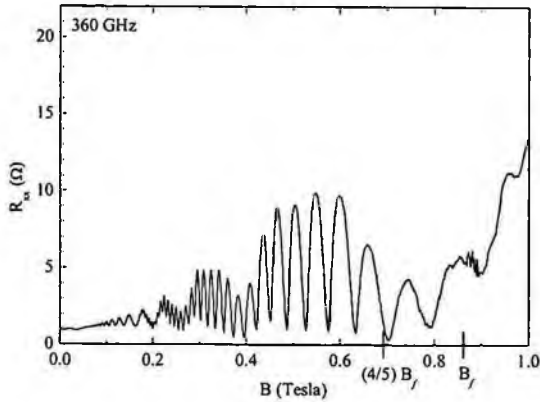


Fig. 3) This figure exhibits radiation induced magneto-resistance oscillations at $f = 360$ GHz. Note that “fast” oscillations, which are due to the Shubnikov-de Haas effect, are superimposed on the “slow” radiation induced component. Cyclotron resonance heating is visible as noise at B_f .

the specimen. Such losses are thought to reduce the microwave power at the sample by approximately 10 dB.

Fig. 4 shows that the source radiation intensity can be reduced from 1000 microwatts [cm^2] to 10 microwatts [cm^2], without a tremendous loss in the oscillatory resistance signal. Indeed, even 1 microwatt [cm^2] yields observable oscillations without supplementary electronic enhancement. We expect that the application of balanced bridge and/or modulation techniques are likely to provide sensitivity to even lower power levels, into the nanowatt [cm^2] level.

The unexpected observation of novel photo-induced oscillations and zero-resistance states³⁻⁸ in a semiconductor system have generated a great deal of theoretical interest, and a number of theoretical studies have attempted to address the origin of this steady-state, effect:⁹⁻²³ Durst and co-workers¹⁰ have identified radiation-induced resistance oscillations with a driven current, similar to Ryzhii.¹⁹ Andreev et al. suggested a physical instability for a possible negative resistivity state, and proposed a scenario for realizing zero-resistance in measurement.¹¹ More recently, Dmitriev et al.²³ have proposed a model, which invokes a change in the distribution function under the steady-state microwave excitation. Although there has been progress in understanding aspects, many basic characteristics could be better understood.

2. Terahertz Sensing

It has sometimes been stated that the new frontier for research in high frequency electronics lies in the Terahertz band, which falls between the microwave and infrared parts of the electromagnetic wave spectrum and includes the frequency interval $0.3 \times 10^{12} < f < 10 \times 10^{12}$ Hz [$300 < f < 10,000$ GHz or $0.3 < f < 10$ THz], corresponding to wavelengths $1000 < \lambda < 30$ micrometers, and energies $1.24 < E < 41.3$ meV. This regime

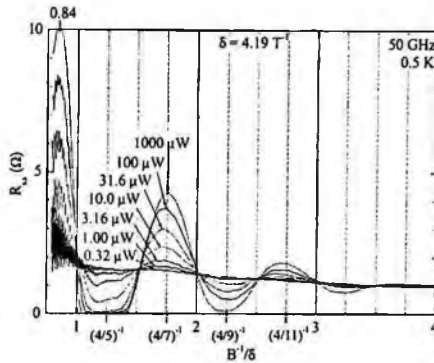


Fig. 4) Radiation induced resistance oscillations at $f = 50$ GHz are exhibited for a number of source intensities in units of μW . Microwave induced resistance oscillations are observable even at $\sim 1 \mu\text{W}$.

comprises a unique set of difficulties because it falls between the upper bound in frequency where electronic approaches continue to operate and the lower limit in f where optical techniques begin to function. Yet, THz detectors and compact radiation sources are needed for interesting applications such as biological agent identification through molecular spectroscopy, large information-bandwidth communications over short distances or at high altitudes, and imaging for security and medical purposes.²⁴⁻²⁷

THz detectors have constituted a challenging problem because the low photon energy in the THz band implies that ambient background thermal noise can easily dominate naturally emitted narrow band signals. Thus, THz detection frequently requires cryogenic cooling or long integration times for improved detection sensitivity. For narrow-band Terahertz sensing, both heterodyne-semiconductor and low temperature heterodyne-superconductor detectors have been employed.^{24,28,29} These detection techniques include the advantage of an increased signal-to-noise ratio due to a reduced detection bandwidth.

Direct detectors are employed for THz detection where spectral resolution is not critical. Associated devices for direct detection include a) Schottky diodes, b) conventional and composite bolometers, c) microbolometers, d) Golay cells, where a gas filled cell changes volume with THz absorption, and the size change is detected by sensing the movement of a mirror, e) acoustic bolometers, which detect changes in pressure due to THz absorption, and f) fast calorimetry, which involves measuring the heating of an absorber filled cavity.²⁴ Non-commercial, cooled, transition-edge bolometers have also played an important role in THz detection. Recently, a single photon quantum dot detector has been demonstrated, where one photon creates an electron-hole pair, and a single electron transistor detects the associated charge polarization.³¹ This sensor exhibits, however, response times in the millisecond range. Semiconductors employed in direct cooled detectors include silicon, germanium, InSb, and HgCdTe. As the THz radiation detection schemes mentioned above often involve complex solutions and cryogenic cooling, it is apparent that even exotic detector solutions could easily gain merit for applications in this band.^{24,25,28,29,31,32}

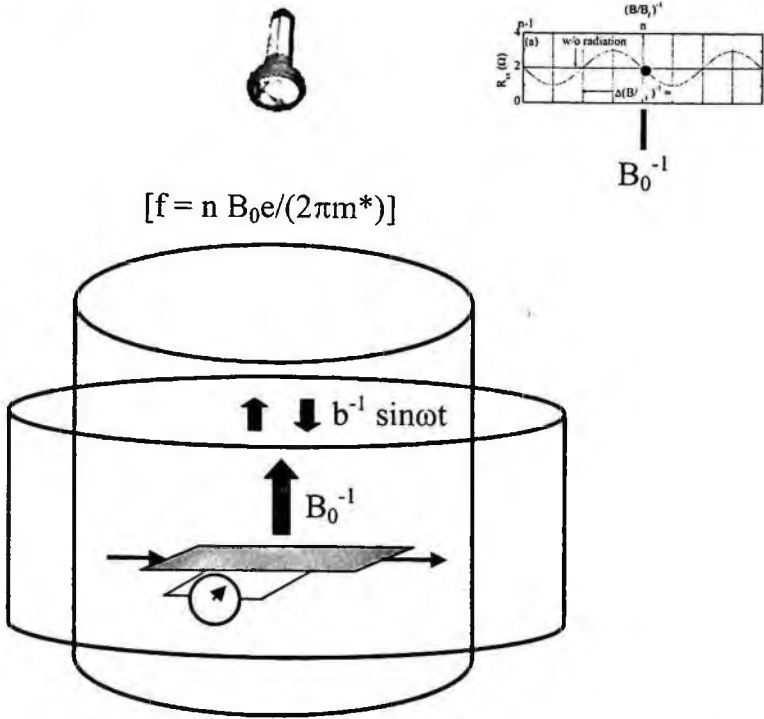


Fig. 5) This figure shows a sketch of the investigated detector concept. An irradiated high mobility two-dimensional electron gas device is subjected to a constant magnetic field B_0 , where B_0^{-1} is chosen to correspond to a fixed point (marked as a dot on the top inset) of the resistance oscillations for incident radiation at a frequency f . The detector device function is realized by superimposing on the static magnetic field, a small time varying component, which has been shown here in blue. Then, a high harmonic, tuned band Terahertz sensor is realized by detecting the device resistance at a odd-harmonic multiple of the field modulation frequency, as the detector is illuminated by Terahertz radiation.

Better *Terahertz radiation sources* are also needed for improved and efficient molecular spectroscopy, THz communications, and imaging.²⁴ For example, in molecular spectroscopy, a narrowband THz source serves as the local oscillator in the heterodyne detection of THz radiation. Yet, narrowband THz sources are difficult to realize because a) electronic solid state sources do not work well at these high frequencies, b) tube sources require high electric and magnetic fields, and c) optical sources like lasers must operate at very low energy levels, comparable to the lattice phonon energy.²⁴ Typically, up to 200 GHz (0.2 THz), electromagnetic radiation is realized by up-converting low-frequency radiation. As high order up converters ($> 4x$) provide poor efficiency, such radiation sources are best realized from chains of low-order multipliers. Yet, such sources exhibit a rapid drop in output power with increasing frequency. Terahertz tube sources, on the other hand, which can provide up to few mW of power at 1 Terahertz, tend to be large in size and expensive in cost, and exhibit lifetimes of 1000's of hours.²⁴ The not-so-compact IR pumped gas laser can provide $\sim 10^3$ mW. A promising approach is the down-conversion from optical sources, through the beating of two slightly detuned lasers on a

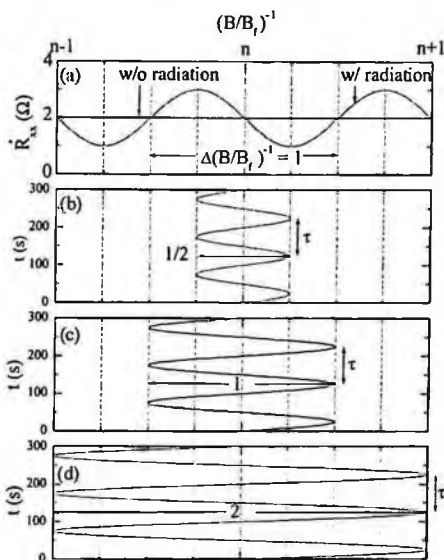


Fig. 6 (a) A sketch of the radiation-induced magnetoresistance R_{xx} vs. the normalized inverse magnetic field. Note that the R_{xx} without radiation crosses the R_{xx} under irradiation at $(B/B_f)^{-1} = n$, as oscillatory minima occur at $n + \pi$ and $n - \pi$. Imagine that $B/B_f = \text{constant} = n^{-1}$, and consider measurements where a small oscillatory variation is superimposed on the static field. Here, the oscillating component is arbitrarily assumed to have a time period, τ , of 100 s. Panels (b)–(d) show the time variation of the total inverse field when the modulation amplitude is 0.5, 1, and 2 times the period of the radiation induced oscillations.

semiconductor point contact, and the coupling of the resulting high frequency (THz) photoinduced current to an antenna.²⁴ Quantum cascade lasers, which operate at low temperatures, offer a good alternative for multi-THz (> 2.5 THz) radiation.³⁰

Due to the existing device limitations, as well as the vast potential for THz applications, it seemed that our new THz physical phenomena ought to be followed and investigated since they might serve to identify new approaches for THz devices.

3. Novel Terahertz (THz) Detector

The experimental results of Fig. 1–4 imply that if radiation at an unknown monochromatic frequency impinges on the high mobility electron specimen under typical experimental conditions, it should be possible to characterize the radiation frequency, f , through measurements of the periodicity, B_f^{-1} , of the resulting magnetoresistance oscillations in a calibrated magnetic field, given the simple empirical relation between B and f , i.e., $B_f = 2\pi f m^*/e$

Indeed, this possibility of obtaining both the radiation frequency (Fig. 2 b) and the intensity (Fig. 4) from a simple electrical resistance measurement suggested the possibility of realizing a simple, tuned, narrowband semiconductor radiation detector, one which operates in the THz regime.

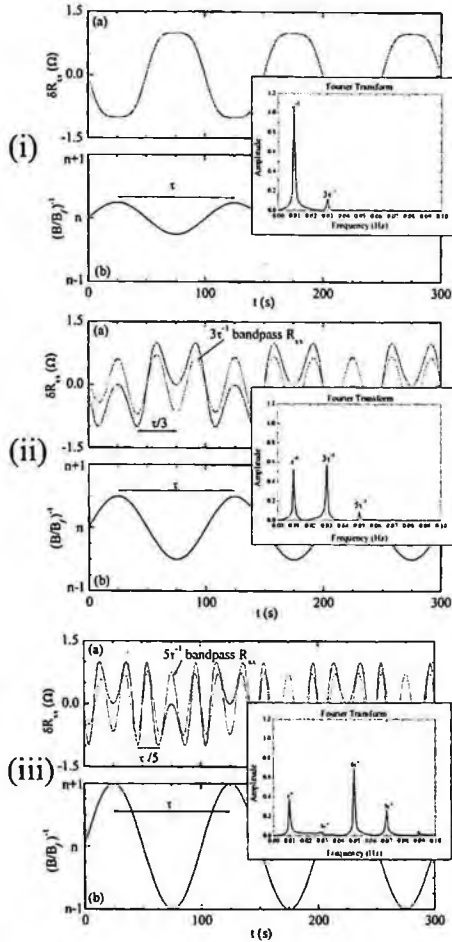


Figure 7: In parts (i) – (iii), the figures (b) exhibit the time variation of the total inverse magnetic field, while the figures (a) illustrate the time varying change in δR_{xx} due to modulation. Here, as the amplitude of the sinusoidal modulation increases, i.e. (i) \rightarrow (iii), there are characteristic changes in δR_{xx} . For example, when the sinusoidal modulation is $\leq \frac{1}{2}$ of the characteristic period (see (i)) of the radiation-induced resistance oscillations, the lineshape of the time varying δR_{xx} (see (i) (a)) looks mostly like that of the modulation (see (i)(b)), within a phase factor. This point is confirmed by the Fourier transform of (i)(a), see inset (i), which shows that the τ^{-1} component, is the dominant component. When the amplitude of the modulation is increased to match the period of the radiation-induced oscillations, see (ii), the lineshape of δR_{xx} begins to differ from that of the driving modulation. Inspection of panel (ii)(a) shows that when the amplitude of modulation matches the period of the radiation-induced oscillatory resistance, the measured signal shows a 3^{rd} harmonic component, see the Fourier transform of (ii) inset, which indicates a peak at $3\tau^{-1}$. As the amplitude of modulation is increased further to twice the period of the radiation-induced resistance oscillations, the 3^{rd} harmonic vanishes, and the 5^{th} harmonic begins to make the dominant contribution, see inset (iii). These characteristic changes in δR_{xx} suggest that a *narrow band* Terahertz sensor can be realized through harmonic detection..

According to the literature, narrow band detection in the THz regime often involves heterodyne techniques or Fourier transform spectroscopy, both of which require exotic supplementary hardware.^{24,33} Our tuned detector involves just a few modifications to the

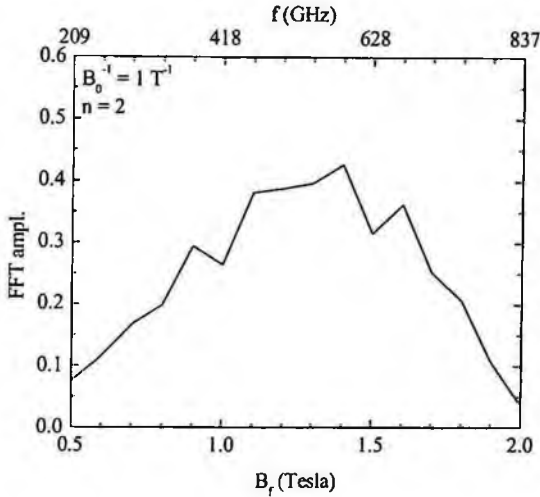


Figure 8: The detector response at the 3rd harmonic as a function of the frequency, when the static magnetic field $B_0^{-1} = 1 \text{ T}^{-1}$, and the working point is the second fixed point, i.e., $n = 2$.

typical experimental configuration. Yet, it provides a frequency selectivity that lies between the extremes of heterodyne- and broadband- detection. Thus, the suggested scheme has the advantage of simplicity.

We have learned that, in order to realize narrow band frequency sensitivity about a radiation frequency f , the magnetic field needs to be first held constant at a value $B_0 = B_f/n$, where n is an integer and $n > 1$. Then, in addition to the static magnetic field, a second sinusoidal oscillating magnetic field component is applied to the specimen, see Fig. 5. As shown below, we find through simulations that when the amplitude of this oscillating field component matches the period of the radiation induced resistance oscillations, there appears a signal at the 3rd harmonic of the modulation frequency, which is proportional to the intensity of the electromagnetic wave excitation in a narrow band about the frequency f . Thus, this 3rd harmonic signal is expected to facilitate tuned narrow band sensing in the THz regime. In this approach, if one wishes to tune the detector sensitivity from the vicinity of one frequency f_0 to another frequency f_1 , one simply changes the static magnetic field from B_{0,f_0} to B_{0,f_1} , and simultaneously adjusts the amplitude of modulation so that it once again corresponds to just one period of the radiation-induced resistance oscillations. Upon making these changes, the 3rd harmonic signal then shows an increased sensitivity to the radiation intensity about f_1 . Such an approach for changing the detection band of interest could turn out to be much simpler than that for existing narrow band heterodyne detection schemes, where a tuned local oscillator is needed for each detection band of interest, when a narrow band IF detector is employed.

The mechanism underlying this approach is summarized in Fig. 6 and 7. Here, in Fig. 6(a), at $(B/B_f)^{-1} = n$, R_{xx} under photoexcitation equals the dark R_{xx} , as in experiment. The static magnetic field is then held fixed at this point, i.e., $B_0 = B_f/n$. In Fig 6(b)-(d), we

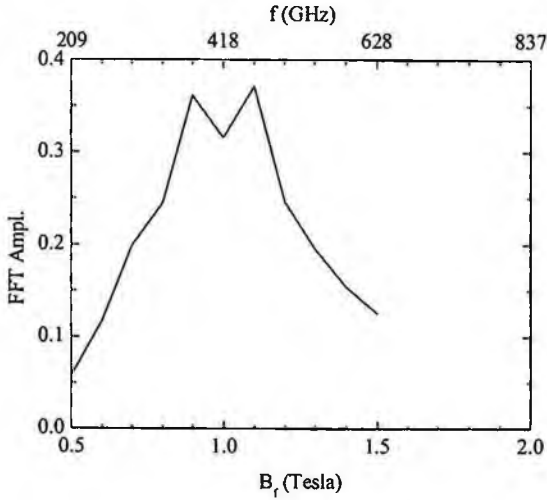


Fig. 9) The detector response at the 5th harmonic as a function of the frequency, when the static magnetic field $B_0^{-1} = 1 \text{ Tesla}^{-1}$, and the working point is the second fixed point.

have conveyed the magnetic field on the specimen, as a function of time, when an oscillating component is superimposed on the static field. We examine next the response of the specimen to the combined effects of the static magnetic field, the oscillatory component, and the photoexcitation in Fig. 7.

When the amplitude of modulation is small, i.e., $\Delta(B/B_0)^{-1} \leq 1/2$, see Fig. 7 (i)(b), the time dependent change in the resistance δR_{xx} under photoexcitation at frequency f , shown in Fig. 7 (i)(a), reflects mostly the time variation of the magnetic field within a phase factor. This situation changes dramatically, however, when the modulation amplitude matches the period of the radiation induced resistance oscillations, see Fig. 6(c), and Fig. 7(ii)(a) and (b). Here, in Fig. 7(ii)(a), the time response of the specimen, i.e., $\delta R_{xx}(t)$, exhibits a strong 3rd harmonic component, which is evident both in the Fourier transform (inset, Fig. 7(ii)) and the 3rd harmonic band-pass filtered portion of $\delta R_{xx}(t)$ (see Fig. 7(ii)(a)). A further increase in the modulation amplitude such that it corresponds to two periods of the radiation induced resistance oscillations (Figs. 6(d) and 7(iii)), leads to the disappearance of the 3rd harmonic component, as a 5th harmonic component takes its place, see inset Fig. 7(iii).

Numerical simulations of this model were carried out to test expectations. The results are shown in Figs. 8 and 9. The figure 8 shows the amplitude of the detected signal at the third harmonic of the modulation frequency when the operating point corresponds to the second node, i.e., $n = 2$, and the B_0 is selected for the detection of radiation in the vicinity of 400 GHz, which lies near the lower edge of the Terahertz band, where the device sensitivity to such radiation has been confirmed by our experiments, see Fig. 3, for example. As confirmed by the simulations, the 3rd harmonic sensing concept yields indeed a narrow band detector, with sensitivity between roughly 200 and 800 GHz, as

indicated by the frequency dependence of the Fourier amplitude. For a center frequency of nearly 400 GHz, the full width at half-maximum of the detector bandwidth appears to be roughly 400 GHz, supporting the point that bandwidth is equal in magnitude to the center frequency, for third harmonic detection. However, as indicated earlier, in this sensor, the bandwidth becomes reduced as one shifts to higher harmonic detection. To make this point, we illustrate also the device sensitivity with 5th harmonic detection in Fig. 9, when the operating point is once again the second fixed point, corresponding to a $B_0 = 1$ T. The results shown in Fig. 9 confirm reduced bandwidth at the higher detection harmonic, in this case the device FWHM corresponds to roughly 200 GHz, about one-half the bandwidth for the 3rd harmonic detector. It appears worth reiterating here that the detector center frequency and the harmonic of detection are set by simply selecting the biasing static field B_0 , and adjusting the modulation field so that it corresponds to the appropriate number cycles of the radiation induced oscillations. Thus, one may change the center frequency of detection by simply changing B_0 . Then, the device response at the 3rd and 5th harmonics will continue to reflect the results shown in Figs. 8 and 9, with the difference that the center frequency would reflect the new biasing static field, and the FWHM of the sensor would reflect the center frequency, similar to the case mentioned above in connection with Figs. 8 and 9.

4. Acknowledgments

This work was funded in part by the L. C. Pegasus Corp., and the Army Research Office.

5. References

1. R. E. Prange and S. M. Girvin, (eds) *The Quantum Hall Effect* (Springer-Verlag, New York, ed.2, 1990).
2. T. Ando, A. B. Fowler, and F. Stern, Electronic Properties of two dimensional systems, *Rev. Mod. Phys.* **54**, 437 (1982).
3. R. G. Mani Zero-resistance states induced by electromagnetic-wave excitation in GaAs/AlGaAs devices, *Nature (London)* **420**, 646-649 (2002).
4. M. A. Zudov, R. R. Du, L. N. Pfeiffer and K. W. West, Evidence for a new dissipationless effect in 2D transport, *Phys. Rev. Lett.* **90**, 046807 (2003).
5. S. I. Dorozhkin, Giant magnetoresistance oscillations caused by cyclotron resonance harmonics, *JETP Lett.* **77**, 577-581 (2003).
6. S. A. Studenikin et al., Microwave radiation induced magneto-oscillations in the longitudinal and transverse resistance of a two-dimensional electron gas, *Sol. St. Comm.* **129**, 341-345 (2004).
7. A. E. Kovalev et al., Observation of a node in the quantum oscillations induced by microwave radiation, *Sol. St. Comm.* **130**, 379-381 (2004).
8. R. L. Willett, K. W. West, and L. N. Pfeiffer, *Bull. Am. Phys. Soc.* **48**, 461 (2003).
9. J. C. Phillips, Microscopic origin of collective exponentially small resistance states, *Sol. St. Comm.* **127**, 233 (2003).
10. A. C. Durst, S. Sachdev, N. Read, and S. M. Girvin, Radiation induced magnetoresistance oscillations in a 2D electron gas, *Phys. Rev. Lett.* **91**, 086803 (2003).
11. A. V. Andreev, I. L. Aleiner, and A. J. Millis, Dynamical symmetry breaking as the origin of the zero-dc-resistance state in an ac-driven system, *Phys. Rev. Lett.* **91**, 056803 (2003).

12. P. W. Anderson and W. F. Brinkman, cond-mat/0302129.
13. J. Shi and X. C. Xie, Radiation-induced “zero-resistance state” and the photon assisted transport, *Phys. Rev. Lett.* **91**, 086801 (2003).
14. A. A. Koulakov and M. E. Raikh, Classical model for the negative dc conductivity of ac-driven two-dimensional electrons near the cyclotron resonance, *Phys. Rev. B* **68**, 115324 (2003).
15. A. F. Volkov and V. V. Pavlovskii, Residual resistance in a two-dimensional electron system, A phenomenological approach, *Phys. Rev. B* **69**, 125305 (2004).
16. S. Mikhailov, Microwave-induced magnetotransport phenomena in two-dimensional electron system: Importance of electrodynamic effects, *Phys. Rev. B* **70**, 165311 (2004).
17. J. Shi and X. C. Xie, cond-mat/0303141.
18. F. S. Bergeret, B. Huckestein, and A. F. Volkov, Current-voltage characteristics and the zero-resistance state in a two-dimensional electron gas, *Phys. Rev. B* **67**, 241303 (2003).
19. V. I. Ryzhii, Microwave photoconductivity in two-dimensional electron systems due to photon assisted interaction of electrons with leaky interface phonons, *Phys. Rev. B* **68**, 193402 (2003); V. Ryzhii et al., Absolute negative conductivity in two-dimensional electron systems associated with acoustic scattering stimulated by microwave radiation, *Phys. Rev. B* **68**, 165406 (2003).
20. X. L. Lei and S. Y. Liu, Radiation-induced magnetoresistance oscillations in a two-dimensional electron gas in faraday geometry, *Phys. Rev. Lett.* **91**, 226805 (2003).
21. P. H. Riviera and P. A. Schulz, Radiation-induced zero-resistance states: Possible dressed electronic structure effects, *Phys. Rev. B* **70**, 075314 (2004).
22. K. Park, Radiation-induced zero-resistance state at low magnetic fields and near half-filling of the lowest Landau level, *Phys. Rev. B* **69**, 201301 (2004); M. G. Vavilov et al., Magnetotransport in a two-dimensional electron gas at large filling factors, *Phys. Rev. B* **69**, 035303 (2004); I. A. Dmitriev et al., Oscillatory ac conductivity and photoconductivity of a two-dimensional electron gas: Quasiclassical transport beyond the Boltzmann equation, *Phys. Rev. B* **70**, 165305 (2004); J. Inarrea and G. Platero, Theoretical approach to microwave-radiation-induced zero-resistance states in 2D electron systems, *Phys. Rev. Lett.* **94**, 016806 (2005).
23. I. A. Dmitriev et al., Theory of the microwave-induced oscillations in the magnetoconductivity of a two-dimensional electron gas, *Phys. Rev. B* **71**, 115316 (2005).
24. P. H. Siegel, Terahertz Technology, *IEEE Trans. On Microwave Theory and Techniques*, **50**, 910 – 928 (2002).
25. T. G. Phillips and J. Keene, Submillimeter Astronomy, *Proc. IEEE*, **80**, 1662 -1678 (1992).
26. B. B. Hu and M. C. Nuss, Imaging with terahertz waves, *Opt. Lett.* **20**, 1716-1718 (1999).
27. R. H. Jacobsen, D. M. Mittleman, and M. C. Nuss, Chemical recognition of gases and gas mixtures with terahertz waves, *Opt. Lett.* **21**, 2011-2013 (1996).
28. P. L. Richards, Bolometers for infrared and millimeter waves, *J. Appl. Phys.* **76**, 1-24 (1994).
29. D. G. Prober, Superconducting terahertz mixer using a transition-edge microbolometer, *Appl. Phys. Lett.* **62**, 2119 – 2121 (1993).
30. J. Faist et al., Quantum Cascade Laser, *Science* **264**, 553-556 (1994).
31. S. Komiyama et al., *Nature (London)* **403 - 407**, 405 (2000).
32. C. Kadow, A. W. Jackson, A. C. Gossard, S. Matsuura, and G. A. Blake, Self-assembled ErAs islands in GaAs for optical-heterodyne THz generation, *Appl. Phys. Lett.* **76**, 3510 (2000).
33. R. J. Bell, *Introductory Fourier Transform Spectroscopy* (New York, Academic, 1972).

TERAHERTZ-BASED DETECTORS USING COLD-ATOM OPTICS

FRANK J. CROWNE*

*Army Research Laboratory, 2800 Powder Mill Road
Adelphi, MD 20783, USA
fcrowne@arl.army.mil*

WILLIAM M. GOLDING

*Army Research Laboratory, 2800 Powder Mill Road
Adelphi, MD 20783, USA
wgolding@arl.army.mil*

CLIFFORD HAZELTON

*Army Research Laboratory, 2800 Powder Mill Road
Adelphi, MD 20783, USA
chazelton@arl.army.mil*

In this paper we explore the design of microwave-based structures that can enhance the interaction of electromagnetic fields with cold-atom ensembles, leading to novel sensing modalities based on the quantum-mechanical behavior of these systems. In particular, we discuss electromagnetically-induced transparency in a single uncondensed cold-atom cloud, and a two-cloud version of a SQUID, where the clouds are BEC's and take the place of the weakly coupled superconductors. These systems are both promising candidates for use in the high-precision detection of chemical contaminants.

Keywords: Cold atoms, millimeter waves, terahertz, EIT, finline, quadrupole trap.

1. Introduction

Rapid advances in the technology of cooling atomic vapors to ultralow temperatures have opened up new avenues for the exploitation of quantum-mechanical phenomena in communications, security, computation, and other areas of engineering. However, applications for these systems in the area of spectral sensing remain largely unexplored. In this paper we will discuss some promising and novel ways that the electromagnetic response of cold-atom systems can be used for spectral sensing at frequencies ranging from rf to far-infrared (terahertz).

Our work at Army Research Laboratory in this area is presently centered on first-principles investigations of cold-atom systems. At normal temperatures, these atomic “clouds” behave like ordinary gases, because the atoms in them are free particles whose energies are primarily kinetic, or “thermal”. At temperatures approaching 10^{-9} degrees Kelvin, however, this kinetic energy is virtually zero, so that their behavior is dominated by potentials that localize or “trap” them. Under these conditions, we would expect ordinary gases to form condensed-matter systems, but that does not happen in cold-atom clouds because the number density of atoms in them is low and interactions between them are weak. What does happen is that the loss of thermal energy causes the atoms to behave like quantum-mechanical objects, despite their large mass. These conditions give rise to a variety of wave phenomena, much like the wave behavior of electrons in a single atom. Because these phenomena are spatially extended over macroscopic distances, they are much easier to observe than their microscopic counterparts. Recent advances in the technology indicate that even large molecules can be made to behave this way.¹

As we explore the interaction of cold-atom systems with microwave and terahertz radiation, we find that they have some unique properties as detectors. A comparison with superconductor-based detectors such as SQUIDs is instructive. Because of the third law of thermodynamics, i.e., a system in a single quantum state has zero entropy, the response of a SQUID is almost free of thermal noise. But an additional property of SQUIDs is that they exhibit the phenomenon of coherence, i.e., wave interference, which leads to entirely new effects, e.g. the AC and DC Josephson effects. Cold atom clouds share this behavior, as we will discuss below.

2. Coherence

In this field, the notion of coherence has at least four distinct meanings:

- Intra-atomic coherence: it is clear that low temperatures can reduce Doppler broadening in an atomic vapor. However, because cooling also suppresses collisions between atoms, the conditions for the Golden Rules to hold are weakened, which greatly lengthens the dephasing times for various intrinsically quantum-mechanical phenomena, such as Rabi oscillations.² This allows the underlying quantum-mechanical nature of atomic systems to come to the fore.
- Interatomic coherence: whereas the phenomena referred to in (1) are associated with the internal degrees of freedom of the atom, its translational motion is still classical at ordinary laboratory temperatures. However, in the nanokelvin regime even the translational degrees of freedom of a trapped atom are quantized, giving such atoms appreciable zero-point energies and extended wave functions within the trap.³

- Electromagnetic coherence: because the dielectric and magnetic response of cold-atom clouds to electromagnetic fields differ profoundly from those of classical systems, novel propagating modes (polaritons) arise in coupled electromagnetic-atomic systems.²
- Many-body coherence: the quantum mechanical statistics of atoms leads either to Fermi statistics in a trapped-atom cloud, or to Bose statistics, the latter giving rise to the celebrated Bose-Einstein condensation discovered a decade or so ago.⁴

In this paper we will concentrate on the detector applications of cases (1) – (3), while briefly discussing our preliminary work involving case (4).

3. Magnetic Trapping

Because the atoms in a vapor are electrically neutral, they cannot be manipulated by electric fields. However, because many of these atomic species have intrinsic angular momentum, they also have magnetic moments. Now, it is well known that a uniform magnetic field acting on an atom with magnetic moment $\vec{\mu}$ cannot produce translational motion; however, a nonuniform field $\vec{H}(\vec{x})$ can give rise to a force

$$\vec{F} = -\nabla V(\vec{x}) \quad (1)$$

where

$$V(\vec{x}) = -\vec{\mu} \cdot \vec{H}(\vec{x}) \quad (2)$$

is the familiar Zeeman potential.⁵ It is this classical force that makes the (semi-) classical Stern-Gerlach experiment possible.

At ultra-low temperatures, where we expect quantum dynamics to govern the translational motion of atoms, this magnetic potential should be included in some sort of Schrodinger equation for the translational motion. However, this inclusion is not trivial, since even the simplest description of such an atom must also include its internal angular-momentum degrees of freedom. For example, the wave function of an atom with angular momentum $\frac{1}{2}$ obeys the following equation:

$$\begin{pmatrix} -\frac{\hbar^2}{2m}\nabla^2 & 0 \\ 0 & -\frac{\hbar^2}{2m}\nabla^2 \end{pmatrix} \begin{pmatrix} \Psi_{\uparrow}(\vec{x}) \\ \Psi_{\downarrow}(\vec{x}) \end{pmatrix} + \mu_B F \begin{pmatrix} H_z(\vec{x}) & H_x(\vec{x}) - iH_y(\vec{x}) \\ H_x(\vec{x}) + iH_y(\vec{x}) & -H_z(\vec{x}) \end{pmatrix} \begin{pmatrix} \Psi_{\uparrow}(\vec{x}) \\ \Psi_{\downarrow}(\vec{x}) \end{pmatrix} = E \begin{pmatrix} \Psi_{\uparrow}(\vec{x}) \\ \Psi_{\downarrow}(\vec{x}) \end{pmatrix} \quad (3)$$

Note that the wave function is now a two-component object. At each point in space, this wave function specifies not only the probability that the atom is present but also the probability that its angular momentum points in a specific direction. The large values of

angular momentum that are routinely encountered in neutral atomic systems require even larger systems of equations than this.

Fortunately, there is a semiclassical simplification to this problem, referred to in the literature as the “adiabatic approximation”.⁶ In this picture, a strong enough magnetic field forces the atomic moment to lie entirely along its lines of force, i.e., there is no quantum-mechanical precession. This causes the atom to have two distinct responses to the field: if its magnetic moment is parallel to the magnetic field (“up-spin”), the atom moves toward a *low* magnetic potential. This is the basis of magnetic trapping: a properly engineered magnetic field can cause the potential $V(\bar{x})$ to have a minimum, and atoms in the “up-spin” state experience a potential well. However, if the magnetic moment of the atom is antiparallel to the magnetic field (“down-spin”), the atom moves towards a *high* magnetic potential, and thus will be repelled by the trap. Hence, when a field is turned on in an atomic cloud, only the “up-spin” fraction of atoms are trapped. Moreover, after the trap is established, any interactions that flip the magnetic moment of an atom cause it to leave the trap.

The simplest type of nonuniform magnetic field that gives rise to a magnetic trap is the field around a current-carrying wire.⁷ Because this type of trap actually confines the atom in only two dimensions, leaving it free to move along the wire, it is classified as an “atomic waveguide,” and is referred to as a “Kepler guide” in the literature. Unfortunately, such traps require very high currents to generate the necessary fields, which heats the wires. In addition, they work “too well”, in the sense that the $1/r$ behavior of the potential causes an atom to fall into the wire and hit its surface, which usually flips its spin and causes it to be lost. To remedy this problem, a constant magnetic field pointing perpendicular to the wire is added to the wire field⁷, which pushes the minimum to one side. This kind of trap is shown in Fig. 1.

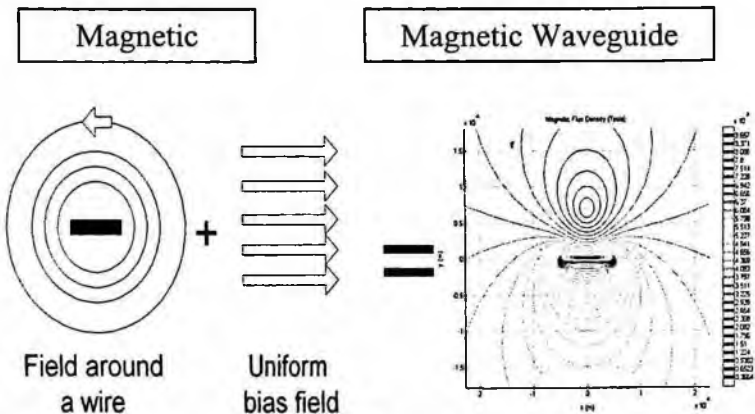


FIG. 1. Simplest type of magnetic waveguide/trap.

An elaboration of the one-wire trap described above is shown in Fig. 2. This trap consists of a square array of four wires, carrying current in the directions shown. Since these currents are all in the z direction (out of the paper), the magnetic field generated by this current configuration is easily calculated.⁸ Near the center of the four-wire square this potential has the form

$$\begin{aligned} B_x &= \partial_y A_z \approx -\frac{\mu I}{4\pi R_0^2} y \\ B_y &= -\partial_x A_z \approx -\frac{\mu I}{4\pi R_0^2} x \end{aligned} \quad (4)$$

where R_0 is the distance from the center of the square to each wire, μ is the permeability of free space, and I is the magnitude of the wire currents. The corresponding potential has the form

$$V \approx \pm \mu_B F \cdot \frac{\mu I}{4\pi R_0^2} \sqrt{x^2 + y^2} \quad (6)$$

where μ_B is the Bohr magneton and F is the angular momentum of the atom; the plus sign is for up-spin atoms, the minus for down-spin atoms. Only the former are trapped.

Our interest in this latter type of trap stems from its remarkable resemblance to a standard electromagnetic waveguide structure used in millimeter-wave technology, called “bilinear finline”.⁹ A bilinear-fineline structure is shown in Fig. 3(a): it consists of four metal strips within a waveguide loaded by a dielectric slab at its center. This slab serves both to confine the wave fields to the center of the guide and to provide mechanical support for the four strips. Figure 3(b) shows a “modified” bilinear finline structure, with air and dielectric regions interchanged. This structure can now accommodate an atomic cloud in the central region, while the strips can carry the dc currents needed to form the trap. The change from wires to strips modifies the potential slightly, but it remains quadrupolar in form at the guide center.

Of course, an unmodified bilinear finline structure has multiple modes associated with it. Fortunately, there is a way to make the cavity favor modes that mimic the field structure of the trap. Fig. 4 shows how to do this by properly connecting (“strapping”) the four strips together. The resulting waveguide cavity will have eigenmodes that produce magnetic fields with the same structure as the dc current – induced fields. This means that the cavity fields will modulate the trap, leading to strong high-frequency interactions with the atomic cloud.

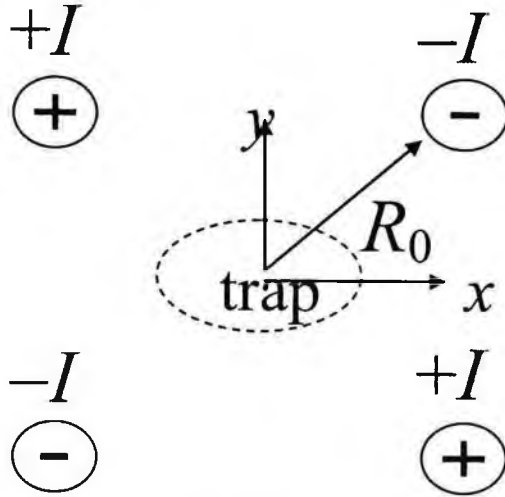
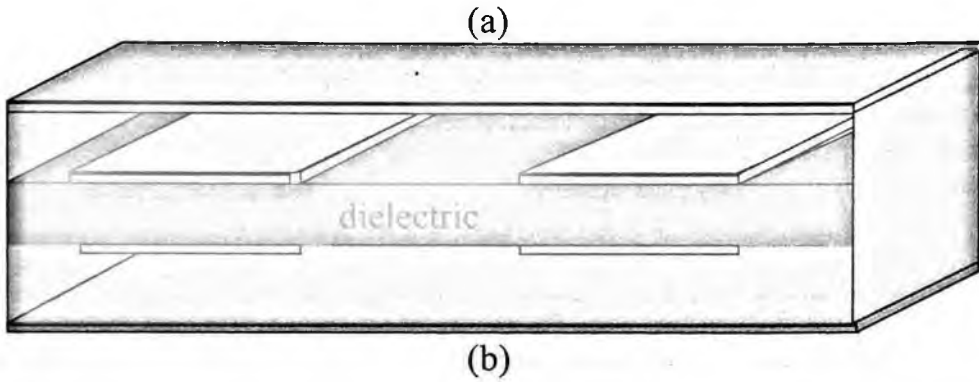


FIG. 2. Four-wire (quadrupolar) magnetic waveguide/trap.



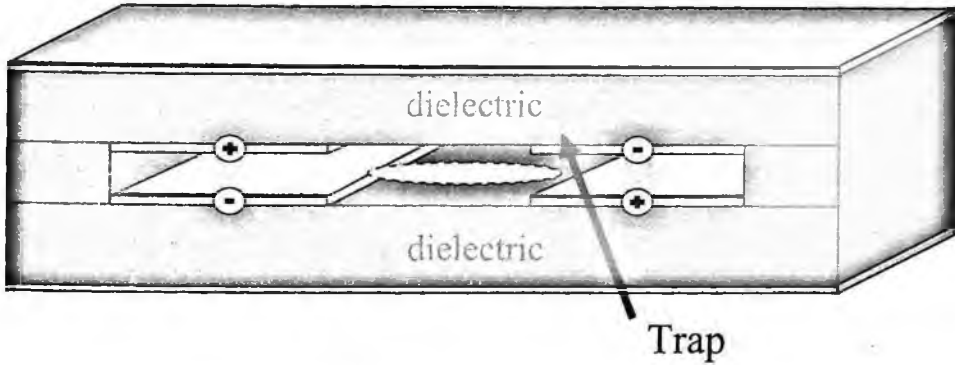


FIG. 3. Comparison of bilinear finline (a) with a four-strip magnetic waveguide/trap (b).

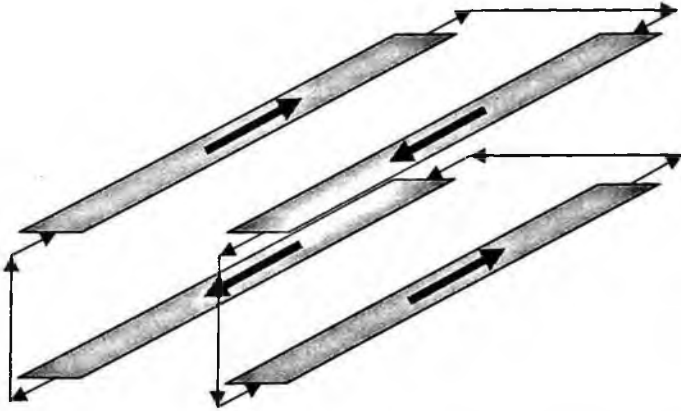


FIG. 4. Strapping pattern for a four-strip cavity designed to produce AC quadrupole trap fields.

4. Detection of Low Gas Concentrations by Electromagnetic Transparency

Because it builds on the well-developed technology of atomic vapor filters, the phenomenon of electromagnetically induced transparency (EIT) is a promising mechanism for detector applications.² This phenomenon makes use of a peculiar type of intra-atomic coherence that arises when a 3-level atom couples to two electromagnetic fields at different frequencies. Fig. 5 shows the energy level diagram of such an atom,

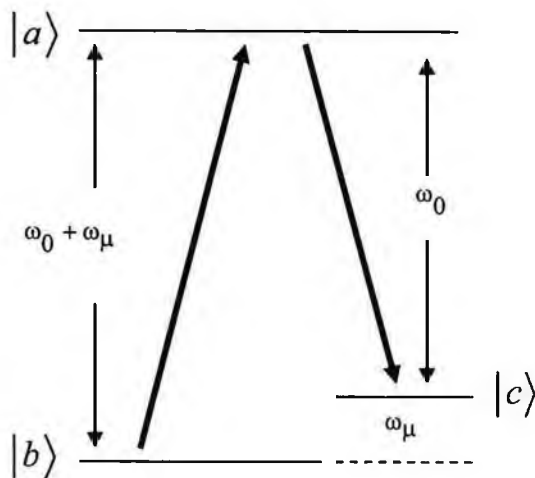


FIG. 5. Atomic 3-level energy diagram for EIT.

whose states are labeled $|a\rangle$, $|b\rangle$, $|c\rangle$. The two transition energies E_{ab} and E_{ac} correspond to optical frequencies, while energy E_{bc} is in the rf to terahertz frequency range. This latter transition, which is between levels of a split ground state, is also characteristic of low lying *molecular* levels (vibrational or rotational levels).¹ It is normally assumed that dipole transitions are not allowed between $|b\rangle$ and $|c\rangle$, a situation that often holds in atomic and molecular systems of this type. When the transition $|a\rangle \leftrightarrow |c\rangle$ is strongly pumped by an optical field, the (complex) dielectric response of a gas of these atoms to a second weak (probe) field \mathcal{E}_{probe} at a frequency ν close to E_{ab} is found to have the form [2]

$$\chi_{ab} = \frac{N_0 \wp_{ab}^2}{\epsilon_0 \hbar} \frac{i(i\Delta + \gamma_{cb})}{(i\Delta + \gamma_{ab})(i\Delta + \gamma_{cb}) + \frac{1}{4}\Omega_{ac}^2} \quad (7)$$

where N_0 is the number density of atoms in the cloud, $\Delta = E_{ab} - \nu$ is the detuning of the probe field from E_{ab} , \wp_{ab} and \wp_{ac} are transition dipole moments for $|a\rangle \leftrightarrow |b\rangle$ and $|a\rangle \leftrightarrow |c\rangle$, $\Omega_{ac} = \frac{\wp_{ac} \mathcal{E}_{pump}}{\hbar}$ is the Rabi frequency corresponding to the pump transition $|a\rangle \leftrightarrow |c\rangle$, \mathcal{E}_{pump} is the pump field amplitude, and γ_{ab} and γ_{cb} are spontaneous decay rates from the excited states to the ground state $|b\rangle$. In general, γ_{cb} will be small for the same reasons that stimulated transitions $|b\rangle \leftrightarrow |c\rangle$ are forbidden; if we assume that $\gamma_{cb} = 0$, the susceptibility becomes

$$\chi_{ab} = \frac{N_0 \rho_{ab}^2}{\epsilon_0 \hbar} \frac{\Delta}{\Delta(\Delta - i\gamma_{ab}) - \frac{1}{4}\Omega_{ac}^2} \quad (8)$$

and it is clear that both the real and imaginary parts of this susceptibility vanish for a probe field in exact resonance with the atomic transition $|a\rangle \leftrightarrow |b\rangle$, i.e., $\Delta = 0$. This phenomenon is called “electromagnetically induced transparency” because the probe field is neither absorbed nor refracted by the atomic medium.

How can we use this effect in a detector? Because the EIT effect is so sensitive, it can be “spoiled” by the presence of uncooled and unpumped atoms in the optical path, which produce absorption losses. A simple photodetector tuned to the probe frequency will then register a drop in current when gas molecules are present in the chamber. We can make this happen by designing the optical path so that it intersects an “interrogation” chamber exposed to the outside air.

For the detector described above to work, the gas of interest must be the same as the working (cold-atom) gas. A technique that allows a variety of gases to be detected using a single working gas can be implemented by simultaneously detuning the pump laser from the EIT condition for the working cold-atom vapor and introducing a microwave field into the main chamber whose frequency equals the detuning. If the cold-atom gas has any nonlinear electromagnetic response, the microwave field will mix with the pump field to generate a sum frequency that satisfies the EIT condition. We are now free to choose the microwave field frequency and pump detuning so as to match an absorption line of the gas of interest. In the presence of this field, the EIT condition will now be satisfied by the sum frequency. If the interrogation chamber is in the microwave path, the presence of outside gases will spoil this condition by modifying the microwave field, either by phase shifting it or attenuating it.

What are the nonlinearities of the cold-atom gas? The standard field-atom equations are highly nonlinear; indeed, their nonlinearity is what gives rise to EIT in the first place. It is noteworthy that the creation of EIT also generates an oscillation in the density matrix component ρ_{cb} at the difference frequency between the pump and the probe fields:

$$\rho_{cb} = \frac{1}{2\hbar} \frac{-i\Omega_{ac}}{(i\Delta + \gamma_1)(i\Delta + \gamma_3) + \frac{1}{4}\Omega_{ac}^2} \mathcal{E}_{probe} e^{-i(\nu - E_{ac})t} \quad (9)$$

Note also that this density component is proportional to the product of \mathcal{E}_{pump} and \mathcal{E}_{probe} . If the level $|c\rangle$ were populated, and if there was a small transition dipole moment \wp_{cb} between the split ground state levels $|b\rangle$ and $|c\rangle$, this would give rise to a microwave susceptibility

$$\chi_{cb} = N_0 \frac{\wp_{cb}}{\hbar} \frac{-i\Omega_{ac}}{(i\Delta + \gamma_1)(i\Delta + \gamma_3) + \frac{1}{4}\Omega_{ac}^2} \mathcal{E}_{probe} \quad (10)$$

It can be shown that the microwave and optical probe fields propagate simultaneously and are spatially coherent, in the sense that their *amplitudes* add – the “polariton” effects alluded to in the Introduction. We are presently exploring the implications of this fact, especially the possibility of phase matching the two signals in the cavity.

5. Detection Modalities based on the Josephson Effect

Because the first models of the physics of superconductivity were based on the phenomenon of Bose-Einstein condensation (BEC) ¹⁰, it is not surprising that the existence of the Josephson effect has also been postulated for cold-atom systems in the BEC state [11]. In both superconductors and BEC cold-atom systems, the Josephson effect arises from the approximation of non-conservation of particle number, which gives rise to a “phase” type of order parameter $\Psi(\bar{x})$ and concomitant wave phenomena, described at $T = 0$ by the Gross-Pitaevski equation ¹¹:

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi + V(\bar{x})\Psi + \lambda \Psi^3 = \mu \Psi \quad (11)$$

Despite its strong resemblance to the Schrodinger equation, this equation encompasses very different physics: the “wave function” $\Psi(\bar{x})$ does not describe a single-particle state in the usual sense, but is rather a building block for the (approximate) ground-state N -particle wave function

$$\Psi_T(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N) = \prod_{n=1}^N \Psi(\bar{x}_n), \quad (12)$$

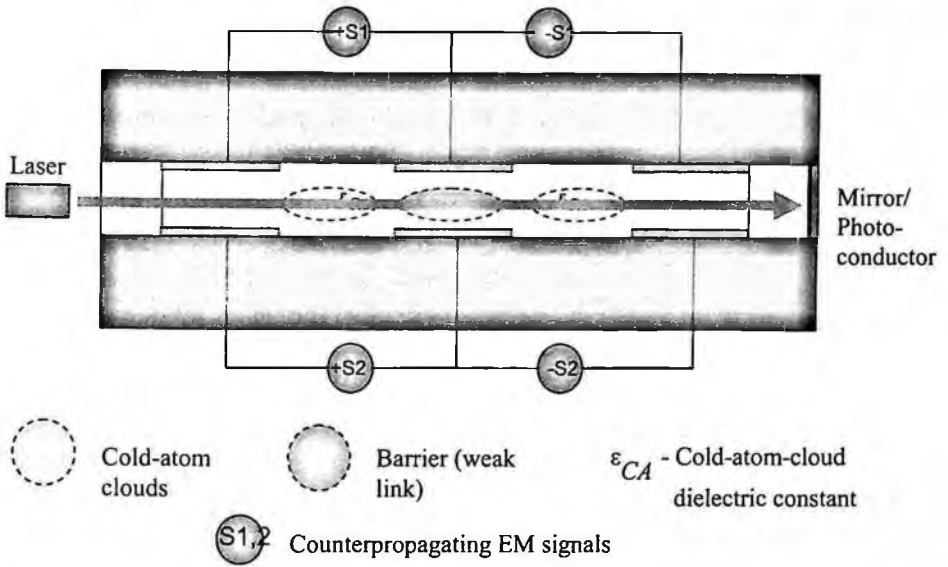


FIG. 6. Generic BEC-microwave interaction structure.

while μ is a chemical potential that contains the Bose statistics, and λ is a scale factor that measures the strength of the attractive “pseudo-force” associated with Bose-Einstein condensation. Field interactions are included by the usual minimal-coupling substitution

$$\nabla \rightarrow \nabla - \frac{e}{c} \vec{A}(\vec{x}) \quad (13)$$

where $\vec{A}(\vec{x})$ is the vector potential of the electromagnetic field.

Of the various superconductivity-inspired phenomena, one of particular interest for detector applications is the AC Josephson effect. In order to create a “weak link” in the language of superconductivity, we need two BEC’s separated by a barrier, and some means to perturb them electromagnetically and observe the result. In Fig. 6 we show a possible geometry for creating and observing this effect: the two atomic clouds are magnetically trapped by striplines, and the low-field region between them forms a barrier. Microwave excitations of this structure apply an AC magnetic field in the proper direction to drive a particle current through the barrier; the resulting oscillatory imbalance of particle density on the two sides can be observed optically through its effect on the dielectric constants of the gases, which are proportional to the atomic density. Again, the sensitivity of this effect to external disturbances, here of the microwave fields driving the Josephson currents, can act as a detection mode.

Our work on this aspect of cold-atom detector phenomenology is in its infancy, but already we have found some interesting differences between the operation of true SQUIDs and their BEC-based counterparts.¹¹ Among them is the important fact that, unlike true superconducting systems, BECs have single-particle excitations that are phononlike, and hence their energy spectrum is gapless. This implies that dissipative processes that are forbidden in a superconductor are allowed in a BEC. We are now trying to evaluate the impact of this new physics on the capabilities of a BEC-based detector.

6. Conclusions

The preliminary nature of our work prevents us from being more specific about the usefulness of cold-atom-based detectors. However, the great variety of phenomena to exploit, coupled with the rapidly evolving technology for creation of cooled quantum-mechanical ensembles, suggests that they hold great promise. In particular, microwave-based techniques, once they have caught up with optical approaches to the detector problem, may hold the key to a new generation of spectral sensing systems.

References

1. D. McGloin and M. H. Dunn, Simple theory of microwave induced transparency in atomic and molecular systems, *J. Modern Optics* **47**, 1887 (2000).
2. M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge University Press, New York, 1997)
3. F. D'Amico, S. Giorgini, L. P. Pitaevskii, and S. Stringari, *Rev. Mod. Phys.* **71**, 463 (1999).
4. See e.g. *Physics Today*, p. 17 (August 1995).
5. E. Merzbacher, *Quantum Mechanics*, 2nd ed., ch. 12 (John Wiley & Sons, New York, 1970).
6. V. I. Yukalov, *Phys. Rev.* **A56**, 5004 (1997)
7. J. Denschlag, D. Cassettari, and J. Schmiedmayer, *Phys. Rev. Lett.* **82**, 2014 (1999).
8. E. A. Hinds and C. Eberlein, *Phys. Rev.* **A61**, 033614-1 (2000)
9. P. J. Meier, Two new integrated-circuit media with special advantages at millimeter wavelengths, in IEEE MTT-S Intl. Microwave Symp. Digest, pp. 221-223 (1972).
10. A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems*, chs. 2, 6 (McGraw-Hill, New York, 1971)
11. A. J. Leggett, Bose-Einstein condensation in the alkali gases: some fundamental concepts, *Rev. Mod. Phys.* **73**, 307 (2001).

DETECTION OF GASEOUS EFFLUENTS FROM AIRBORNE LWIR HYPERSPECTRAL IMAGERY USING PHYSICS-BASED SIGNATURES

DAVID W. MESSINGER

*Digital Imaging and Remote Sensing Laboratory, Chester F. Carlson Center for Imaging Science, Rochester
Institute of Technology,
54 Lomb Memorial Dr, Rochester, NY 14623, USA
messinger@cis.rit.edu*

CARL SALVAGGIO

*Digital Imaging and Remote Sensing Laboratory, Chester F. Carlson Center for Imaging Science, Rochester
Institute of Technology,
54 Lomb Memorial Dr, Rochester, NY 14623, USA
salvaggio@cis.rit.edu*

NATALIE M. SINISGALLI

*Digital Imaging and Remote Sensing Laboratory, Chester F. Carlson Center for Imaging Science, Rochester
Institute of Technology,
54 Lomb Memorial Dr, Rochester, NY 14623, USA
nms7102@cis.rit.edu*

Detection of gaseous effluent plumes from airborne platforms provides a unique challenge to the remote sensing community. The measured signatures are a complicated combination of phenomenology including effects of the atmosphere, spectral characteristics of the background material under the plume, temperature contrast between the gas and the surface, and the concentration of the gas. All of these quantities vary spatially further complicating the detection problem. In complex scenes simple estimation of a "residual" spectrum may not be possible due to the variability in the scene background. A common detection scheme uses a matched filter formalism to compare laboratory-measured gas absorption spectra with measured pixel radiances. This methodology can not account for the variable signature strengths due to concentration path length and temperature contrast, nor does it take into account measured signatures that are observed in both absorption and emission in the same scene. We have developed a physics-based, forward model to predict in-scene signatures covering a wide range in gas / surface properties. This target space is reduced to a set of basis vectors using a geometrical model of the space. Corresponding background basis vectors are derived to describe the non-plume pixels in the image. A Generalized Likelihood Ratio Test is then used to discriminate between plume and non-plume pixels. Several species can be tested for iteratively. The algorithm is applied to airborne LWIR hyperspectral imagery collected by the Airborne Hyperspectral Imager (AHI) over a chemical facility with some ground truth. When compared to results from a clutter matched filter the physics-based signature approach shows significantly improved performance for the data set considered here.

Keywords: hyperspectral; longwave infrared; gaseous effluent; detection

1. Introduction

Airborne hyperspectral imagery in the longwave infrared (LWIR) provides a useful tool for interrogation of gaseous effluents in the atmosphere. Molecular gases typically exhibit unique spectral absorption features in this spectral regime (8 – 12 μm) allowing for both detection and species identification given a sensor with an appropriate spectral response. However, the physics underlying the at-sensor signature phenomenology is complicated and exploitation of such imagery to detect and characterize gaseous effluents is challenging. Under some circumstances a “residual spectrum” can be computed by measuring or estimating the background spectrum and simply removing this signature from the on-plume measurement. This methodology is not possible for many realistic scenarios such as measurements in a complex industrial facility where there are many materials in the background and the spatial/spectral variability of the background is high. Detection in these scenarios requires a target signature that is well matched to the at-sensor radiance measurement. Such a methodology is presented here.

Detection of plumes in hyperspectral imaging applications has been previously studied. Ifarraguerri (1998)¹ describes a passive standoff chemical agent detection system in which Principle Components Analysis (PCA) and Convex Cone Analysis are used to detect plume pixels in the scene. The Projection Pursuit algorithm has been developed by Ifarraguerri & Chang (1998)² for unsupervised detection of gaseous plumes. They show that in a LWIR hyperspectral imager, an SF₆ plume is readily detectable in “early” projections. Matched filter detection of weak gas plumes was investigated by Funk, et al. (2001)³ on synthetic data. They show that the signal-to-clutter ratio is increased using a modified *k*-means clustering algorithm that in turn improves the matched filter detection. Recently, Foy & Theiler (2004)⁴ considered Independent Component Analysis (ICA) as a method of detecting plumes in passive hyperspectral and active LIDAR imagery. They show that it can be effective for weak plumes in characterizing the background clutter and thus improves the ability of matched filters to detect the plume pixels. All of these methods rely on the statistical differences in the scene and do not account for the physics of the target signatures.

Gittins & Marinelli (1998)⁵ describe the AIRIS system used for standoff detection of chemical agents. The AIRIS sensor is typically deployed as a ground-based sensor staring across a field of regard allowing for an accurate estimate of the background radiance to be made. Thus, a residual spectrum can be estimated and matched to known gas target species laboratory-measured spectra⁶. Cosofret, et al. (2004)⁷ describe how the AIRIS system has been used successfully to detect methane leaks in a horizontal viewing geometry with a standoff of as much as 200 m.

The method presented here is based off previous research into the use of physics-based target signatures in a scheme where detection is performed in the native image radiance space. Healey & Slater (1999)⁸ and Thai & Healy (2002)⁹ first presented the method as a way to overcome deficiencies in atmospheric compensation of visible / near infrared / shortwave infrared (Vis / NIR / SWIR) hyperspectral imagery. In this case, variability in the at-sensor target signature manifestations is modeled through variability in properties

of the atmosphere. The method was shown to provide improved detection results particularly for targets in difficult illumination conditions (*i.e.*, in shadow). Ientilucci & Schott (2005)¹⁰ and Ientilucci (2005)¹¹ extended this method to use a more realistic forward target model, a more tightly constrained set of input parameters to the forward model, and included the use of an “infeasibility” metric as a false alarm mitigation tool.

These methods were limited to the reflective portion of the electromagnetic spectrum, but have been extended to the LWIR by O’Donnell, *et al.* (2004, 2005)^{12,13}. The variability in the target signatures was modeled with variations in the gas concentration and temperature state instead of the atmospheric contributions to the signal. This work used synthetic data over a range of gas temperatures and concentrations and considered both single-species and mixed-species plumes. The research presented here extends this work to application to real LWIR hyperspectral imagery of plumes in a complex, industrial facility.

This paper is presented in the following way. Section 2 describes the phenomenology underlying the at-sensor signature manifestations for pixels affected by the presence of a gas plume and the physics-based model used to predict the target signatures. Section 3 presents the detection algorithm implemented. Section 4 describes the test data used and the detection methodology comparison investigated. Section 5 presents the results from the testing and the paper concludes with a brief summary of the work.

2. Physics-Based Target Model

Typical target detection in visible / near infrared / short-wave infrared (Vis / NIR / SWIR) hyperspectral imagery uses a matched filter formalism¹⁴ where the reflectance spectrum of the target of interest is matched to each pixel in an atmospherically compensated image. This provides good detection results if the image can be accurately transformed into the pixel reflectance space. Another methodology uses physics-based models to predict target signatures in the image radiance space^{8,9,10,11}. Here, uncertainties in the atmospheric and illumination conditions on a per-pixel basis are built into a forward model predicting not one, but several target manifestations in the image radiance space. In this manner, a target sub-space is built up describing the possible target signatures in the image. This space is reduced through either statistical or geometric methods¹⁵ to a set of basis vectors or end-member spectra. The background space of the image is similarly characterized and every pixel is tested as to whether it is more “like” the background space or more “like” the target space. In this manner, a likelihood map of target detections can be created.

This physics-based signature detection methodology has been applied here to the detection of gaseous effluents in longwave infrared (LWIR) imagery^{12,13}. Here, the variability in the at-sensor signature is not based in illumination and atmospheric uncertainties, but rather in the gas concentration path length and temperature contrast with the surface. For any release, these physical parameters will vary spatially within the scene and thus, a matched filter detection scheme with a single target spectrum may not be optimal.

The longwave infrared spectral regime is dominated by thermal emission from both solid surfaces and the atmosphere. Gas plumes can be detected through their unique spectral absorption and emission characteristics. The absorption spectrum of a gas can be measured in the laboratory¹⁶. Such signatures determine the “location” of the spectral signature of gases, but not the magnitude or “direction” (*i.e.*, whether the gas feature is observed in emission or absorption). These are determined by the gas concentration path length and the temperature contrast between the column of gas and the surface beneath the column. For an observed pixel that does not contain the effects of a gas plume, the at-sensor radiance can be approximated as

$$L(\lambda) = \varepsilon_s(\lambda)B(\lambda, T_s)\tau_{atm}(\lambda) + L_u(\lambda) \quad (1)$$

where $L(\lambda)$ is the measured radiance as a function of wavelength λ , $\varepsilon_s(\lambda)$ is the surface spectral emissivity, $B(\lambda, T_s)$ is the Planckian blackbody radiance for the surface at a temperature T_s , $\tau_{atm}(\lambda)$ is the atmospheric transmission, and $L_u(\lambda)$ is the atmospheric upwelling radiance. Atmospheric downwelling radiance that is reflected off the surface is ignored and the surface pixel is assumed to have only a single material at constant temperature.

The at-sensor signature of a pixel containing the effects of a layer of gaseous effluent containing a single species can be approximated in the following way. We write the at-sensor measured radiance as

$$L(\lambda) = \{ \varepsilon_s(\lambda)B(\lambda, T_s)\tau_g(\lambda) + \varepsilon_g(\lambda)B(\lambda, T_g) \} \tau_{atm}(\lambda) + L_u(\lambda) \quad (2)$$

where subscripts g refer to the gas plume quantities. Further simplifications can be made. The gas emissivity can be related to the concentration path length of the gas and the absorption coefficient of the gas species by

$$\varepsilon_g(\lambda) \approx ck(\lambda). \quad (3)$$

Here the units of c are parts-per-million-meter (ppm-m) and the units of $k(\lambda)$ are 1/[ppm-m]. Additionally, for optically thin gas layers, the gas transmission can be approximated as

$$\tau_g(\lambda) = 1 - \varepsilon_g(\lambda). \quad (4)$$

These simplifications lead to a radiance model of a pixel containing a layer of gas that can be written as

$$L(\lambda) = \{ \varepsilon_s(\lambda)B(\lambda, T_s)(1 - ck(\lambda)) + ck(\lambda)B(\lambda, T_g) \} \tau_{atm}(\lambda) + L_u(\lambda). \quad (5)$$

Re-arranging terms in this model demonstrates some signature phenomenology that leads to the physics-based signature detection approach for the gas detection problem. If the at-sensor radiance is written as

$$L(\lambda) = \{ \varepsilon_s(\lambda)B(\lambda, T_s) + ck(\lambda)[B(\lambda, T_g) - \varepsilon_s(\lambda)B(\lambda, T_s)] \} \tau_{atm}(\lambda) + L_u(\lambda) \quad (6)$$

we see that the signature strength of the gas depends on several factors. First, there will only be unique signatures in spectral regions where the gas has absorption / emission features as measured in $k(\lambda)$. Outside these spectral regions, the gas is transparent. Second, the magnitude of the signature is dependent on two factors. The concentration path length of the gas, c , scales the overall signature strength as long as the concentration is low enough that the gas is optically thin. Also, the temperature contrast between the layer of gas and the surface below the layer is of dramatic importance as it not only influences the “strength” of the feature, but also the “shape” of the feature. One can easily see from eq. 6 that for pixels where the gas temperature is greater than the effective temperature of the surface, the gas signature will be in emission. In pixels with a lower gas temperature than the surface effective temperature, the gas signature will be in absorption. This also gives rise to the phenomena that for optically thin plumes where the gas temperature is equal to the effective surface temperature, there is *no signature* due to the gas regardless of the concentration path length of the layer. This signature variability based on concentration path length and temperature contrast with the surface motivates a detection scheme that does not search for a single target signature, but many. Downwind from the release point, as the gas cools and diffuses, the signatures may be dramatically different from those near the stack where the concentration and temperature are the greatest.

The previously mentioned radiometric model of a pixel containing a layer of gas is implemented in the detection scheme described here with one further definition. The gas temperature is described explicitly in terms of a temperature contrast with the surface temperature by writing it as

$$T_g = T_s + \Delta T \quad (7)$$

where ΔT is the difference between the surface and gas temperatures. Of course, accurately estimating the surface temperature without knowledge of the surface material emissivity is a problem in and of itself, so an approximate method is used here. The goal in the physics-based signature prediction model is to generate a large number (several hundred) of target signature manifestations that describe (or bound) the possible ways in which the signature could appear in the scene. These manifestations are then searched for in the image and compared with a characterization of the background to contrast each pixel as more “target like” or more “background like”.

3. Detection Algorithm

The detection scheme implemented here uses the physics-based model shown in equation 6 to generate a large number of target signature manifestations. The atmosphere is not varied (as was the case in the reflective implementation of this approach^{8,9,10,11}). Instead, a single estimate of the atmospheric contributions to the at-sensor signal is used in the forward model. The surface temperature is estimated through the identification of a “background” region of interest assumed to be free of the effects of the gas layer, but representative of the materials in the scene. Over this region, the brightness temperature spectrum, $T_b(\lambda)$ is computed for each pixel by inverting the Planck function for each wavelength in the pixel. The maximum brightness temperature in each pixel $T_{b,max}$ is then

compiled into a list, and the mean value of the set of $T_{b,max}$ is computed over the region. This value is used as an estimate to the surface temperature in equation 6. The background material in the forward model is assumed to be a blackbody.

The variability in the predicted radiance signatures derives from computing the signatures over a range of concentration path lengths and ΔT . Typically, the concentration path length is varied from between 0.1 [ppm-m] and 100 [ppm-m], simulating cases of very high and relative low concentrations. ΔT is generally allowed to vary between -10K and +10K. The range of each parameter is subdivided into a fixed number of intervals (10 were chosen in this study). Signatures are then predicted for every combination of concentration path length and temperature contrast. In this way, signatures are predicted for hot, dense portions of the plume as well as cool, diffuse portions of the plume and a range of possibilities in between.

Once the target space is created, it is reduced to a set of end member spectra using a geometric projection scheme. The Maximum Distance Method (MaxD)¹⁵ is used to determine the extrema spectra that define the convex hull enclosing the data in the space. For this work, the target space was reduced to eight end member spectra. The background is characterized in a similar way. The background region of interest described above is reduced to a set of background end member spectra using the same MaxD procedure as the target space. Here, the background was characterized with 15 end member spectra. Efforts were not made to optimize the background or target space characterizations in terms of the number of end member spectra used to describe the space. Research into this topic is ongoing.

Given the background and target space characterizations, a test is applied to each pixel to determine the likelihood that it contains the effects of the target gas in question. The test used is the Generalized Likelihood Ratio Test, GLRT¹⁵. The GLRT as formulated here is based on the matched subspace detector, MSD, written as

$$MSD(\mathbf{x}) = \frac{\mathbf{x}^T (\mathbf{P}_B^\perp - \mathbf{P}_Z^\perp) \mathbf{x}}{\mathbf{x}^T \mathbf{P}_Z^\perp \mathbf{x}} \quad (8)$$

where \mathbf{x} is the test pixel vector, and \mathbf{P}_Y^\perp represents the projection matrix orthogonal to the subspace Y . The subscript Z denotes the subspace matrix containing the background space basis vectors augmented with the target space basis vectors and the subscript B represents the background basis vectors. The GLRT is then written as

$$GLR(\mathbf{x}) = [MSD(\mathbf{x}) + 1]^{\frac{p}{2}} \quad (9)$$

where p is the number of bands in the pixel spectrum. The use of other detectors within the physics-based signatures scheme is under investigation.

The detection scheme is run iteratively on several gas species of interest for a single scene. Separate target spaces are created for each species using the same input parameters to the signature model and the individual species' laboratory absorption

spectrum. As a result, individual detection planes are generated for each target species considered. Mixed plumes are not considered in the detection scheme (*i.e.*, target signatures containing mixtures of gases are not generated and used in the detection process for this study).

4. Test Data

The data used in this study were collected with the Airborne Hyperspectral Imager (AHI) by the University of Hawaii as part of an experiment for the US Environmental Protection Agency. AHI is a longwave infrared (LWIR) pushbroom imager that collects 256 cross-track pixels in 256 spectral channels ranging from 8-11.5 μm ¹⁷. The data collection also used a color infrared (CIR) line scanner producing images for context. Imagery of an industrial facility with known release points was collected during April 2004. Ground truth of the releases consists of the general locations of the releases and the list of species possibly contained in the plumes. No controlled release experiments were performed. All plumes in the imagery were live releases. As such, exact knowledge of the locations and constituent species is unknown making the declaration of “detections” relatively subjective.

Several flightlines were collected during the flight campaign, including both day and night flights, as well as flights at several altitudes. The data used here were collected during the day and were at an altitude of 5000 ft providing a ground sample distance (GSD) of approximately 2 m. Preprocessing of the data was required to remove known bad bands from the data (*e.g.*, at the beginning and end of the spectral region covered) and image edge effects. The data were not spectrally binned for this application as is sometimes done with AHI data to improve the signal-to-noise ratio. Here, the native spectral resolution of the sensor was used.

For comparison, a simple clutter matched filter was applied to the data as well as the physics-based signatures algorithm. The clutter-matched filter was implemented as

$$CMF(\mathbf{x}) = \frac{\mathbf{d}^T \boldsymbol{\Sigma}^{-1} \mathbf{x}}{\mathbf{d}^T \boldsymbol{\Sigma}^{-1} \mathbf{d}} \quad (10)$$

where \mathbf{d} is the target spectrum of interest, here taken to be the laboratory-measured gas absorption spectrum, \mathbf{x} is the de-measured pixel under test, and $\boldsymbol{\Sigma}$ is the background covariance. Both the mean and covariance of the background were estimated from the same set of pixels used to characterize the background in the physics-based approach described above (*i.e.*, the background region of interest).

5. Results

The physics-based signature detection scheme was applied to the dataset as described above.

Application of the physics-based signatures detection requires an estimate of the atmospheric contributions to the at-sensor signatures (see eq. 6). Here, the atmosphere used was the MODTRAN mid-latitude summer model¹⁸ for the appropriate sensor altitude, day of the year, and time of day. No attempt was made to optimize the

atmosphere for the data. Improvements in the match between the atmosphere used in the target space prediction and that actually contributing to the data should improve the detection results.

Target signatures used were taken from a commercially available database of laboratory-measured absorption spectra¹⁶ and were convolved with the spectral response of the sensor (as were the atmospheric contributions described previously). Target species were determined from the ground truth listing of potential releases provided with the dataset. Species in the target library used were methane, propane, butane, ethane, sulfur dioxide, ethylene, propylene, and benzene.

A region of the facility expected to contain methane and ethane releases was investigated as described above and the results are shown below. The physics-based signatures approach and the clutter-matched filter were applied to this image. Figure 1 shows the CIR image of the area and the corresponding 10.5 μm band from the AHI hyperspectral cube. The region is very cluttered containing several background materials: roads, soils, several structures, cooling towers, and extensive metal piping. The full detection images for both target species and both methods are also shown.

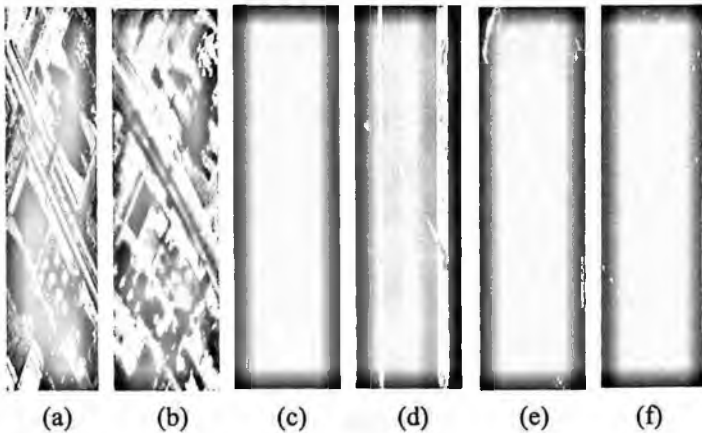


Figure 1: (a) Color IR image of the facility investigated for propane and ethane releases. (b) 10.5 μm band of the corresponding area from the AHI hyperspectral cube. (c) Physics-based signatures detection image for ethane. (d) Physics-based signatures detection for methane. (e) Clutter matched filter detection of ethane. (f) Clutter matched filter detection of methane. The physics-based signature detection images are thresholded to show the top 5% of detection in the image and the clutter matched filter results are thresholded to show the top 10% of detections in the image.

In the full detection images, several features of note are apparent. The clutter matched filter results are highest on the very “bright” objects in the scene. This is a common source of false alarms for this method as it is essentially a projection operator and the test statistic can be large for pixels with a large vector magnitude regardless of spectral similarity to the target spectrum. The physics-based signatures method detects false alarms on two different types of pixels. The diagonal stripe through the center of the detection image corresponds to the surface feature seen in the image. This material may

not have been correctly characterized in the background ROI and as such, the detection scheme has difficulty differentiating it from the target space. The vertical stripes correspond to sensor artifacts in the imagery itself. The physics-based method predicts signatures in the radiometric space of the sensor, but here, sensor artifacts or miscalibrations were not included in the model. Consequently, the manifestations of the targets in these pixels were not correctly predicted and the detection scheme has difficulty distinguishing between target and background. Closer inspection of the physics-based signature detection results show three isolated regions of a high detection statistic outside these regions, described in more detail below (the detections are spatially small making them difficult to distinguish in the full-size images shown in Figure 1). The first is of an ethane release and the remaining two show methane releases. Of particular note is the observation that the physics-based detections are at approximately the same detection level as the false alarms making mitigation relatively simpler. No other species in the target library were detected at appreciable levels.

Figure 2 shows a region highlighting an ethane detection in the physics-based signatures method. The corresponding region in the clutter matched filter detection is also presented along with the higher resolution CIR image of the area providing scene context of the detection.

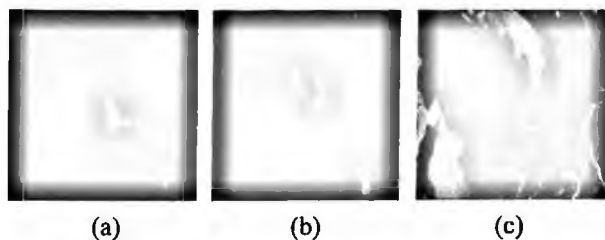


Figure 2: Zoom of the region showing the ethane detection. (a) Physics-based signatures detection. (b) Clutter matched filter detection. (c) CIR image of region of detection for context. The detection images are perfectly registered (they are based on the same input LWIR image) but the CIR image is not and has a different spatial resolution. The detections are approximately in the center of the CIR image.

Here, the detection using the physics-based signatures algorithm produces a small, compact, plume-shaped object with a conical shape indicative of a source and diffusion downwind. The detection strength across the plume in the physics-based signature result is relatively uniform and continuous. This is in contrast to the clutter matched filter results which, while showing strong detections, are discontinuous and do not exhibit the same spatial characteristics.

Figure 3 shows one of the two methane releases detected in this region of the facility. Again, the clutter matched filter results are shown as well as the higher resolution CIR image to provide scene context. Here, the clutter matched filter results (thresholded to only show the top 10% of all detections) do not show any detection of methane. If the matched filter has detected the plume, it is so far below the highest detections in the image that it is not observable at this level. Similar to the ethane detection above, the plume detected in the physics-based signatures detection scheme is small, compact, and

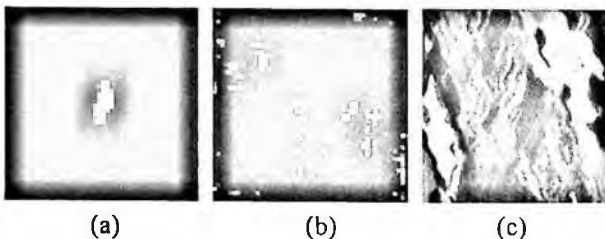


Figure 3: Zoom of the region showing the first methane detection. (a) Physics-based signatures detection. (b) Clutter matched filter detection. (c) CIR image of region of detection for context. The detection images are perfectly registered (they are based on the same input LWIR image) but the CIR image is not and has a different spatial resolution. The detections are approximately in the center of the CIR image.

has spatial characteristics indicative of a small release diffusing due to the effects of the local wind.

Figure 4 shows results for the second methane plume detection with the physics-based signatures detection algorithm. Again, the clutter-matched filter fails to detect any methane at this level while the physics-based signatures algorithm detects a spatially compact, well-defined source.

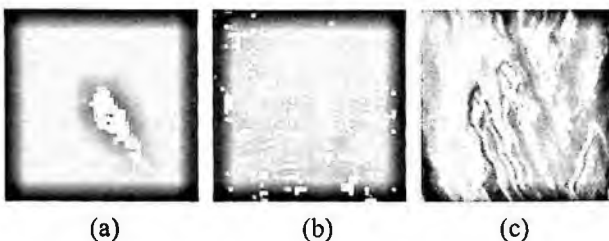


Figure 4: Zoom of the region showing the second methane detection. (a) Physics-based signatures detection. (b) Clutter matched filter detection. (c) CIR image of region of detection for context. The detection images are perfectly registered (they are based on the same input LWIR image) but the CIR image is not and has a different spatial resolution. The detections are approximately in the center of the CIR image.

6. Summary

An algorithm for the detection of gaseous effluents from airborne LWIR hyperspectral imagery of complex scenes has been developed based on the use of physics-based signature predictions. This method accounts for the complex phenomenology underlying the gas signatures manifested at the sensor including variations in the gas concentration path length and temperature contrast with the surface. Detection results for imagery of a facility collected with the AHI airborne LWIR hyperspectral sensor have been presented for both the physics-based signatures approach and the more common clutter matched filter. In the latter case, the target spectrum was the laboratory-measured gas absorption spectrum, a common approach used when a “residual” spectrum is easily estimated. Here, due to the complex nature of the scene such an approach is more difficult and the physics-based approach has been shown to outperform the simple matched filter.

While the physics-based approach provides detections where the clutter matched filter does not, it is also important to note the nature of the false alarms generated by each method. The clutter-matched filter is essentially a projection operator and is susceptible to false alarms on pixels with large spectral magnitude (*i.e.*, “bright” pixels). The physics-based approach tends to false alarm on background materials not well characterized and on sensor artifacts. The latter is because the detection is in the native radiance space of the image and in the presence of miscalibration or artifacts, the predicted signatures will not match well with the measurement.

Improvements to the algorithm are under investigation. Incorporation of real background materials in the target model (as opposed to the assumed blackbody used here) could improve performance. Additionally, target and background space characterization, as well as a detection scheme, that uses a geometric model has not been sufficiently proven to be the optimal method for detection of gas plume in LWIR imagery. Other methods, such as unstructured detection schemes will be tested to characterize performance. Finally, no efforts were made here to optimally estimate the atmospheric contributions to the at-sensor target signature predictions. Use of more accurate estimates of these quantities will better match the target signatures to the measured signatures and should improve performance over that presented here.

Acknowledgements

The authors would like to thank Mr. David Williams of the US Environmental Protection Agency for providing the data used in this study.

References

1. Ifarraguerri, A, Exploitation of imaging spectrometry data for chemical cloud detection, identification and mapping, Proceedings of the 1997 ERDEC Conference on Chemical Defense Research, Special Publication ERDEC-SP-063, 401-406 (1998)
2. Ifarraguerri, A. and Chang, C-I, Projection pursuit analysis of hyperspectral scenes, Conference on Algorithms for Multispectral and Hyperspectral Imagery IV, Proceedings of SPIE, vol. 3372, 1-59 (1998)
3. Funk, C.C., Theiler, J., Roberts, D.A, and Borel, C.C., Clustering to improve matched filter detection of weak gas plumes in hyperspectral thermal imagery, IEEE Transactions on Geoscience and Remote Sensing, 39(7), 410-420 (2001)
4. Foy, B.R. and Theiler, J., Scene analysis and detection in thermal infrared remote sensing using independent component analysis, Independent Component Analyses, Wavelets, Unsupervised Smart Sensors, and Neural Networks II, Proceedings of SPIE vol. 5439, 131-139 (2004)
5. Gittins, C.M. and Marinelli, W., AIRIS multispectral imaging chemical sensor, Electro-Optical Technology for Remote Chemical Detection and Identification III, Proceedings of SPIE, vol. 3383, 65-74 (1998)
6. Jensen, J.O., Thierault, J-M., Bradette, C., Gittins, C. and Marinelli, W., Results from the pronghorn field test using passive infrared spectroradiometers - CATSI and AIRIS, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery VIII, Proceedings of SPIE, vol. 4725, 177-183 (2002)

7. Cosofret, et al. Passive infrared imaging sensor for standoff detection of methane leaks, *Chemical and Biological Standoff Detection II*, Proceedings of SPIE, vol. 5584, 93-99 (2004)
8. Healey, G. and Slater, D., Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions, *IEEE Transactions on Geoscience and Remote Sensing*, **37**(6), 2706-2717 (1999)
9. Thai, B. and Healey, G., Invariant subpixel material detection in hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing*, **40**(3), 599-608 (2002)
10. Ientilucci, E.J. and Schott, J.R., Target detection in a structured background environment using an infeasibility metric in an invariant space, *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, Proceedings of SPIE, vol. 5806, 491-502 (2005)
11. Ientilucci, E.J., Hyperspectral subpixel target detection using hybrid algorithms and physics based modeling, Ph.D. thesis, Rochester Institute of Technology (2005)
12. O'Donnell, E.M., Messinger, D.W., Salvaggio, C., and Schott, J.R., Identification and detection of gaseous effluents from hyperspectral imagery using invariant algorithms, *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, Proceedings of SPIE, vol. 5425, 573-582 (2004)
13. O'Donnell, E.M., Messinger, D.W., Salvaggio, C., and Schott, J.R., The invariant algorithm for identification and detection of multiple gas plumes and weak releases, *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, Proceedings of SPIE, vol. 5806, 206-217 (2005)
14. Manolakis, D., Marden, D., and Shaw, G., Hyperspectral image processing for automatic target detection applications, *Lincoln Laboratory Journal*, **14**(1), 79-116 (2003)
15. Bajorski, P., Ientilucci, E.J., and Schott, J.R., Comparison of basis-vector selection methods for target and background subspaces as applied to subpixel target detection, *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, Proceedings of SPIE, vol. 5425, 97-108 (2004)
16. Sharpe, S., et al., Creation of 0.10 cm⁻¹ resolution, quantitative, infrared spectral libraries for gas samples, *Vibrational Spectroscopy-Based Sensor Systems*, Proceedings of SPIE, vol. 4577, 12-24 (2001)
17. Lucey, P., et al., AHI: An airborne longwave infrared hyperspectral imager, *Airborne Reconnaissance XXII*, Proceedings of SPIE, vol. 3431, 36-43 (1998)
18. Berk, A., Bernstein, L.S., and Robertson, D.C., MODTRAN: A moderate resolution model for LOWTRAN 7, Air Force Geophysics Laboratory, GL-TR-89-0122 (1988)

NONLINEAR CHEMICAL PLUME DETECTION USING KERNEL-BASED MATCHED SUBSPACE DETECTORS

HEESUNG KWON, NASSER M. NASRABADI and PATTI GILLESPIE

U.S. Army Research Laboratory,

2800 Powder Mill Road, Adelphi, MD 20783.

hkwon@arl.army.mil, nnasraba@arl.army.mil, pgillesp@arl.army.mil

In this paper, we compare several detection algorithms that are based on spectral matched (subspace) filters. Nonlinear (*kernel*) versions of these spectral matched (subspace) detectors are also discussed and their performance is compared with the linear versions. These kernel-based detectors exploit the nonlinear correlations between the spectral bands that are ignored by the conventional detectors. Several well-known matched detectors, such as matched subspace detector, orthogonal subspace detector, spectral matched filter and adaptive subspace detector (adaptive cosine estimator) are extended to their corresponding kernel versions by using the idea of kernel-based learning theory. In kernel-based detection algorithms the data is implicitly mapped into a high dimensional kernel feature space by a nonlinear mapping which is associated with a kernel function. The detection algorithm is then derived in the feature space which is *kernelized* in terms of the kernel functions in order to avoid explicit computation in the high dimensional feature space. Experimental results based on simulated toy-examples and real hyperspectral imagery shows that the kernel versions of these detectors outperform the conventional linear detectors.

Keywords: Matched signal detectors, nonlinear subspace matched filters, kernel-based matched subspace detectors

1. Introduction

Detecting signals of interest, particularly with wide signal variability, in noisy environments has long been a challenging issue in various fields of signal processing. Among a number of previously developed detectors, the well-known matched subspace detector (MSD),¹ orthogonal subspace detector (OSD),² spectral matched filter (SMF),^{3,4} and adaptive subspace detectors (ASD) also known as adaptive cosine estimator (ACE),^{5,6} have been widely used to detect a desired signal (target).

Matched signal detectors, such as spectral matched filter and matched subspace detectors (whether adaptive or non-adaptive), only exploit second order correlations, thus completely ignoring nonlinear (higher order) spectral inter-band correlations that could be crucial to discriminate between target and background. In this paper, our aim is to introduce nonlinear versions of MSD, OSD, SMF and ASD detectors which effectively exploit the higher order spectral inter-band correlations in a high (possibly infinite)

dimensional feature space associated with a certain nonlinear mapping via kernel-based learning methods.⁷ A nonlinear mapping of the input data into a high dimensional feature space is often expected to increase the data separability and reduce the complexity of the corresponding data structure. The nonlinear versions of a number of signal processing techniques such as principal component analysis (PCA),⁸ Fisher discriminant analysis,⁹ linear classifiers,¹⁰ kernel matched subspace detectors for target detection^{11,12} and kernel-based anomaly detection¹³ have already been defined in a kernel space.

This paper is organized as follows. Sec. 2 provides the background to the kernel-based learning methods and kernel trick. Sec. 3 introduces a linear matched subspace and its kernel version. The orthogonal subspace detector is defined in Sec. 4 as well as its kernel version. In Sec. 5 we describe the conventional spectral matched filter and its kernel version in the feature space and reformulate the expression in terms of the kernel function using the kernel trick. Finally, in Sec. 6 the adaptive subspace detector and its kernel version are introduced. Performance comparison between the conventional and the kernel versions of these algorithms is provided in Sec. 7 and conclusions are given in Sec. 8.

2. Kernel-based Learning and Kernel Trick

Suppose that the input hyperspectral data is represented by the data space ($\mathcal{X} \subseteq \mathcal{R}^P$) and \mathcal{F} is a feature space associated with \mathcal{X} by a nonlinear mapping function Φ

$$\Phi : \mathcal{X} \rightarrow \mathcal{F}, \quad \mathbf{x} \mapsto \Phi(\mathbf{x}) \quad (1)$$

where \mathbf{x} is an input vector in \mathcal{X} which is mapped into a potentially much higher -- (could be infinite) -- dimensional feature space. Due to the high dimensionality of the feature space, it is computationally not feasible to implement any algorithm directly in feature space. However, kernel-based learning algorithms use an effective kernel trick given by Eq. 2 to implement dot products in feature space by employing kernel functions.⁷ The idea in kernel-based techniques is to obtain a nonlinear version of an algorithm defined in the input space by implicitly redefining it in the feature space and then converting it in terms of dot products. The kernel trick is then used to implicitly compute the dot products in \mathcal{F} without mapping the input vectors into \mathcal{F} ; therefore, in the kernel methods, the mapping Φ does not need to be identified.

The kernel representation for the dot products in \mathcal{F} is expressed as

$$k(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) \quad (2)$$

where k is a kernel function in terms of the original data. There are a large number of Mercer kernels that have the kernel trick property, see⁷ for detailed information about the properties of different kernels and kernel-based learning. Our choice of kernel in this paper is the Gaussian RBF kernel and the associated nonlinear function Φ with this kernel generates a feature space of infinite dimensionality.

3. Linear MSD and Kernel MSD

3.1. Linear MSD

In this model the target pixel vectors are expressed as a linear combination of target spectral signature and background spectral signature, which are represented by subspace target spectra and subspace background spectra, respectively. The hyperspectral target detection problem in a p -dimensional input space is expressed as two competing hypotheses H_0 and H_1

$$\begin{aligned}
 H_0 : \mathbf{y} &= \mathbf{B}\boldsymbol{\zeta} + \mathbf{n}, && \text{Target absent} && (3) \\
 H_1 : \mathbf{y} &= \mathbf{T}\boldsymbol{\theta} + \mathbf{B}\boldsymbol{\zeta} + \mathbf{n} = [\mathbf{T} \ \mathbf{B}] \begin{bmatrix} \boldsymbol{\theta} \\ \boldsymbol{\zeta} \end{bmatrix} + \mathbf{n}, && \text{Target present}
 \end{aligned}$$

where \mathbf{T} and \mathbf{B} represent orthogonal matrices whose p -dimensional column vectors span the target and background subspaces, respectively; $\boldsymbol{\theta}$ and $\boldsymbol{\zeta}$ are unknown vectors whose entries are coefficients that account for the abundances of the corresponding column vectors of \mathbf{T} and \mathbf{B} , respectively; \mathbf{n} represents Gaussian random noise ($\mathbf{n} \in \mathcal{R}^p$) distributed as $\mathcal{N}(0, \sigma^2 \mathbf{I})$; and $[\mathbf{T} \ \mathbf{B}]$ is a concatenated matrix of \mathbf{T} and \mathbf{B} . The numbers of the column vectors of \mathbf{T} and \mathbf{B} , N_t and N_b , respectively, are usually smaller than p ($N_t, N_b < p$).

The generalized likelihood ratio test (GLRT) for the model 3 was derived in,¹ given as

$$L_2(\mathbf{y}) = \frac{\mathbf{y}^T (\mathbf{I} - \mathbf{P}_B) \mathbf{y}}{\mathbf{y}^T (\mathbf{I} - \mathbf{P}_{TB}) \mathbf{y}} \underset{H_0}{\overset{H_1}{>}} \eta. \tag{4}$$

where $\mathbf{P}_B = \mathbf{B}(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T = \mathbf{B} \mathbf{B}^T$ is a projection matrix associated with the N_b dimensional background subspace $\langle \mathbf{B} \rangle$; \mathbf{P}_{TB} is a projection matrix associated with the $(N_{bt} = N_b + N_t)$ -dimensional target-and-background subspace $\langle \mathbf{TB} \rangle$:

$$\mathbf{P}_{TB} = [\mathbf{T} \ \mathbf{B}] [[\mathbf{T} \ \mathbf{B}]^T [\mathbf{T} \ \mathbf{B}]]^{-1} [\mathbf{T} \ \mathbf{B}]^T. \tag{5}$$

3.2. Linear MSD in the Feature Space and its Kernel Version

The hyperspectral detection problem based on the target and background subspaces can be described in the feature space \mathcal{F} as

$$\begin{aligned}
 H_{0\Phi} : \Phi(\mathbf{y}) &= \mathbf{B}_\Phi \boldsymbol{\zeta}_\Phi + \mathbf{n}_\Phi, && \text{Target absent} && (6) \\
 H_{1\Phi} : \Phi(\mathbf{y}) &= \mathbf{T}_\Phi \boldsymbol{\theta}_\Phi + \mathbf{B}_\Phi \boldsymbol{\zeta}_\Phi + \mathbf{n}_\Phi = [\mathbf{T}_\Phi \ \mathbf{B}_\Phi] \begin{bmatrix} \boldsymbol{\theta}_\Phi \\ \boldsymbol{\zeta}_\Phi \end{bmatrix} + \mathbf{n}_\Phi, && \text{Target present}
 \end{aligned}$$

where \mathbf{T}_Φ and \mathbf{B}_Φ represent full-rank matrices whose column vectors span target and background subspaces $\langle \mathbf{B}_\Phi \rangle$ and $\langle \mathbf{T}_\Phi \rangle$ in \mathcal{F} , respectively; $\boldsymbol{\theta}_\Phi$ and $\boldsymbol{\zeta}_\Phi$ are unknown vectors whose entries are coefficients that account for the abundances of the corresponding column vectors of \mathbf{T}_Φ and \mathbf{B}_Φ , respectively; \mathbf{n}_Φ represents Gaussian random noise; and $[\mathbf{T}_\Phi \mathbf{B}_\Phi]$ is a concatenated matrix of \mathbf{T}_Φ and \mathbf{B}_Φ . In general, any sets of basis vectors that span the corresponding subspace can be used as the column vectors of \mathbf{T}_Φ and \mathbf{B}_Φ . Using a similar reasoning as described in the previous subsection, the GLRT of the hyperspectral detection problem depicted by the model in (6) is given by

$$L_{\text{GLRT}}(\Phi(\mathbf{y})) = \frac{\Phi(\mathbf{y})^T (\mathbf{P}_{\mathbf{T}_\Phi} - \mathbf{P}_{\mathbf{B}_\Phi}) \Phi(\mathbf{y})}{\Phi(\mathbf{y})^T (\mathbf{P}_{\mathbf{T}_\Phi \mathbf{B}_\Phi} - \mathbf{P}_{\mathbf{T}_\Phi \mathbf{B}_\Phi}) \Phi(\mathbf{y})}, \quad (7)$$

where $\mathbf{P}_{\mathbf{T}_\Phi}$ represents an identity projection operator in \mathcal{F} ; $\mathbf{P}_{\mathbf{B}_\Phi} = \mathbf{B}_\Phi (\mathbf{B}_\Phi^T \mathbf{B}_\Phi)^{-1} \mathbf{B}_\Phi^T = \mathbf{B}_\Phi \mathbf{B}_\Phi^T$ is a background projection matrix; and $\mathbf{P}_{\mathbf{T}_\Phi \mathbf{B}_\Phi}$ is a joint target-and-background projection matrix in \mathcal{F}

$$\begin{aligned} \mathbf{P}_{\mathbf{T}_\Phi \mathbf{B}_\Phi} &= [\mathbf{T}_\Phi \mathbf{B}_\Phi] \left[[\mathbf{T}_\Phi \mathbf{B}_\Phi]^T [\mathbf{T}_\Phi \mathbf{B}_\Phi] \right]^{-1} [\mathbf{T}_\Phi \mathbf{B}_\Phi]^T \\ &= [\mathbf{T}_\Phi \mathbf{B}_\Phi] \begin{bmatrix} \mathbf{T}_\Phi^T \mathbf{T}_\Phi & \mathbf{T}_\Phi^T \mathbf{B}_\Phi \\ \mathbf{B}_\Phi^T \mathbf{T}_\Phi & \mathbf{B}_\Phi^T \mathbf{B}_\Phi \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{T}_\Phi^T \\ \mathbf{B}_\Phi^T \end{bmatrix} \end{aligned} \quad (8)$$

To kernelize (7) we will separately kernelize the numerator and the denominator. First consider its numerator,

$$\Phi(\mathbf{y})^T (\mathbf{P}_{\mathbf{T}_\Phi} - \mathbf{P}_{\mathbf{B}_\Phi}) \Phi(\mathbf{y}) = \Phi(\mathbf{y})^T \mathbf{P}_{\mathbf{T}_\Phi} \Phi(\mathbf{y}) - \Phi(\mathbf{y})^T \mathbf{B}_\Phi \mathbf{B}_\Phi^T \Phi(\mathbf{y}). \quad (9)$$

Each column of \mathbf{B}_Φ and \mathbf{T}_Φ can be written in terms of its corresponding data space⁷ as

$$\mathbf{B}_\Phi = [\mathbf{e}_b^1 \mathbf{e}_b^2 \dots \mathbf{e}_b^{N_b}] = \Phi_{\mathbf{Z}_B} \mathcal{B}, \quad (10)$$

$$\mathbf{T}_\Phi = [\mathbf{e}_t^1 \mathbf{e}_t^2 \dots \mathbf{e}_t^{N_t}] = \Phi_{\mathbf{Z}_T} \Gamma, \quad (11)$$

where \mathbf{e}_b^i and \mathbf{e}_t^j are the significant eigenvectors of $\mathbf{C}_{\mathbf{B}_\Phi}$ and $\mathbf{C}_{\mathbf{T}_\Phi}$, respectively; $\Phi_{\mathbf{Z}_B} = [\Phi(\mathbf{y}_1) \Phi(\mathbf{y}_2) \dots \Phi(\mathbf{y}_M)]$, $\mathbf{y}_i \in \mathbf{Z}_B$ and $\Phi_{\mathbf{Z}_T} = [\Phi(\mathbf{y}_1) \Phi(\mathbf{y}_2) \dots \Phi(\mathbf{y}_N)]$, $\mathbf{y}_i \in \mathbf{Z}_T$; the column vectors of \mathcal{B} and Γ represent only the significant normalized eigenvectors $(\beta_1, \beta_2, \dots, \beta_{N_b})$ and $(\alpha_1, \alpha_2, \dots, \alpha_{N_t})$ of the background centered kernel matrix $\mathbf{K}(\mathbf{Z}_B, \mathbf{Z}_B) = (\mathbf{K})_{ij} = k(\mathbf{y}_i, \mathbf{y}_j)$, $\mathbf{y}_i, \mathbf{y}_j \in \mathbf{Z}_B$ and the target centered kernel matrix $\mathbf{K}(\mathbf{Z}_T, \mathbf{Z}_T) = (\mathbf{K})_{ij} = k(\mathbf{y}_i, \mathbf{y}_j)$, $\mathbf{y}_i, \mathbf{y}_j \in \mathbf{Z}_T$, respectively. Using (10) the projection of $\Phi(\mathbf{y})$ onto \mathbf{B}_Φ becomes $\mathbf{B}_\Phi^T \Phi(\mathbf{y}) = \mathcal{B}^T \mathbf{K}(\mathbf{Z}_B, \mathbf{y})$ and, similarly, using (11) the projection onto \mathbf{T}_Φ is $\mathbf{T}_\Phi^T \Phi(\mathbf{y}) = \Gamma^T \mathbf{K}(\mathbf{Z}_T, \mathbf{y})$ where $\mathbf{K}(\mathbf{Z}_B, \mathbf{y})$ and $\mathbf{K}(\mathbf{Z}_T, \mathbf{y})$ are referred to as the empirical kernel maps in the machine learning literature,⁷ are column vectors whose entries are $k(\mathbf{x}_i, \mathbf{y})$ for $\mathbf{x}_i \in \mathbf{Z}_B$ and $\mathbf{x}_i \in \mathbf{Z}_T$, respectively. Now we can write

$$\Phi(y)^T \mathbf{B}_\Phi \mathbf{B}_\Phi^T \Phi(y) = \mathbf{K}(\mathbf{Z}_B, y)^T \mathcal{B} \mathcal{B}^T \mathbf{K}(\mathbf{Z}_B, y) \quad (12)$$

The projection onto the identity operator $\Phi(y)^T \mathbf{P}_{\mathbf{I}_\Phi} \Phi(y)$ also needs to be kernelized which is given by

$$\Phi(y)^T \mathbf{P}_{\mathbf{I}_\Phi} \Phi(y) = \Phi(y)^T \Phi_{\mathbf{Z}_{\text{TB}}} \Delta \Delta^T \Phi_{\mathbf{Z}_{\text{TB}}}^T \Phi(y) = \mathbf{K}(\mathbf{Z}_{\text{TB}}, y)^T \Delta \Delta^T \mathbf{K}(\mathbf{Z}_{\text{TB}}, y), \quad (13)$$

where $\Phi_{\mathbf{Z}_{\text{TB}}} = \Phi_{\mathbf{Z}_T} \cup \Phi_{\mathbf{Z}_B}$ and Δ is a matrix whose columns are the eigenvectors $(\kappa_1, \kappa_2, \dots, \kappa_{N_{\Phi}})$ of the centered kernel matrix given by the expression: $\mathbf{K}(\mathbf{Z}_{\text{TB}}, \mathbf{Z}_{\text{TB}}) = (\mathbf{K})_{ij} = k(y_i, y_j)$, $y_i, y_j \in \mathbf{Z}_T \cup \mathbf{Z}_B$ with nonzero eigenvalues, normalized by the square root of their associated eigenvalues, normalized by the square root of their associated eigenvalues and $\mathbf{K}(\mathbf{Z}_{\text{TB}}, y)$ is a concatenated vector $[\mathbf{K}(\mathbf{Z}_T, y)^T \mathbf{K}(\mathbf{Z}_B, y)^T]^T$. To complete the kernelization process the denominator of (7) is given by

$$\begin{aligned} \Phi(y)^T \mathbf{P}_{\mathbf{T}_\Phi \mathbf{B}_\Phi} \Phi(y) &= \Phi(y)^T [\mathbf{T}_\Phi \mathbf{B}_\Phi] \begin{bmatrix} \mathbf{T}_\Phi^T \mathbf{T}_\Phi & \mathbf{T}_\Phi^T \mathbf{B}_\Phi \\ \mathbf{B}_\Phi^T \mathbf{T}_\Phi & \mathbf{B}_\Phi^T \mathbf{B}_\Phi \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{T}_\Phi^T \\ \mathbf{B}_\Phi^T \end{bmatrix} \Phi(y) \\ &= [\mathbf{K}(\mathbf{Z}_T, y)^T \Gamma \mathbf{K}(\mathbf{Z}_B, y)^T \mathcal{B}] \begin{bmatrix} \Gamma^T \mathbf{K}(\mathbf{Z}_T, \mathbf{Z}_T) \Gamma & \Gamma^T \mathbf{K}(\mathbf{Z}_T, \mathbf{Z}_B) \mathcal{B} \\ \mathcal{B}^T \mathbf{K}(\mathbf{Z}_B, \mathbf{Z}_T) \Gamma & \mathcal{B}^T \mathbf{K}(\mathbf{Z}_B, \mathbf{Z}_B) \mathcal{B} \end{bmatrix}^{-1} \begin{bmatrix} \Gamma^T \mathbf{K}(\mathbf{Z}_T, y) \\ \mathcal{B}^T \mathbf{K}(\mathbf{Z}_B, y) \end{bmatrix}. \end{aligned} \quad (14)$$

Finally, substituting (12), (13), and (14) into (7) the kernelized GLRT is given by

$$\mathbf{L}_{2K} = \frac{\mathbf{K}(\mathbf{Z}_{\text{TB}}, y)^T \Delta \Delta^T \mathbf{K}(\mathbf{Z}_{\text{TB}}, y) - \mathbf{K}(\mathbf{Z}_B, y)^T \mathcal{B} \mathcal{B}^T \mathbf{K}(\mathbf{Z}_B, y)}{\mathbf{K}(\mathbf{Z}_{\text{TB}}, y)^T \Delta \Delta^T \mathbf{K}(\mathbf{Z}_{\text{TB}}, y) - [\mathbf{K}(\mathbf{Z}_{\text{TB}}, y)^T \Gamma \mathbf{K}(\mathbf{Z}_{\text{TB}}, y)^T \mathcal{B}] \Lambda_1^{-1} \begin{bmatrix} \Gamma^T \mathbf{K}(\mathbf{Z}_T, y) \\ \mathcal{B}^T \mathbf{K}(\mathbf{Z}_B, y) \end{bmatrix}}, \quad (15)$$

$$\text{where } \Lambda_1 = \begin{bmatrix} \Gamma^T \mathbf{K}(\mathbf{Z}_T, \mathbf{Z}_T) \Gamma & \Gamma^T \mathbf{K}(\mathbf{Z}_T, \mathbf{Z}_B) \mathcal{B} \\ \mathcal{B}^T \mathbf{K}(\mathbf{Z}_B, \mathbf{Z}_T) \Gamma & \mathcal{B}^T \mathbf{K}(\mathbf{Z}_B, \mathbf{Z}_B) \mathcal{B} \end{bmatrix}$$

In the above derivation (15) we assumed that the mapped input data was centered in the feature space by removing the sample mean. However, the original data is usually not centered and the estimated mean in the feature space can not be explicitly computed, therefore, the kernel matrices have to be properly centered. The resulting centered $\hat{\mathbf{K}}$ is shown in⁷ to be given by

$$\hat{\mathbf{K}} = (\mathbf{K} - \mathbf{1}_N \mathbf{K} - \mathbf{K} \mathbf{1}_N + \mathbf{1}_N \mathbf{K} \mathbf{1}_N), \quad (16)$$

where the $N \times N$ matrix $(\mathbf{1}_N)_{ij} = \frac{1}{N}$. The empirical kernel maps $\mathbf{K}(\mathbf{Z}_T, y)^T, \mathbf{K}(\mathbf{Z}_B, y)^T$ and $\mathbf{K}(\mathbf{Z}_{\text{TB}}, y)$ have also to be centered by removing their corresponding empirical kernel map mean. (e.g. $\tilde{\mathbf{K}}(\mathbf{Z}_T, y)^T = \mathbf{K}(\mathbf{Z}_T, y)^T - \frac{1}{N} \sum_{i=1}^N k(y_i, y)$, $y_i \in \mathbf{Z}_T$.)

4. OSP and Kernel OSP Algorithms

4.1. Linear spectral mixture model

The OSP algorithm² is based on maximizing the SNR (signal-to-noise ratio) in the subspace orthogonal to the background subspace and only depends on the noise second-order statistics. It also does not provide any estimate of the abundance measure for the desired end member in the mixed pixel. A linear mixture model for pixel \mathbf{y} consisting of p spectral bands is described by

$$\mathbf{y} = \mathbf{M}\boldsymbol{\alpha} + \mathbf{n}, \quad (17)$$

where the $p \times l$ matrix \mathbf{M} represent l endmembers spectra, $\boldsymbol{\alpha}$ is a $l \times 1$ column vector whose elements are the coefficients that account for the proportions (abundances) of each endmember spectrum contributing to the mixed pixel, and \mathbf{n} is an $p \times 1$ vector representing an additive zero-mean Gaussian noise with covariance matrix $\sigma^2 \mathbf{I}$ and \mathbf{I} is the $p \times p$ identity matrix.

Assuming now we want to identify one particular signature (e.g. a military target) with a given spectral signature \mathbf{d} and a corresponding abundance measure α_p , we can represent \mathbf{M} and $\boldsymbol{\alpha}$ in partition form as $\mathbf{M} = (\mathbf{B} : \mathbf{d})$ and $\boldsymbol{\alpha} = [\alpha_p]$ then model (17) can be rewritten as

$$\mathbf{y} = \mathbf{d}\alpha_p + \mathbf{B}\boldsymbol{\gamma} + \mathbf{n}, \quad (18)$$

where the columns of \mathbf{B} represent the undesired spectral signatures (background signatures or eigenvectors) and the column vector $\boldsymbol{\gamma}$ is the abundance measures for the undesired spectral signatures. The OSP operator that maximizes the signal to noise ratio is given by

$$\mathbf{q}_{osp}^T = \mathbf{d}^T (\mathbf{I} - \mathbf{B}\mathbf{B}^\#) \quad (19)$$

which consists of a background signature rejecter followed by a matched filter. The output of the OSP classifier is now given by

$$D_{osp} = \mathbf{q}_{osp}^T \mathbf{y} = \mathbf{d}^T (\mathbf{I} - \mathbf{B}\mathbf{B}^\#) \mathbf{y} \quad (20)$$

4.2. OSP in feature space and its kernel version

The mixture model in the high dimensional feature space \mathcal{F} is given by

$$\Phi(\mathbf{y}) = \mathbf{M}_\Phi \boldsymbol{\alpha} + \mathbf{n}_\Phi, \quad (21)$$

where \mathbf{M}_Φ is a matrix whose columns are the endmember's spectra in the feature space and \mathbf{n}_Φ is an additive zero-mean Gaussian noise with covariance matrix $\sigma^2 \mathbf{I}_\Phi$ and \mathbf{I}_Φ is the identity matrix in the feature space. The model (21) can also be rewritten as

$$\Phi(\mathbf{y}) = \Phi(\mathbf{d})\alpha_p + \mathbf{B}_\Phi \boldsymbol{\gamma}_\Phi + \mathbf{n}_\Phi, \quad (22)$$

where $\Phi(\mathbf{d})$ represent the spectral signature of the desired target in the feature space and the columns of \mathbf{B}_Φ represent the undesired background signatures in the feature space.

The output of the OSP classifier in the feature space is given by

$$D_{OSP\Phi} = \mathbf{q}_{OSP\Phi}^T \Phi(\mathbf{y}) = \Phi(\mathbf{d})^T (\mathbf{I}_\Phi - \mathbf{B}_\Phi \mathbf{B}_\Phi^\#) \Phi(\mathbf{y}). \quad (23)$$

This output (23) is very similar to the numerator of (7). It can easily be shown that the kernelized version of (23) is given by

$$D_{KOSP} = \mathbf{K}(\mathbf{Z}_{Bd}, \mathbf{d})^T \Delta \Delta^T \mathbf{K}(\mathbf{Z}_{Bd}, \mathbf{d}) - \mathbf{K}(\mathbf{Z}_B, \mathbf{d})^T \mathbf{B} \mathbf{B}^T \mathbf{K}(\mathbf{Z}_B, \mathbf{d}) \quad (24)$$

where $\mathbf{Z}_B = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ correspond to N input background spectral signatures and $\mathbf{B} = (\beta_1, \beta_2, \dots, \beta_{N_b})$ are the N_b significant eigenvectors of the centered kernel matrix (Gram matrix) $\mathbf{K}(\mathbf{Z}_B, \mathbf{Z}_B)$ normalized by the square root of their corresponding eigenvalues.⁸ $\mathbf{K}(\mathbf{Z}_B, \mathbf{y})$ and $\mathbf{K}(\mathbf{Z}_B, \mathbf{d})$ are column vectors whose entries are $k(\mathbf{x}_i, \mathbf{y})$ and $k(\mathbf{x}_i, \mathbf{d})$ for $\mathbf{x}_i \in \mathbf{Z}_B$, respectively. $\mathbf{Z}_{Bd} = \mathbf{Z}_B \cup \mathbf{d}$ and Δ is a matrix whose columns are the N_{bd} eigenvectors of the centered kernel matrix $\mathbf{K}(\mathbf{Z}_{Bd}, \mathbf{Z}_{Bd}) = (\mathbf{K})_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, $\mathbf{x}_i, \mathbf{x}_j \in \mathbf{Z}_{Bd}$ with nonzero eigenvalues, normalized by the square root of their associated eigenvalues. Also $\mathbf{K}(\mathbf{Z}_{Bd}, \mathbf{y})$ is the concatenated vector $[\mathbf{K}(\mathbf{Z}_B, \mathbf{y})^T \mathbf{K}(\mathbf{d}, \mathbf{y})^T]^T$ and $\mathbf{K}(\mathbf{Z}_{Bd}, \mathbf{d})$ is the concatenated vector $[\mathbf{K}(\mathbf{Z}_B, \mathbf{d})^T \mathbf{K}(\mathbf{d}, \mathbf{d})^T]^T$. In the above derivation (24) we assumed that the mapped input data was centered in the feature space. The kernel matrices and the empirical kernel maps have to be properly centered as was shown in subsection 3.2.

5. Linear SMF and Kernel Spectral Matched Filter

5.1. Linear Spectral Matched Filter

In this section, we introduce the concept of linear SMF. The constrained least squares approach is used to derive the linear SMF. Let the input spectral signal \mathbf{x} be $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_p]$ consisting of p spectral bands. We can model each spectral observation as a linear combination of the target spectral signature and noise

$$\mathbf{x} = a\mathbf{s} + \mathbf{n} \quad (25)$$

where a is an attenuation constant (target abundance measure). When $a=0$ no target is present and when $a>0$ target is present, vector $\mathbf{s} = [s_1, s_2, \dots, s_p]^T$ contains the spectral signature of the target and vector \mathbf{n} contains the added background clutter noise.

Let us define \mathbf{X} to be a $p \times N$ matrix of the N mean-removed background reference pixels (centered) obtained from the input image. Let each centered observation spectral pixel to be represented as a column in the sample matrix \mathbf{X}

$$\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_p]. \quad (26)$$

We can design a linear matched filter such that the desired target signal \mathbf{s} is passed through while the average filter output energy is minimized. The solution to this minimization problem was shown in¹⁴ and was called Constrained Energy Minimization (CEM) filter. The output of the linear matched filter for a test input \mathbf{r} , given the estimated covariance matrix is given by

$$\mathbf{y}_r = \mathbf{w}^T \mathbf{r} = \frac{\mathbf{s}^T \hat{\mathbf{C}}^{-1} \mathbf{r}}{\mathbf{s}^T \hat{\mathbf{C}}^{-1} \mathbf{s}}, \quad (27)$$

where $\hat{\mathbf{C}}$ is the estimated covariance matrix. In^{4,5} it was shown that using the GLRT the same expression for the linear matched filter (27) can be obtained.

5.2. SMF in Feature Space and its Kernel Version

Consider the linear model of the input data in a kernel feature space which is equivalent to a non-linear model in the input space

$$\Phi(\mathbf{r}) = a_\Phi \Phi(\mathbf{s}) + \mathbf{n}_\Phi, \quad (28)$$

where Φ is the non-linear mapping that maps the input data into a kernel feature space, a_Φ is an attenuation constant (abundance measure), the high dimensional vector $\Phi(\mathbf{s})$ contains the spectral signature of the target in the feature space, and vector \mathbf{n}_Φ contains the added noise in the feature space.

Using the constrained least squares approach it can easily be shown that the output of the desired matched filter for the input $\Phi(\mathbf{r})$ is given by

$$\mathbf{y}_{\Phi(\mathbf{r})} = \frac{\Phi(\mathbf{s})^T \hat{\mathbf{C}}_\Phi^{-1} \Phi(\mathbf{r})}{\Phi(\mathbf{s})^T \hat{\mathbf{C}}_\Phi^{-1} \Phi(\mathbf{s})}, \quad (29)$$

where $\hat{\mathbf{C}}_\Phi$ is the estimated covariance of pixels in the feature space.

We now show how to kernelize the matched filter expression (29) where the resulting non-linear matched filter is called the kernel matched filter. The pseudoinverse (inverse) of the estimated background covariance matrix can be written in terms of its eigenvector decomposition as¹⁰

$$\hat{\mathbf{C}}_\Phi^{-1} = \mathbf{X}_\Phi \mathbf{B} \mathbf{A}^{-1} \mathbf{B}^T \mathbf{X}_\Phi^T, \quad (30)$$

where $\mathbf{X}_\Phi = [\Phi(\mathbf{x}_1) \Phi(\mathbf{x}_2) \dots \Phi(\mathbf{x}_N)]$ is a matrix whose columns are the mapped background reference data in the feature space and $\mathbf{B} = [\beta^1 \beta^2 \dots \beta^N]$ are the nonzero eigenvectors of the

centered kernel matrix (Gram matrix) $\mathbf{K}(\mathbf{x}, \mathbf{x})$ normalized by the square root of their corresponding eigenvalues.

Inserting Eq. (30) into (29) it can be rewritten as

$$y_{\Phi(\mathbf{r})} = \frac{\Phi(\mathbf{s})^T \mathbf{X}_{\Phi} \mathbf{B} \Lambda^{-1} \mathbf{B}^T \mathbf{X}_{\Phi}^T \Phi(\mathbf{r})}{\Phi(\mathbf{s})^T \mathbf{X}_{\Phi} \mathbf{B} \Lambda^{-1} \mathbf{B}^T \mathbf{X}_{\Phi}^T \Phi(\mathbf{s})} \quad (31)$$

Also using the properties of the Kernel PCA,⁷ we have the relationship $\mathbf{K}^{-1} = (1/N) \mathbf{B} \Lambda^{-1} \mathbf{B}^T$. We denote $\mathbf{K}(\mathbf{X}, \mathbf{X}) = (\mathbf{K})_{ij}$ an $N \times N$ Gram kernel matrix whose entries are the dot products $\langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle$. Finally, the kernelized version of SMF is now given by

$$y_{\mathbf{K}_r} = \frac{\mathbf{K}(\mathbf{X}, \mathbf{s})^T \mathbf{K}^{-1} \mathbf{K}(\mathbf{X}, \mathbf{r})}{\mathbf{K}(\mathbf{X}, \mathbf{s})^T \mathbf{K}^{-1} \mathbf{K}(\mathbf{X}, \mathbf{s})} = \frac{\mathbf{K}_s^T \mathbf{K}^{-1} \mathbf{K}_r}{\mathbf{K}_s^T \mathbf{K}^{-1} \mathbf{K}_s} \quad (32)$$

where the empirical kernel maps $\mathbf{k}_s = \mathbf{k}(\mathbf{X}, \mathbf{s})$ and $\mathbf{k}_r = \mathbf{k}(\mathbf{X}, \mathbf{r})$. As in the previous section the kernel matrix \mathbf{k} as well as the empirical kernel maps need to be properly centered.

6. Adaptive Subspace Detector and Kernel Adaptive Subspace Detector

6.1. Linear ASD

In this section, the GLRT under the two competing hypotheses H_0 and H_1 for a certain mixture model is described. The subpixel detection model for a measurement \mathbf{x} (a pixel vector) is expressed as

$$\begin{aligned} H_0 : \mathbf{x} &= \mathbf{n}, & \text{Target absent} \\ H_1 : \mathbf{x} &= \mathbf{U}\boldsymbol{\theta} + \sigma \mathbf{n} & \text{Target present} \end{aligned} \quad (33)$$

where \mathbf{U} represents an orthogonal matrix whose column vectors are the eigenvectors that span the target subspace $\langle \mathbf{U} \rangle$; $\boldsymbol{\theta}$ is an unknown vector whose entries are coefficients that account for the abundances of the corresponding column vectors of \mathbf{U} ; \mathbf{n} represents Gaussian random noise distributed as $\mathcal{N}(\mathbf{0}, \mathbf{C})$.

In the model, \mathbf{x} is assumed to be a background noise under H_0 and a linear combination of a target subspace signal and a scaled background noise, distributed as $\mathcal{N}(\mathbf{U}\boldsymbol{\theta}, \sigma^2 \mathbf{C})$ under H_1 . The background noise under the two hypotheses is represented by the same covariance but different variances because of the existence of subpixel targets under H_1 . The GLRT for the subpixel problem as described in⁵ (so called ASD) is given by

$$\mathbf{D}_{ASD}(\mathbf{x}) = \frac{\mathbf{x}^T \hat{\mathbf{C}}^{-1} \mathbf{U} (\mathbf{U}^T \hat{\mathbf{C}}^{-1} \mathbf{U})^{-1} \mathbf{U}^T \hat{\mathbf{C}}^{-1} \mathbf{x}}{\mathbf{x}^T \hat{\mathbf{C}}^{-1} \mathbf{x}} \underset{H_0}{\overset{H_1}{>}} \eta \quad (34)$$

where $\hat{\mathbf{C}}$ is the MLE (maximum likelihood estimate) of the covariance \mathbf{C} and η represents a threshold.

6.2. ASD in the Feature Space and its Kernel Version

We define a new subpixel model by assuming that the input data has been implicitly mapped by a nonlinear function Φ into a high dimensional feature space \mathcal{F} . The model in \mathcal{F} is then given by

$$\begin{aligned} H_{0\Phi} : \Phi(\mathbf{x}) &= \mathbf{n}_\Phi, & \text{Target absent} & \quad (35) \\ H_{1\Phi} : \Phi(\mathbf{x}) &= \mathbf{U}_\Phi \boldsymbol{\theta}_\Phi + \sigma_\Phi \mathbf{n}_\Phi, & \text{Target present} & \end{aligned}$$

where \mathbf{U}_Φ represents a full-rank matrix whose M_l column vectors are the eigenvectors that span target subspace $\langle \mathbf{U}_\Phi \rangle$ in \mathcal{F} ; $\boldsymbol{\theta}_\Phi$ is unknown vectors whose entries are coefficients that account for the abundances of the corresponding column vectors of \mathbf{U}_Φ ; \mathbf{n}_Φ represents Gaussian random noise distributed by $\mathcal{N}(0, \mathbf{C}_\Phi)$; and σ_Φ is the noise variance under $H_{1\Phi}$. The GLRT for the model (35) in \mathcal{F} is now given by

$$D_{ASD}(\Phi(\mathbf{x})) = \frac{\Phi(\mathbf{x})^T \hat{\mathbf{C}}_\Phi^{-1} \mathbf{U}_\Phi (\mathbf{U}_\Phi^T \hat{\mathbf{C}}_\Phi^{-1} \mathbf{U}_\Phi)^{-1} \mathbf{U}_\Phi^T \hat{\mathbf{C}}_\Phi^{-1} \Phi(\mathbf{x})}{\Phi(\mathbf{x})^T \hat{\mathbf{C}}_\Phi^{-1} \Phi(\mathbf{x})}, \quad (36)$$

where $\hat{\mathbf{C}}_\Phi$ is the MLE of \mathbf{C}_Φ .

The kernelized expression of (36) is given by¹²

$$D_{KASD}(\mathbf{x}) = \frac{\mathbf{K}_x [\Gamma^T \mathbf{K}(\mathbf{X}, \mathbf{Y})^T \mathbf{K}(\mathbf{X}, \mathbf{X})^{-1} \mathbf{K}(\mathbf{X}, \mathbf{Y}) \Gamma]^{-1} \mathbf{K}_x^T}{\mathbf{k}(\mathbf{x}, \mathbf{X})^T \mathbf{K}(\mathbf{X}, \mathbf{X})^{-1} \mathbf{k}(\mathbf{x}, \mathbf{X})}, \quad (37)$$

where $\mathbf{K}_x = \mathbf{k}(\mathbf{x}, \mathbf{X})^T \mathbf{K}(\mathbf{X}, \mathbf{X})^{-1} \mathbf{K}(\mathbf{X}, \mathbf{Y}) \Gamma$, background spectral signatures is denoted by $\mathbf{X} = [x_1 \ x_2 \ \dots \ x_N]$, target spectral signatures are denoted by $\mathbf{Y} = [y_1 \ y_2 \ \dots \ y_M]$ and $\Gamma = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_{M_l}]$, $M_l < M$, is a matrix consisting of the M_l eigenvectors of the kernel matrix $\mathbf{K}(\mathbf{X}, \mathbf{Y})$. As in the previous section, all the kernel matrices as well as the empirical kernel maps need to be properly centered.⁷

7. Experimental Results

In this section, the kernel-based matched signal detectors, such as the kernel MSD (KMSD), kernel ASD (KASD), kernel OSP (KOSP) and kernel SMF (KSMF) as well as the corresponding conventional detectors are implemented based on two different types of data sets -- illustrative toy data sets and a real hyperspectral image that contains military targets. The Gaussian RBF kernel, $\mathbf{k}(\mathbf{x}, \mathbf{y}) = \exp(-\|\mathbf{x} - \mathbf{y}\|^2 / c)$, was used to implement the kernel-based detectors. c represents the width of the Gaussian distribution and the value of c was chosen such that the overall data variations can be fully exploited by the Gaussian RBF function. In this paper, the values of c were determined experimentally.

7.1. Illustrative Toy Examples

Figs. 1 and 2 show contour and surface plots of the conventional detectors and the kernel-based detectors, on two different types of two-dimensional toy data sets: a Gaussian mixture in Fig. 1 and nonlinearly mapped data in Fig. 2. In the contour and surface plots,

data points for the desired target were represented by the star-shaped symbol and the background points were represented by the circles. In Fig. 2 the two-dimensional data points $x=(x,y)$ for each class were obtained by nonlinearly mapping the original Gaussian mixture data points $x_0=(x_0,y_0)$ in Fig. 1. All the data points in Fig. 2 were nonlinearly mapped by $x_0=(x,y)=(x_0,x_0^2+y_0)$. In the new data set the second component of each data point is nonlinearly related to its first component.

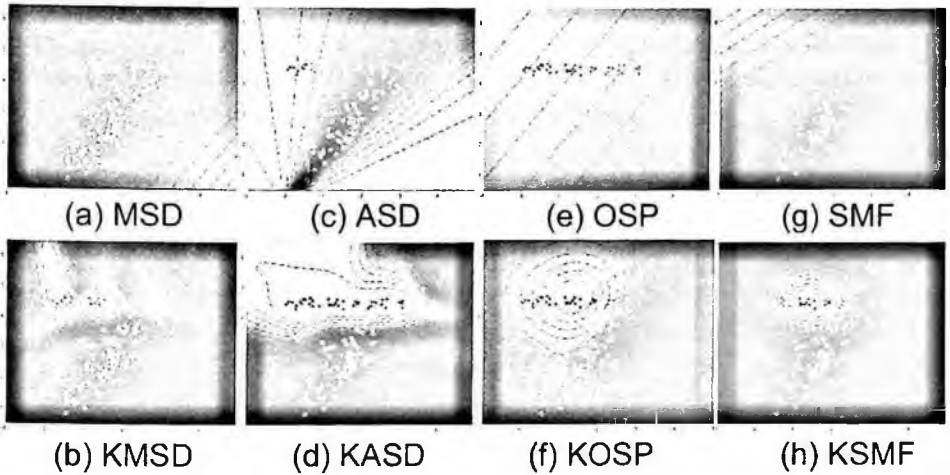


Fig.1. Contour and surface plots of the conventional matched signal detectors and their kernel versions on a toy dataset (a mixture of Gaussian).

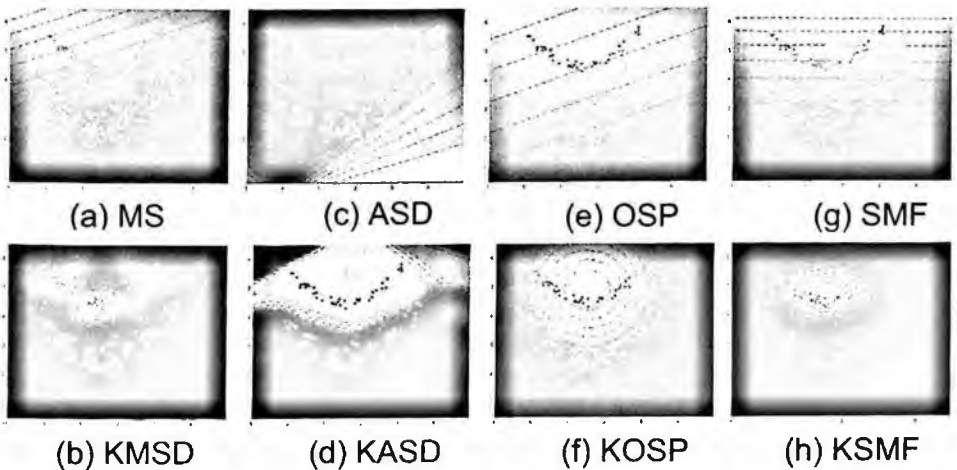


Fig.2. Contour and surface plots of the conventional matched signal detectors and their kernel versions on a toy dataset: in this toy example, the Gaussian mixture data shown in Fig. 1 was modified to generate nonlinearly mixed data.

For both data sets, the contours generated by the kernel-based detectors are highly nonlinear and naturally following the dispersion of the data and thus successfully separating the two classes, as opposed to the linear contours obtained by the conventional detectors. Therefore, the kernel-based detectors clearly provided significantly improved discrimination over the conventional detectors for both the Gaussian mixture and nonlinearly mapped data. Among the kernel-based detectors, KMSD and KASD outperform KOSP and KSMF mainly because targets in KMSD and KASD are better represented by the associated target subspace than by a single spectral signature used in KOSP and KSMF. Note that the contour plots for MSD (Fig. 1(a) and Fig. 2(a)) represent only the numerator of Eq. 4 because the denominator becomes unstable for the two-dimensional cases: i.e., for the two-dimensional data $(\mathbf{I} - \mathbf{P}_{\text{TB}})$ becomes zero.

7.2. Hyperspectral Images

In this section, a HYDICE (HYperspectral Digital Imagery Collection Experiment) image from the Desert Radiance II data collection (DR-II) was used to compare detection performance between the kernel-based and conventional methods. The HYDICE imaging sensor generates 210 bands across the whole spectral range (0.4 -- 2.5 μm) which includes the visible and short-wave infrared (SWIR) bands. But we only use 150 bands by discarding water absorption and low signal to noise ratio (SNR) bands; the spectral bands used are the 23rd--101st, 109th--136th, and 152nd--194th for the HYDICE images. The DR-II image includes 6 military targets along the road, as shown in the sample band images in Fig. 3. The detection performance of the DR-II image was provided in both the qualitative and quantitative -- the receiver operating characteristics (ROC) curves -- forms. The spectral signatures of the desired target and undesired background signatures were directly collected from the given hyperspectral data to implement both the kernel-based and conventional detectors.

All pixel vectors in a test image are first normalized by a constant, which is a maximum value obtained from all the spectral components of the spectral vectors in the corresponding test image, so that the entries of the normalized pixel vectors fit into the interval of spectral values between zero and one. The rescaling of pixel vectors was mainly performed to effectively utilize the dynamic range of Gaussian RBF kernel.

Figs. 4-5 show the detection results including the ROC curves generated by applying the kernel-based and conventional detectors to the DR-II image. In general, the detected targets by the kernel-based detectors are much more evident than the ones detected by the conventional detectors, as shown in Fig. 4. Fig. 5 shows the ROC curve plots for the kernel-based and conventional detectors; the kernel-based detectors clearly outperformed the conventional detectors. In particular, KMSD performed the best of all the kernel-based detectors detecting all the targets and significantly suppressing the background. The performance superiority of KMSD is mainly attributed to the utilization of both the target and background kernel subspaces representing the target and background signals in the feature space, respectively.

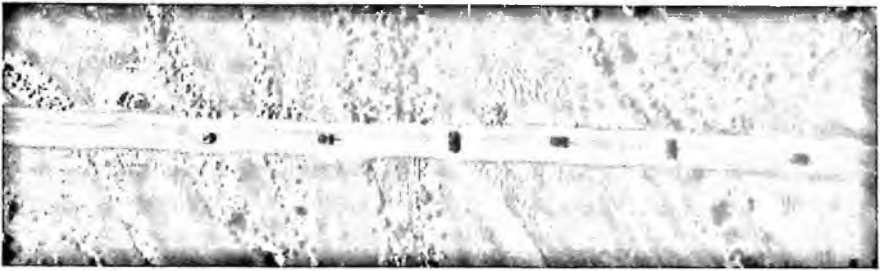


Fig. 3. A sample band from DR-II data.

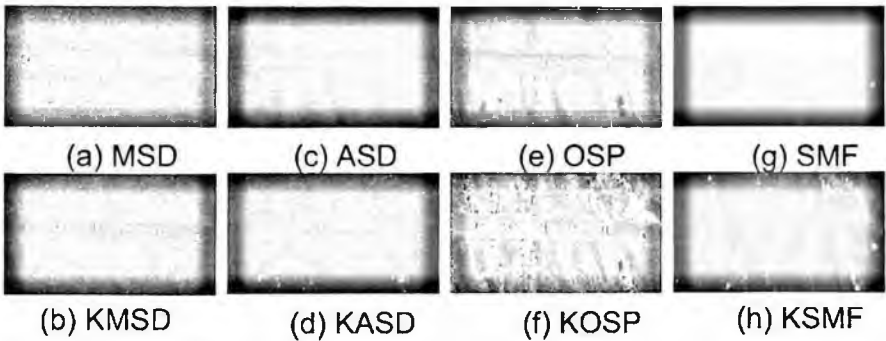


Fig.4. Detection results for the DR-II image using the conventional detectors and the corresponding kernel versions.

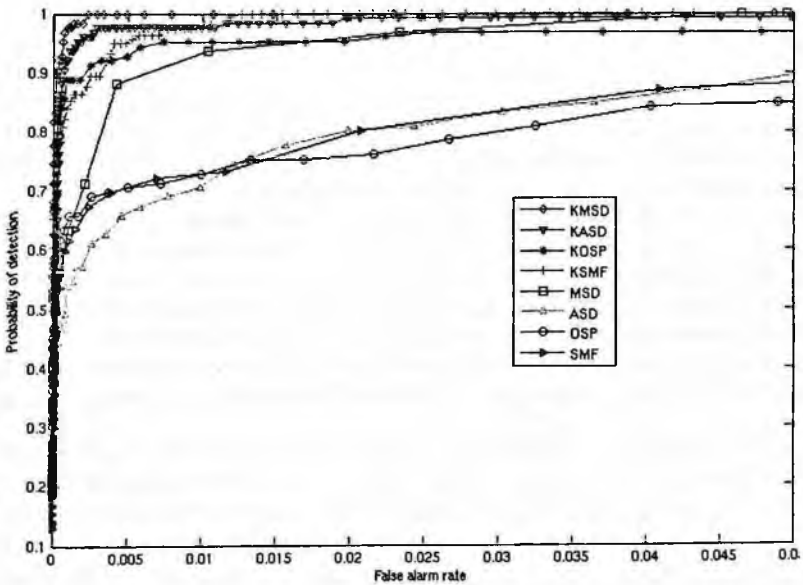


Fig.5. ROC curves obtained by conventional detectors and the corresponding kernel versions for the DR-II image.

8. Conclusions

In this paper, nonlinear versions of several matched signal detectors, such as KMSD, KOSP, KSMF and KASD have been implemented using the kernel-based learning theory. Performance comparison between the matched signal detectors and their corresponding nonlinear versions was conducted based on two-dimensional toy-examples as well as a real hyperspectral image. It is shown that the kernel-based nonlinear versions of these detectors outperform the linear versions.

References

1. L. L. Scharf and B. Friedlander, Matched subspace detectors, *IEEE Trans. Signal Process.* **42**(8):2146-2157, (1994).
2. J. C. Harsanyi and C.-I. Chang, Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach, *IEEE Trans. Geosci. Remote Sensing* **32**(4):779-785, (1994).
3. Δ. Μανολακίς, Γ. Σηαω ανδ Ν. Κεσηαα, Χομπαρατιπε αναλψισ οφ ηψπερσπεχτρα λαδαπιπε ματχηεδ φίλτερ δετεχτορ, *Προχ. ΣΠΙΕ*, πολυμε 4049, παγεσ 2-17, (2000).
4. F. C. Robey, D. R. Fuhrmann, E. J. Kelly, an R. Nitzberg, A CFAR adaptive matched filter detector, *IEEE Trans. on Aerospace and Elect. Syst.* **28**(1):208-216, (1992).
5. S. Kraut and L. L. Scharf. The CFAR adaptive subspace detector is a scale-invariant GLRT, *IEEE Trans. Signal Process.* **47**(9):2538-2541, (1999).
6. S. Kraut, L. L. Scharf and T. McWhorter, Adaptive subspace detectors, *IEEE Trans. Signal Process.* **49**(1):1-16, (2001).
7. B. Scholkopf and A. J. Smola, *Learning with Kernels* (The MIT Press 2002).
8. B. Scholkopf, A. J. Smola and K.-R. Muller. Nonlinear component analysis as a kernel eigenvalue problem, *Neural Computation*, **10**(5):1299-1319, (1998).
9. G. Baudat and F. Anouar, Generalized discriminant analysis using a kernel approach, *Neural Computation*, **12**(12):2385-2404, (2000).
10. A. Ruiz and E. Lopez-de Teruel, kernel-based statistical pattern analysis, *IEEE Trans. on Neural Networks*, **12**(1):16-32, (2001).
11. H. Kwon and N. M. Nasrabadi, Kernel matched subspace detectors for hyperspectral target detection. *IEEE Trans. Pattern Anal. Machine Intell.* **28**(2):178-194, (2006).
12. H. Kwon and N. M. Nasrabadi, Kernel adaptive subspace detector for hyperspectral target detection, *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, Philadelphia, PA, USA, March (2005).
13. H. Kwon and N. M. Nasrabadi, Kernel RX-algorithm : A nonlinear anomaly detector for hyperspectral imagery, *IEEE Trans. Geosci. Remote Sensing*, **43**(2):388-397, (2005).
14. J. C. Harsanyi, Detection and Classification of Subpixel Spectral Signatures in Hyperspectral Image Sequences, Ph.D. thesis, Univ. of Maryland, Baltimore County, USA, (1993).

VIBRATIONAL SPECTROSCOPY OF CHEMICAL AGENTS SIMULANTS, DEGRADATION PRODUCTS OF CHEMICAL AGENTS AND TOXIC INDUSTRIAL COMPOUNDS

SAMUEL P. HERNÁNDEZ-RIVERA, LEONARDO C. PACHECO-LONDOÑO,
OLIVA M. PRIMERA-PEDROZO, ORLANDO RUIZ, YADIRA SOTO-FELICIANO and WILLIAM ORTIZ

MURI Center for Chemical Sensors Development

Department of Chemistry

University of Puerto Rico-Mayagüez

PO Box 9019, Mayagüez, PR 00681

Corresponding author: sp_hernandez_uprm@yahoo.com

This paper focuses on the measurement of spectroscopic signatures of Chemical Warfare Agent Simulants (CWAS), degradation products of chemical agents and Toxic Industrial Compounds (TIC) using vibrational spectroscopy. Raman Microscopy, Fourier Transform Infrared Spectroscopy in liquid and gas phase and Fiber Optics Coupled-Grazing Angle Probe-FTIR were used to characterize the spectroscopic information of target threat agents. *Ab initio* chemical calculations of energy minimization and FTIR spectra of Chemical Warfare Agents were accompanied by Cluster Analysis to correlate spectral information of real agents and simulants.

Keywords: Chemical Warfare Agents Simulants, degradation compounds, Toxic Industrial Compounds, Cluster Analysis

1. Introduction

Since September 11, 2001 the world focuses its attention to the possibility of use of chemical and biological threat agents by terrorist organizations directed against troops and/or civilians. Countering future attacks requires a wide array of detection capabilities for a range of potential deployment scenarios^{1,2}. The Chemical Warfare Agents (CWA) most likely to be used as terrorist weapons, according to terrorism experts include nerve agents, blister agents, choking agents, and blood agents. The nerve agents are a group of particularly toxic chemical warfare agents³⁻⁴. They were developed just before and during

World War II and are related chemically to the organophosphorus insecticides. The principle agents in this group are: GA (Tabun), GB (Sarin), GD (Soman), GF and VX (methylphosphonothioic acid). Table 1 lists the most well known chemical agents. Although it is expected that chemical agents will be used mainly against personnel, food and water supplies may easily become contaminated and may be contaminated deliberately. This may occur from contact with chemical agents in the form of vapor, aerosol, drops or splashes of liquid or particulate smokes⁵. The characterization of hydrolysis and or degradation products of CWA provides the capability to detect possible attacks to water systems^{2,3}.

Various chemicals with potential terrorist groups usage have been identified as Toxic Industrial Compounds. TIC's are chemicals toxic to humans that are widely used in manufacturing or primary material processing. They have received more attention in recent years due to ease of accessibility of large quantities by potential terrorists¹. Chemical agents (CWA) differ from TIC's in that they are intended to immediately incapacitate as many soldiers as possible when released against an enemy in war. To do this, the CWA must be toxic enough to cause an instant response when it is inhaled or comes into contact with the skin.

Table 1. Chemical Warfare Agents: structure and Chemical Abstracts Service registry number.

CWA	STRUCTURE	CAS NO.
Sarin, GB	$\begin{array}{c} \text{O} \quad \text{CH}_3 \\ \parallel \quad \\ \text{H}_3\text{C}-\text{P}-\text{O}-\text{CH}-\text{CH}_3 \\ \\ \text{F} \end{array}$	107-44-8
Mustard Gas, HD	$\text{ClCH}_2\text{CH}_2-\text{S}-\text{CH}_2\text{CH}_2\text{Cl}$	505-60-2
Soman, GD	$\begin{array}{c} \text{O} \quad \text{CH}_3 \quad \text{CH}_3 \\ \parallel \quad \quad \\ \text{H}_3\text{C}-\text{P}-\text{O}-\text{CH}-\text{C}-\text{CH}_3 \\ \quad \quad \\ \text{F} \quad \quad \text{CH}_3 \end{array}$	96-64-0
Tabun, GA	$\begin{array}{c} \text{O} \\ \parallel \\ \text{CH}_3\text{CH}_2-\text{O}-\text{P}-\text{C}\equiv\text{N} \\ \\ \text{CH}_3-\text{N}-\text{CH}_3 \end{array}$	77-81-6
VX	$\begin{array}{c} \text{O} \\ \parallel \\ \text{CH}_3\text{CH}_2-\text{O}-\text{P}-\text{CH}_3 \\ \\ (\text{Isopropyl})_2\text{NCH}_2\text{CH}_2-\text{S} \end{array}$	5078269-9

Several techniques have been employed to detect and identify chemical agents including HPLC/MS⁶, GC/MS⁷⁻⁸, IMS⁹⁻¹⁰, Infrared Spectroscopy, Raman Spectroscopy²⁻¹¹, among others. Vibrational spectroscopy has the advantage of providing chemical information and has the potential for remote sensing. Chemical Point Detection, although limited in spatial coverage, provides the sensitivity and selectivity required for CWA detection. Technology could vastly improve by coupling with wide area, early detection standoff and near field proximity detectors such as infrared detectors that could be used to continuously monitor air in orthogonal sensor modality. Basic research is required in detection technology using vibrational spectroscopy as Point/Proximity Detectors. This promises to open new sensing avenues of CWA *in situ* with added bonuses of identification and quantification in real-time frames. Vehicle borne and possibly, man portable systems could soon be within the reach of Defense and Security Organizations and in particular, to the soldier. Raman spectroscopy, in all its forms, is a vibrational spectroscopic method that has the inherent ability to distinguish between molecules with high degrees of similarity.

In this paper several Chemical Warfare Agents Simulants (CWAS), some of their hydrolysis and degradation products, and Toxic Industrial Compounds (TIC's) were analyzed using vibrational spectroscopy. The objective of this work is to characterize the spectroscopic signatures of these threat compounds and to demonstrate, at laboratory scale, the capability of Infrared and Raman spectroscopy for generating methodologies for detection of CWA and TIC's.

2. Experimental

2.1. Materials and Reagents

Due to the high toxicity of chemical agents, structural analogs with less toxicity were used for spectroscopic simulation of the real CWA. Various chemical warfare agent simulants and TIC's were analyzed to establish the analytical response to Raman and Infrared spectroscopy. A total of six CWA, six CWA degradation products and seven TIC were studied. Table 1-4 summarizes the list of the compounds studied. The first six compounds are considered CWA simulants: TEP, DMMP and DIMP are GB simulants; 2-CEES and 1,6-DCH are HD simulants and 2-BAET is a VX simulant. 98% 1,4-thioxane, 98% cyclohexylamine, 99% cycloheptylamine, 99% dipropylamine, 97% dihexylamine, 98% 2-chloroethyl sulfide, thiodiglycol sulfoxide and 2-(butylamino)-ethanethiol were obtained from Sigma Aldrich Chemical Company, St. Louis, MO. Dichloromethane, chloroform, ethylene glycol, carbon disulfide, dimethyl sulfoxide, *N,N*-dimethylformamide and dimethylmethyl phosphonate were purchased from Fisher Scientific International, Chicago, IL. Triethylphosphate and acetone were obtained from Across Organics, Fisher Scientific International, Chicago, IL. 1,6-dichlorohexane, diisopropyl methylphosphonate were purchased from Alfa Aesar, Ward Hill, MA

2.2. Raman Experiments

For Raman Vibrational Analysis a Renishaw Raman Microscope RM2000 was used. This system was equipped with CCD viewing camera aligned with the microscope that

allowed capturing images of the sample with visible light and used to focus the laser on the sample. The monochromatic excitation sources used were a Coherent INNOVA 310-8 Ar⁺ laser that produced a beam of a 514 nm (green), 488 nm (Blue), a Coherent 899 RING titanium-sapphire laser that produced a beam of 785 nm (Red) and a Coherent diode laser Verdi-V6 (532 nm, green). For the deep ultraviolet excitation (DUV) a frequency doubled 488 nm Coherent FRED laser system operating at 244 nm was used. A Jobin-Yvon LABRAM HR-UV was employed for the UV measurements. Samples of standards of CWA's and TIC's were placed in a capillary glass tube for VIS and NIR Raman excitation and a quartz capillary tube was used for UV experiments. Spectra were recorded in the Raman Shift range of 100-3600 cm⁻¹ at the same laser power for all the excitation lines. The laser beams were focused onto the samples through objective with 10x magnifications. An integration time of 10 seconds was used. Two acquisitions were averaged per scan in order to improve the signal to noise ratio.

2.3. Infrared Experiments

2.3.1. KBr Cells Analysis

Samples were prepared by using a KBr cell. Salt plates were cleaned before the experiments using Acetone. Using a micropipette, a drop of the CWA was placed on one salt plate. Then the second plate was put over it and the liquid was spread into a thin film. Samples were analyzed in a Bruker Vector 22 FTIR with a DTGS detector. The Infrared experimental conditions were 20 scans and a 4 cm⁻¹ resolution. A scanned background spectrum was acquired before each measurement session using a clean test salt plate at the same instrumental conditions used for sample spectra acquisition. All spectra were recorded in absorbance mode.

2.3.2. IR Surface Analysis

Threat compounds were analyzed as traces deposited on surfaces. To carry out these experiments, samples were deposited on thin stainless steel plates by applying a known amount of the compound to the surface. Taking into account the surface area covered by the target analyte, the amount of the compound was weighted and diluted in order to have loading concentrations of 5 µg/cm² and 10 µg/cm². Samples were diluted with Methanol. Aliquots of 20 µL were deposited on one end of the plate. Methanol was used as the solvent. Using a Teflon sheet, the samples were smeared from one plate end to the other^{12, 13}. Finally, the samples were allowed to dry for a few seconds before continuing on with the spectroscopic measurements.

Measurements of IR spectra were done using a Remspec grazing angle probe (GAP)¹⁴⁻¹⁶ interfaced to a Bruker VECTOR-22 Fourier Transform Infrared interferometer and equipped with external, liquid nitrogen cooled, Mercury-Cadmium-Telluride (MCT) infrared detector. The IR beam was steered from the interferometer to the GAP using a high IR transmission Mid-IR fiber bundle composed of a 19-fiber chalcogenide glass optical bundle of the As-Se-Te system and was focused on the surface at approximately 80° from the surface normal. The software used for IR data acquisition was Bruker Optics OPUS™, Version 4.2. The wavenumber range measured in the experiments was

1000–4000 cm^{-1} at a 4 cm^{-1} resolution. A total of 10 scans were acquired for each spectrum.


2.3.3. Gas Phase IR Analysis

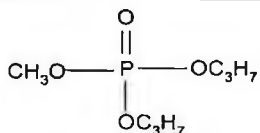
Gas phase experiments of the six CWAS were performed in a Bruker Optics FTIR model IFS 66v/S spectrometer equipped with a DTGS detector and a potassium bromide (KBr) beamsplitter. A gas cell was placed in the macro compartment and adapted to a micro pump that removed background air and transferred the sample to the cell. DMMP and DIMP were used for these analyses. One gram of the CWAS was deposited on an Erlenmeyer. A typical spectroscopic measurement averaged 20 scans at a resolution of 4 cm^{-1} in the range of 400 – 7500 cm^{-1} . Gas phase IR spectra were also acquired using Bruker Optics OPUSTTM, Version 4.2. A background of air in the cell was recorded before each run of the CWAS experiments.

3. Results

The tables that follow contain information on the Chemical Warfare Agents Simulants (Table 2), degradation products of actual chemical agents (Table 3) and on the Toxic Industrial Compounds (Table 4) studied. Raman Spectroscopy and Fourier Transform Infrared Spectroscopy were used for the vibrational analysis. Spectra were compared to data obtained from *ab initio* calculations using Gaussian 03 platform¹⁷.

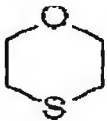
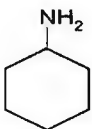
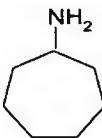
Table 2. Chemical structures and CAS number of the CAWS studied.

Chemical Warfare Agents Simulants		
Compound	Chemical Structure	CAS number
2-chloroethyl ethyl sulfide (2-CEES)	$\text{ClCH}_2\text{CH}_2\text{SCH}_2\text{CH}_3$	693-07-2
dimethyl methyl phosphonate (DMMP)	$\begin{array}{c} \text{O} \\ \\ \text{MeO}-\text{P}-\text{Me} \\ \\ \text{OMe} \end{array}$	756-79-6
triethylphosphate (TEP)	$\begin{array}{c} \text{O} \\ \\ \text{CH}_3\text{CH}_2\text{O}-\text{P}-\text{OCH}_2\text{CH}_3 \\ \\ \text{OCH}_2\text{CH}_3 \end{array}$	78-40-0
1, 6-dichlorohexane (1,6-DCH)		2163-00-0
2-(butylamino)ethanethiol (2-BAET)	$\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_2\text{NHCH}_2\text{CH}_2\text{SH}$	5842-00-2

di-isopropylmethyphosphonate
(DIMP)

1445-75-6

Table 3. Chemical structures and CAS numbers for degradation products of CWS studied.

Degradation Products of Chemical Warfare Agents		
Compound	Chemical Structure	CAS number
1,4- thioxane (1,4-TO)-HD		15980-15-1
thiodiglycol sulfoxide -HD (TDS)	$\text{OHCH}_2\text{CH}_2-\text{S}-\text{CH}_2\text{CH}_2\text{OH}$	3085-45-8
dihexylamine (DHA)	$\begin{array}{l} \text{CH}_3(\text{CH}_2)_4\text{CH}_2 \\ \quad \quad \quad \diagdown \\ \quad \quad \quad \text{NH} \\ \quad \quad \quad \diagup \\ \text{CH}_3(\text{CH}_2)_4\text{CH}_2 \end{array}$	143-16-8
dipropylamine (DPA)	$\begin{array}{l} \text{CH}_3\text{CH}_2\text{CH}_2 \\ \quad \quad \quad \diagdown \\ \quad \quad \quad \text{NH} \\ \quad \quad \quad \diagup \\ \text{CH}_3\text{CH}_2\text{CH}_2 \end{array}$	142-84-7
cyclohexylamine (CHA)		108-91-8
cycloheptylamine (CHPA)		5452-35-7

3.1. Raman Spectroscopy Measurements

Normal or spontaneous vibrational Raman spectra were obtained for all of the compounds reported in this study. Excitation wavelengths spanned from the Near IR (NIR) at 785 nm to the VIS at 532, 514.5 and 488 nm and even the deep ultraviolet region at 244 nm. These discrete lines were used to collect Raman shift spectra from 100 to 3600 cm^{-1} . By varying the excitation line, at a same laser power, it was possible to determine which incident source produced the best Raman signatures of the threat compounds studied. Typical Raman Shift spectra are shown in Figure 1 for 1,6-dichlorohexane (1,6-DHC). For this CWAS the optimum excitation lines were 488 nm and 532 nm with a laser power of 25 mW. The most prominent bands were at these incident wavelengths were located at 652 and 1444 cm^{-1} .

Table 4. Chemical structures and CAS numbers for degradation products of TIC studied.

Toxic Industrial Compounds		
Compound	Chemical Structure	CAS number
dichloromethane	CH_2Cl_2	75-09-2
acetone	$\begin{array}{c} \text{O} \\ \\ \text{CH}_3 - \text{C} - \text{CH}_3 \end{array}$	67-64-1
chloroform	CHCl_3	67-66-3
ethylene glycol	$\begin{array}{c} \text{H} \quad \text{H} \\ \quad \\ \text{H}-\text{O}-\text{C}-\text{C}-\text{O}-\text{H} \\ \quad \\ \text{H} \quad \text{H} \end{array}$	107-21-1
carbon disulfide	CS_2	75-15-0
dimethyl sulfoxide	$\begin{array}{c} \text{O} \\ \\ \text{CH}_3 - \text{S} - \text{CH}_3 \end{array}$	67-68-5
<i>N,N</i> -dimethylformamide	$\begin{array}{c} \text{O} \\ \\ \text{H} - \text{C} - \text{N} - \text{CH}_3 \\ \\ \text{CH}_3 \end{array}$	68-12-2

The Raman signature located at 652 cm^{-1} represents the dominant spectroscopic feature at 244 and 785 nm excitation and appears intense in all excitation lines studied. This band has been tentatively assigned to C-Cl stretching that also appears in the structure of CWA HD. All vibrational signatures of the chemical agent (HD) are present in the simulant, with the exception of C-S-C group vibration. The vibrational band at ca. 1444 cm^{-1} is the dominant figure under visible excitation (532 and 488 nm). Most signatures are also prominent at DUV (244 nm) excitation. The same analysis was done for the remaining CWAS, CWA degradation products and TIC and is summarized in Table 5.

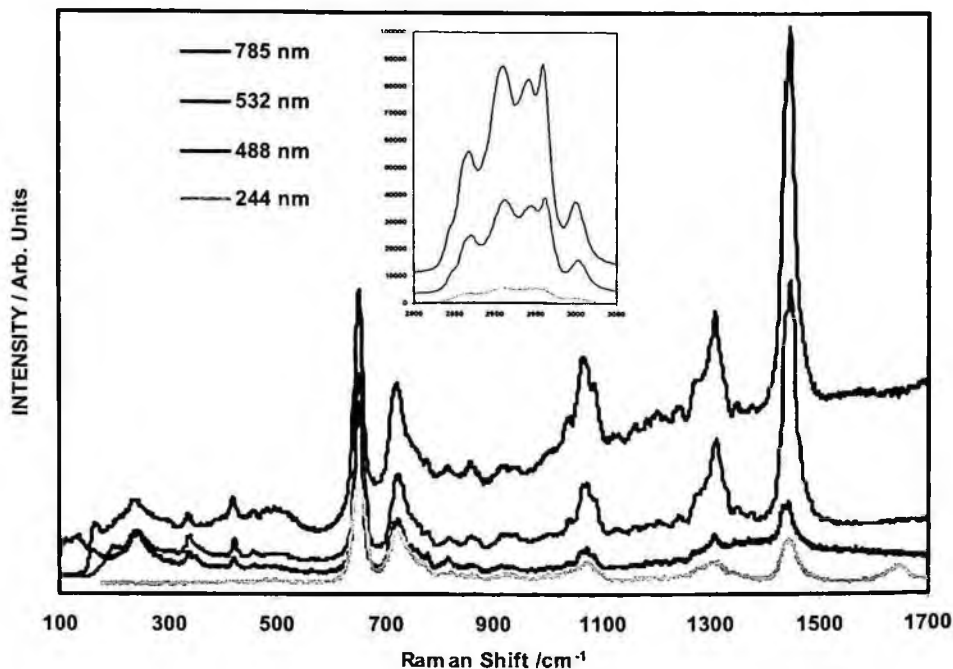


Figure 1. Raman spectra of 1,6-dichlorohexane obtained at different excitation lines.

3.2. Infrared

Infrared spectra were recorded for all threat compounds using a KBr cell. In all the cases, well defined IR signatures were obtained. Figure 2 shows the comparison of the spectra in gas phase (air) and liquid phase for DMMP and DIMP. According to the infrared spectra the region between $600\text{-}900\text{ cm}^{-1}$ can be used for analysis of these compounds in air. Besides the interferences present in air, the chemical agent's simulants could be easily detected in the vapor phase. For DIMP two predominant bands were observed, 796 cm^{-1} and 916 cm^{-1} . For DMMP bands at 775 cm^{-1} , 815 cm^{-1} and 912 cm^{-1} were detected.

When the theoretical and experimental spectra of DMMP and DIMP were compared, a good prediction was found in the infrared region for wavenumbers lower than 2000 cm^{-1} and acceptable agreement was obtained in the C-H region, as expected (see Figure 3). By

taking into account this result, we can extrapolate than between theoretical and experimental spectra of GB (Sarin) the result can also be found. Due the fact that the IR spectrum of Sarin was not recorded, the theoretical spectrum was used to do the cluster analysis together with the experimental spectra of the simulants and CWA degradation products. In addition, the theoretical spectra of GB and DMMP and DIMP were compared (Figure 4).

Table 5. Optimal conditions and Raman Bands for the CWA and TIC

Compound	Excitation line (nm)	Laser power (mW)	Most intense bands (cm ⁻¹)
1,4- thioxane (1,4-TO)	785	5	663 (a), 825 (a, b, c); 2957 (b)
2-chloroethyl ethyl sulfide (2-CEES)	785, 532	8	699 (a) 2934 (a, b)
dipropyl amine (DPA)	785, 532	8	1454 (a, b); 2936 (b)
dihexyl amine (DHA)	785, 532	8	1726 (a); 2905 (a, b)
cyclohexyl amine (CHA)	785, 532	8	1446 (a, b); 780 (a, b); 1029 (a, b)
cycloheptyl amine (CHPA)	785, 532	8	1446 (a, b); 722 (a, b); 1010 (a, b)
2-(butylamino)-ethanethiol (2-BAET)	785, 532	8	1444 (a, b); 664 (a, b) 2574 (a, b); 2934 (a, b)
dimethyl methyl phosphonate (DMMP)	785, 532, 488	20	715 (a, c); 1421 (b, c) 2900 (b, c)
triethylphosphate (TEP)	532, 488	20	734 (b); 1100 (b) 1458 (b); 2936 (c)
1,6-dichlorohexane (1,6-DCH)	785, 532 488, 244	25	1444 (b, c); 652 (a, b, c, d)
di-isopropylmethyphosphoante (DIMP)	785, 532 488, 244	20	721 (a, b, c, d); 881 (a, b, c, d) 1357 (a, b, c, d); 1455 (b, c) 2936 (b, c)
thiodiglycol sulfoxide (TDS)	785 514	20	1043 (a, e); 1460 (a, e) 2835 (a, e); 2951 (a, e)
acetone	785, 514	20	792 (e); 1452 (e); 1720 (e); 2928 (e)
chloroform	785, 514	20	258 (a); 363 (a); 664 (a)
ethylene glycol	785, 514	20	874 (e); 1102 (e), 2945 (e)
carbon disulfide	785, 514	20	659 (e); 801 (e)

dichloromethane	785, 514	20	288 (e); 709 (e); 2992 (e);
dimethyl sulfoxide	785 514	20	1678 (e); 1415 (e); 1100(e); 872 (e); 666 (e); 2939 (e);
<i>N,N</i> -dimethylformamide	785, 514	20	1403 (e); 862 (e); 656(e); 402 (e)

a = 785 nm; b = 532 nm; c = 488 nm; d = 244 nm; e = 514 nm

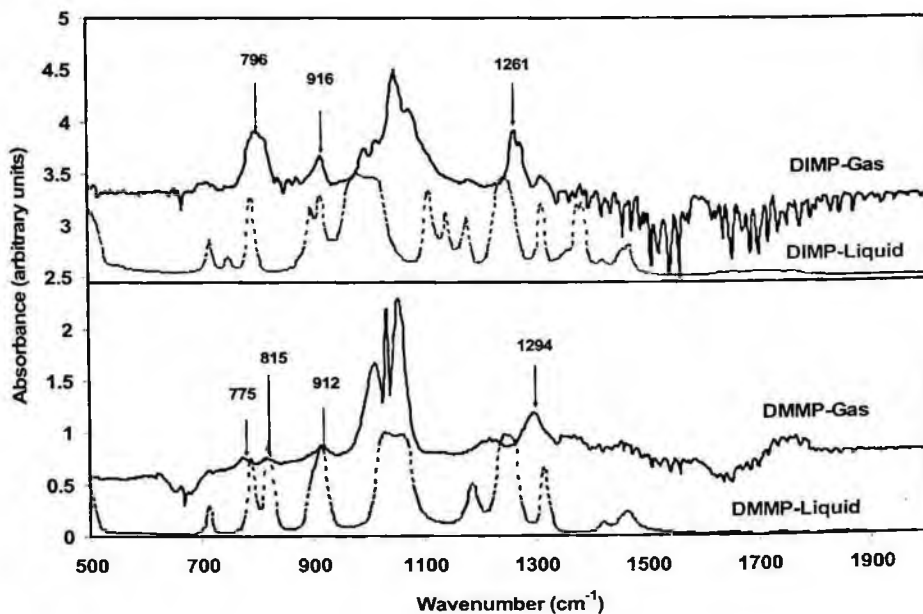


Figure 2. Infrared spectra of DIMP and DMMP in air (gas cell) and liquid (KBr cell)

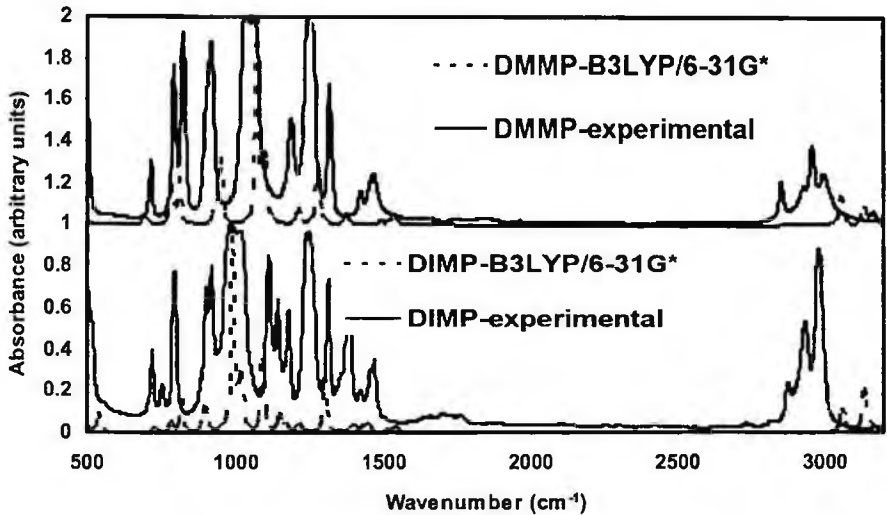


Figure 3. Comparison of theoretical and experimental spectra of DMMP and DIMP

It can be seen that the spectra are very similar. This was also observed within the experimental spectra. For HD and VX theoretical spectra were also used for the cluster analysis. The Cluster Analysis was done by making the assumption that the theoretical spectra of GB, HD and VX are more closely related to the real spectra of GB, HD and VX than the experimental spectra of the simulants and degradation products. This assumption is based on the comparisons that were done before and on the fact that this is the procedure used by most researchers where the theoretical and experimental spectra are compared using DFT²⁰.

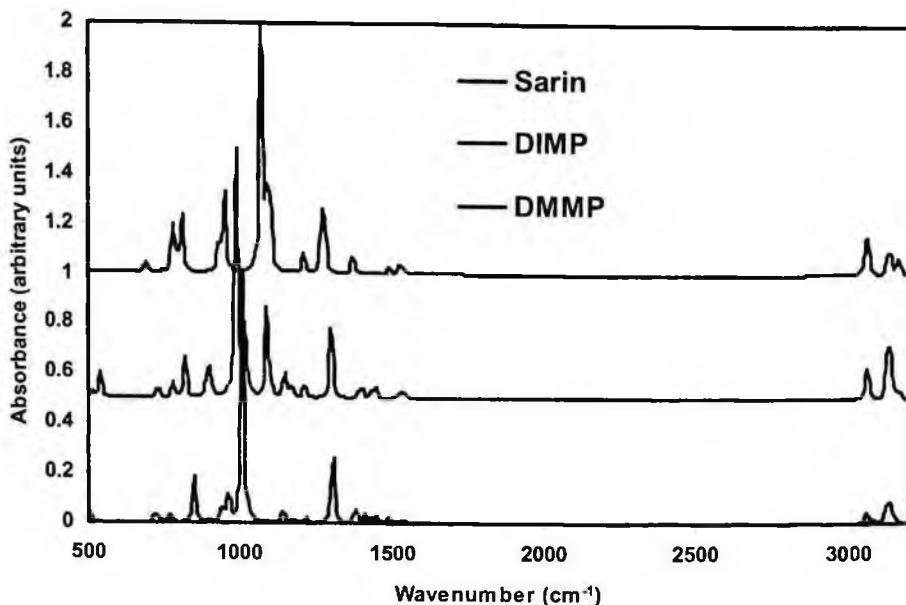


Figure 4. Comparison of theoretical spectra of DMMP and DIMP with Sarin (GB).

3.2.1. Comparison of Theoretical IR spectra of CWA with CWA simulants, CWAs degradation products

Theoretical IR spectra for Chemical Agents GB, HD and VX and for Simulants DMMP and DIMP were calculated using the Density Functional Theory (DFT) with the B3LYP method¹⁷⁻¹⁹. Target chemicals' geometries and frequencies were optimized and calculated using B3LYP/6-31G* basis set of Gaussian 03¹⁷. First, the theoretical IR spectra of DMMP and DIMP were compared with experimental IR spectra and then the theoretical IR spectra of GB were compared with theoretical IR spectra of its simulant. Second, the theoretical IR spectra of CWA and experimental IR spectra of CWA simulants were analyzed by cluster analysis in order to find similarities.

Cluster analysis.

The Spectral data were normalized (Min/Max normalization) and the spectra were shifted so that the minimum occurring y-value was set to zero. The spectra were then expanded in the y-direction so that the maximum occurring y-value was placed at 2 absorbance units. This data preprocessing was needed for all spectra (experimental and theoretical) in the same range at the y-axis. Next, the cluster analysis was generated for OPUS program for the Factorization Method²¹⁻²². This method of calculating the spectral distances by cluster analysis differs from methods based on calculating by the identity test. Overlapping frequency ranges will not be merged when using cluster analysis. An artificial spectrum is derived from the selected spectral ranges of the measured spectrum. The artificial spectrum is used for the calculation of the spectral distances (and the data preprocessing) and includes numerous data points of the overlapping frequency regions.

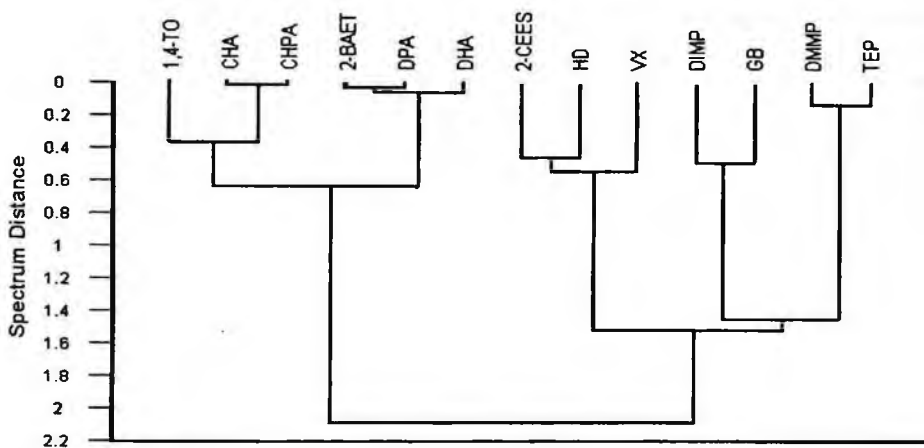


Figure 5. Cluster Analysis Dendrogram of CWA and CWAS and CWA degradation products studied.

The cluster analysis provides a graph which is called Dendrogram, where the y axis gives the distance that measures the similarity of the things tested. In this study, Infrared spectra were analyzed. The spectral distance indicates the degree of spectral similarity. Two spectra with a spectral distance of 0 are entirely identical (within the frequency ranges tested). The higher the distance between two spectra, the lower the similarity between spectra.

Figure 5 shows the high spectroscopic similarity that exists between GB and its simulant DIMP and also a very high similarity between TEP and DMMP. This fact can be explained from the Dendrogram distances observed. This indicates that the best simulant for the infrared detection for GB is DIMP. Simulant DIMP was found to be closely related to the TEP, another simulant and less related to GB, VX and HD. A high similarity was obtained between 2-CEES and HD. VX and HD are shown to be spectroscopically closely related to 2-CEES. This can be explained by the presence of the -S- group these compounds. Model compounds 2-BAET, DPA and DHA are highly similar to each other. This similarity arises from the fact that they are secondary amines. Although 2-BAET is typically used as a VX simulant, their spectroscopic similarity is not good. This is due to the fact that VX is a tertiary amine. CHA and CHPA have a high similarity for being cyclic amines and 1,4-TO is closely related to them because it is also a cyclic compound.

3.2.2. Surface detection

In this work the first method for detection and quantification of a chemical agent's simulants on metallic surfaces using sample smearing on surface as transfer method was evaluated as a proof of concept experiment. Spectroscopic characterization of thin layer deposits was achieved using the powerful technique of Grazing Angle Probe-Fiber Optic Coupled-FTIR developed for surface analysis¹²⁻¹⁶. This methodology relies on

Reflection-Absorption Infrared Spectroscopy (RAIS) operating at the grazing-angle which is one of the most powerful resonance absorption phenomenon and is the most sensitive optical absorption technique available for measuring low concentrations of chemical compounds adhered to surfaces and even allows for the non-destructive mono layer chemical analysis on the surface. Figure 6 shows the FOC-GAP-FTIR spectra 5 and 10 $\mu\text{g}/\text{cm}^2$ 2-CEES deposited as a trace residue on stainless steel surface. It is clear from the spectra shown that low limits of detection (LOD) achievable could be as low as nanograms/ cm^2 (milligrams/ m^2).

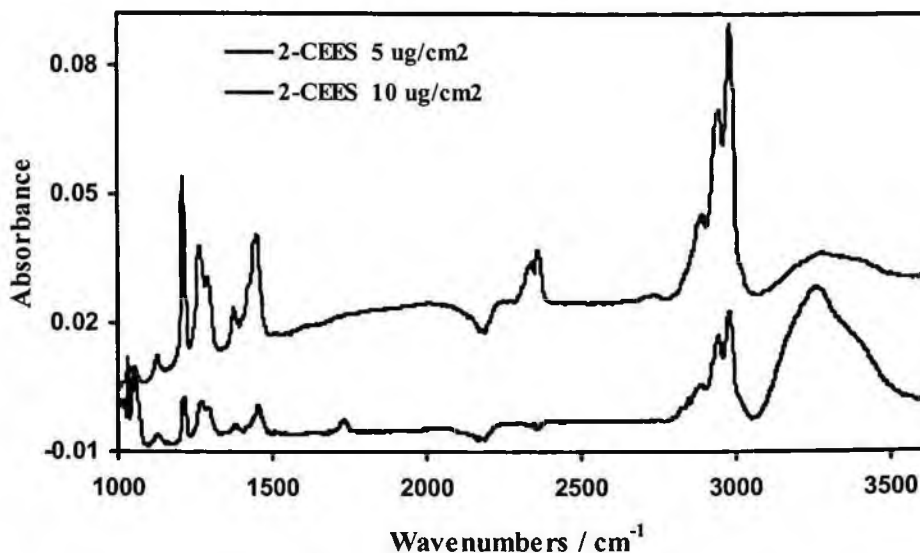


Figure 6. Fiber optic coupled-FTIR of 2-CEES on stainless steel surface.

Comparing the Fiber Coupled-GAP-FTIR spectra obtained with KBr cell spectra measured in a bench instrument, it can be concluded that they are identical. This methodology can be applied for detecting CWA on surfaces such as metals, plastics, glass, wood and others¹²⁻¹³. Other applications of the technique in the arena of chemical agents and TIC are in decontamination and cleaning validation of these highly toxic and polluting compounds. Fiber coupling allows working in the far field at a 2-5 m from the source, providing safe management for the operator. Figure 7 shows the FOC-GAP-FTIR spectra of TIC: acetone, chloroform and dimethylformamide. The only practical limitation of this spectroscopic analysis is the residence time of analytes on surfaces. For threat compounds with high vapor pressures (liquids) or high sublimation rates (solids) are subject to low surface residence time which will downgrade the excellent sensitivity and robustness of the methodology.

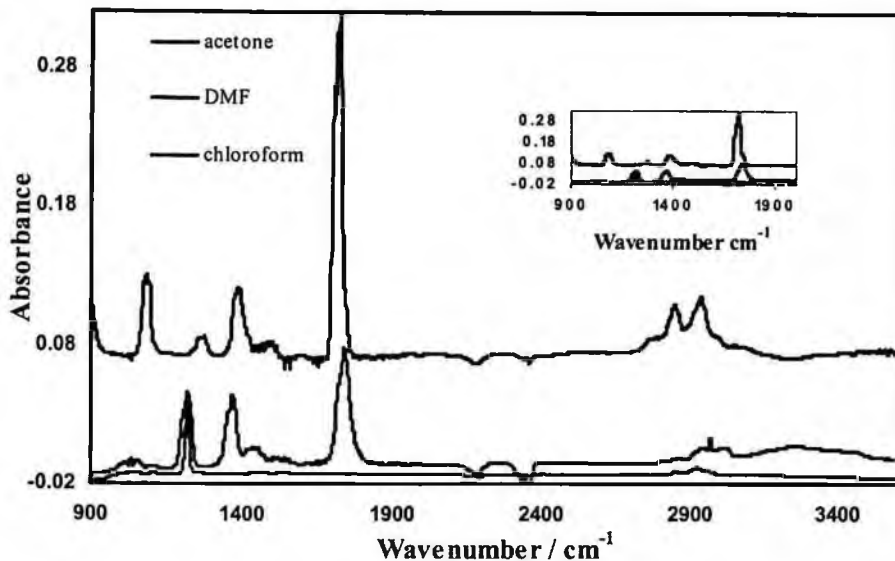


Figure 7. Fiber optic coupled-FTIR of several TIC on stainless steel surface.

4. Conclusions

The contributed work represents the continuation of a Chemical Point Detection Program at the Center for Chemical Sensors Development of the University of Puerto Rico – Mayagüez. The focal point of the program is detection of chemical treat agents such as explosives, chemical and biological agents and toxic industrial compounds for Defense and security applications. Vibrational signatures of simulants of Chemical Warfare Agents, degradation products of CWA and Toxic Industrial Compounds have been measured using Raman Microscopy, FTIR in the liquid and gas phases and Fiber Coupled-Grazing Angle Probe-FTIR in the case of liquids deposited on metal surfaces as trace residues. Raman excitation in the NIR (785 nm), in the VIS (532, 514 and 488 nm) and in the DUV (244 nm) provided clear and strong signatures that can be used for Chemical Detection in the near or far field. FTIR measurements in gas phase were correlated with liquid phase measurements and with *ab initio* computed IR vibrational spectra of simulants. Similar computations were performed on chemical agents that could not be measured experimentally. Cluster Analysis of IR vibrational signatures were performed on all threat agents measured and of calculated real chemical agents. The best simulant found for GB was DIMP. Simulant DIMP was found to be closely related to the TEP, another simulant, and less related to GB, VX and HD. A high similarity was obtained between 2-CEES and HD.

In this work the first method for detection and quantification of the chemical agent's simulants on metallic surfaces using sample smearing on the surface as the transfer method was evaluated as a proof of concept experiment. Comparison of the Fiber Coupled-GAP-FTIR spectra obtained with KBr cell spectra measured in a bench

instrument concluded that they are identical. This methodology can be applied for detecting CWA on surfaces such as metals, plastics, glass, wood and others. Other applications of the technique in the arena of chemical agents and TIC are in surface decontamination and cleaning validation of these highly toxic and polluting compounds

5. Acknowledgements

This work was supported by the U.S. Department of Defense, University Research Initiative (URI)-MURI Program, under grant number: DAAD19-02-1-0257. The authors also acknowledge contributions by the Chemical Imaging Center at the Chemistry Department of the University of Puerto Rico – Mayagüez.

References

1. Y. Sun and K. Y. Ong, *Detection Technologies for Chemical Warfare Agents and Toxic Vapors* (CRC Press, Boca Raton, FL, 2005).
2. S. Farquharson, A. Gift, P. Maksymiuk and F. Inscore, Surface-enhanced Raman spectra of VX and its hydrolysis products, *Appl. Spec.*, **59**, 654-659 (2005).
3. V. S. Smentkowski, P. Hagans and J. T. J. Yates, Study of the Catalytic Destruction of Dimethyl Methylphosphonate: Oxidation over Mo(II), *Phys. Chem.*, **92**, 6351-6357, (1988).
4. T. C. Marrs, R. L. Maynard and F. R. Sidell, *Chemical Warfare Agents: Toxicology and Treatment*, (John Wiley & Sons Ltd., London, UK, 1996).
5. U.S. Department of Defense, Department of the Army, NATO Handbook on the Medical Aspects of NBC Defensive Operations, Part III – Chemical (February 1996).
6. P. A. D'Agostino, J. R. Hancock, C. L. Chenier and C. R. Jackson-Lepage, Liquid chromatography electrospray tandem mass spectrometric and desorption electrospray ionization tandem mass spectrometric analysis of chemical warfare agents in office media typically collected during a forensic investigation, *J. Chrom. A*, **1110**, 86-94 (2006).
7. D. D. Pardasani, M. Palit, A.K. Gupta, Pankaj K. Kanaujia and D.K. Dubey. Gas chromatography–mass spectrometry analysis of trifluoroacetyl derivatives of precursors of nitrogen and sulfur mustards for verification of chemical weapons convention, *J. Chrom. A*, **1059**, 157-164, (2004).
8. P. A. Smith, D. Koch, G. L. Hook, R. P. Erickson, C. R. Jackson, L. Haley, D. M. Wyatt, G. Betsinger and B. A. Eckenrode, Detection of gas-phase chemical warfare agents, *Anal. Chem.*, **23**, 296-306 (2004).
9. W. E. Steiner, C. S. Harden, F. Hong, S. J. Klopsch, H. H. Hill, Jr. and V. M. McHugh. Detection of Aqueous Phase Chemical Warfare Agent Degradation Products by Negative Mode Ion Mobility Time-of-Flight Mass Spectrometry [IM(TOF)MS], *J. Am. Soc. Mass Spec.*, **17** 241-245 (2006).
10. A. B. Kanua, Paul E. Haigh and Herbert H. Hill. Surface detection of chemical warfare agent simulants and degradation products *Anal. Chim. Acta.*, **553**, 148–159 (2005).
11. W. F. Pearman and A. W. Fountain, Classification of Chemical and Biological Warfare Agent Simulants by Surface-Enhanced Raman Spectroscopy and Multivariate Statistical Techniques, *Appl. Spec.*, **60**, 356-36 (2001).
12. O. M. Primera-Pedrozo, L. C. Pacheco-Londoño, L. F. De la Torre-Quintana, S. P. Hernandez-Rivera, R T. Chamberlain and R. T. Lareau, Use of fiber optic coupled FT-IR in detection of explosives on surfaces, *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense III*; Edward M. Carapezza; Ed., *Proc. SPIE*, **5403**, 237-245 (2004).

13. O. M. Primera-Pedrozo, L. C. Pacheco-Londoño, O. Ruiz, M. Ramirez, Y. M. Soto-Feliciano, L. F. De La Torre-Quintana, S. P. Hernandez-Rivera, Characterization of thermal inkjet technology TNT deposits by fiber optic-grazing angle probe FTIR spectroscopy, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IV; Edward M. Carapezza; Ed., *Proc. SPIE*, **5778**, 543-552 (2005).
14. Melling, P.J.; Shelley, P. Spectroscopic Accessory for Examining Films and Coatings on Solid Surfaces, U.S Patent 6,3,10,348, United States Patent and Trademark Office, Washington, DC..
15. P. Melling and M. Thomson, Fiber optic probes for mid-infrared spectrometry, (John Wiley & Sons Ltd. 2002).
16. N. K. Mehta, J. E. Goenaga-Polo, S. P. Hernández-Rivera, D. Hernández, M. A. Thomson, P. J. Melling, Development of an In-Situ Spectroscopic Method for Cleaning Validation Using Mid-IR Fiber Optics, *Bio Pharm.* **15**(5), 36-42 (2002).
17. Gaussian 03, Revision B.04, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 2003.
A. D. Becke, Correlation energy of an inhomogeneous electron gas: A coordinate-space model, *J. Chem. Phys.*, **88**: 1053-1062, (1988).
18. A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, *J. Chem. Phys.*, **98**, 5648. (1993).
19. W. Koch, M. C. Holthausen, A Chemists guide to density functional theory, pp. 130-133, Second edition, (Wiley-VCH, Germany, 2001).
20. R. G Brereton, Chemometrics: Data Analysis for the Laboratory and Chemical Plant, pp. 224-230 (John Wiley & Sons, Ltd., London, UK, 2003).
21. K.R Beebe, R.J Pell and M.B Seasholtz. Chemometrics: A Practical Guide. John Wiley & Sons, INC New York (1998) 65-83

HYPERSPECTRAL AND POLARIZATION IMAGING WITH DOUBLE-TRANSDUCER AOTFS FOR WIDER SPECTRAL COVERAGE

NEELAM GUPTA

*U.S. Army Research Laboratory, 2800 Powder Mill Road,
Adelphi, Maryland 20783-1197, USA
ngupta@arl.army.mil*

We have developed a number of programmable hyperspectral imagers operating from the ultraviolet to the longwave infrared using an acousto-optic tunable filter (AOTF) in conjunction with a suitable camera. Each of these AOTFs is fabricated with a single transducer to operate over only one octave in wavelength. Recently, we have developed two separate tellurium dioxide (TeO_2) noncollinear AOTF cells that operate over two octave range. Each cell has an aperture of $1.5 \times 1.5 \text{ cm}^2$ with a field-of-view of 4.2° , and two thin-plate lithium niobate transducers are bonded on it. The tuning range of the first cell is $0.43\text{--}2.1 \mu\text{m}$, and $0.69\text{--}4.0 \mu\text{m}$ for the second cell. We have used each of these cells to carry out spectral imaging experiments with CCD, InGaAs, and InSb cameras to cover the spectral region from the visible to the midwave infrared. We image the two orthogonally polarized diffracted beams to get spectropolarimetric information from the scene, as well as a broadband image by using one of the undiffracted beams. In this paper, we discuss the two AOTF cells, the imaging experiments, and present the results obtained.

Keywords: Hyperspectral imaging; spectropolarimetric imaging; acousto-optic tunable filter; AOTF; TeO_2 ; double-transducers.

1. Introduction

The U.S. Army Research Laboratory has developed a family of random access wavelength agile hyperspectral imagers using electronically tunable acousto-optic tunable filters (AOTFs) as programmable dispersive elements in combination with a camera and optics appropriate for the spectral region of operation covering wavelengths from the ultraviolet (UV) to the longwave infrared (LWIR).¹⁻⁹ Such imagers can be used for a variety of applications including chemical/biological agent detection, target detection, medical diagnostics, atmospheric monitoring, and in astronomy to study galaxies, stars, and planets. In general, an AOTF-based hyperspectral imager operates over one octave in wavelength due to the operational limits of the piezoelectric transducer bonded on the crystal. Such AOTFs can be used with either two cameras to image the two diffracted beams to obtain both spectral and polarization information, as well as a third camera to image one of the two undiffracted beams to obtain a broadband image,⁴ or with only one

camera to image only one diffracted beam (by blocking the second diffracted beam and the two undiffracted beams) and placing an electronically tunable liquid crystal variable retarder (LCVR) before the AOTF in the optical train to obtain spectropolarimetric images.¹

Recently, we have designed hyperspectral imagers with more than two octave coverage in wavelength.⁴ To achieve such an operation, a novel noncollinear AOTF cell using a TeO₂ crystal had to be designed with two transducers bonded on the crystal.⁹ Two such AOTF cells were fabricated. Each AOTF cell was fabricated in a single crystal of TeO₂ with a square aperture of 1.5×1.5 cm² and an angular aperture of 4.2°. Each cell has two thin-plate lithium niobate transducers bonded on it. Each transducer has a separate radio-frequency (rf) connector, and the frequency and power of the applied rf signal to each transducer are independently controlled. Each of the transducer is designed with carefully matched electrical and acoustic impedances such that almost all applied electric power is converted to acoustic power. The first cell covers 0.43–2.1 μm, corresponding to an rf frequency range from 123 to 18 MHz and the second cell has a tuning range 0.69–4.0 μm with a corresponding rf frequency range from 61 to 10 MHz. We used these cells in a number of different imaging experiments with CCD, InGaAs, and InSb cameras to cover the spectral region from the visible to the midwave IR. The tuning of the applied frequency, and the image acquisition and storage was carried out under full computer control. We obtained excellent spectropolarimetric images during these imaging experiments. In this paper, we will first give a short background and operational principles for an AOTF, then describe the imaging experiments carried out in our laboratory and present the results obtained.

2. Background

An AOTF is fabricated by bonding a piezoelectric transducer to a specially cut birefringent crystal.^{10–13} When an rf signal is applied to the transducer, it produces an ultrasonic wave that travels through the crystal with an acoustic frequency matching the applied rf. This sets up a moving diffraction grating in the crystal with the acoustic velocity specific to the material, and the grating period equal to the acoustic wavelength in the material. An acoustic absorber absorbs the sound wave after it traverses the crystal. When polarized light is incident on such a crystal, it is diffracted by the traveling acoustic wave, and produces an orthogonally polarized diffracted beam with a Doppler shift for a particular wavelength based on the phase-matching condition. The diffracted wavelength is inversely proportional to the acoustic frequency. There are two main types of AOTFs: collinear and noncollinear. In a collinear AOTF, the incident, diffracted, and undiffracted optical beams, as well as the sound wave travel in the same direction while in a noncollinear design the diffracted beams and the sound wave travel in different directions relative to the incident and the undiffracted optical beams. Most of the filters used in imaging applications are noncollinear and filter unpolarized broadband light into three spatially separated beams: a frequency upshifted beam, a frequency downshifted beam, and undiffracted incident light at all wavelengths except the diffracted wavelength that

contains two separate beams with orthogonal polarizations corresponding to the ordinary and the extraordinary beam propagation in the birefringent crystal. The two frequency shifted beams have orthogonal polarizations, and they are spatially separated from each other on either side of the incident beam direction. In general, an AOTF has a single transducer bonded to it and due to transducer limitations; it operates over a frequency range of one octave or less. If more than one transducer is bonded to the crystal, the operating frequency range can be expanded to more than one octave. With a careful design of a transducer having matching acoustic and electrical impedances, an imaging AOTF cell can be fabricated with two transducers to operate over greater than two-octave range.¹⁴ In general, AOTF cells for hyperspectral imaging applications are designed with a large linear, as well as angular aperture and a relatively broad bandpass to be able to collect enough photons in each narrow spectral band to obtain high quality spectral images with good signal-to-noise-ratio.

We show the spectral filtering operation of a noncollinear AOTF cell with two transducers in figure 1. As shown in this drawing, the first transducer works for the higher rf frequency input and the second one for the lower rf frequency. Both diffracted beams cover a narrow spectral band defined by the applied rf and are polarized orthogonal to each other. Since the zero order diffracted light consists of two orthogonally polarized beams as explained earlier, a linear polarizer can be used to pass through only one of these beams to obtain a broadband polarized image of the scene on a CCD camera.

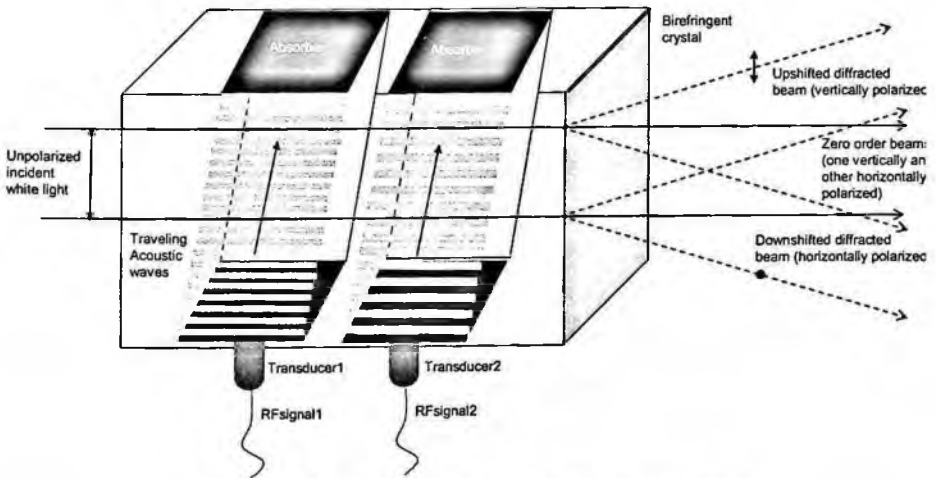


Figure 1. Filtering operation of an AOTF cell with two transducers.

3. Theory of operation

The tuning relationship for a noncollinear filter can be derived based on the conservation of momentum in a photon-phonon interaction given by the following expression

$$\mathbf{k}_{io} - \mathbf{K} = \mathbf{k}_{de} \quad (1)$$

where the lengths of the wavevectors of the ordinary polarized incident and the extraordinary polarized diffracted light are given by $k_{io} = \frac{2\pi n_o}{\lambda_0}$ and $k_{de} = \frac{2\pi n_d}{\lambda_0}$, as shown

in the wavevector diagram for the wide-angle acousto-optic (AO) interaction geometry in TeO₂ in figure 2. The length of the ultrasound wavevector with frequency f is given by $K = \frac{2\pi f}{V}$. Here, n_o is the refractive index for the ordinary beam and n_d is the refractive

index for the extraordinary beam as shown in figure 2. λ_0 is the center wavelength of the diffracted light, and V is the acoustic velocity in the material. n_d is equal to n_e on the [110] axis, n_o on the [001] axis, and at other points it is given by the equation for the index ellipse as follows

$$n_d(\theta_d) = \frac{n_e n_o}{\sqrt{n_e^2 \cos^2(\theta_d + \alpha) + n_o^2 \sin^2(\theta_d + \alpha)}} \quad (2)$$

where α is the angle the acoustic wavevector makes with the [110] axis and the diffracted wavevector is at an angle θ_d relative to a direction normal to \mathbf{K} .

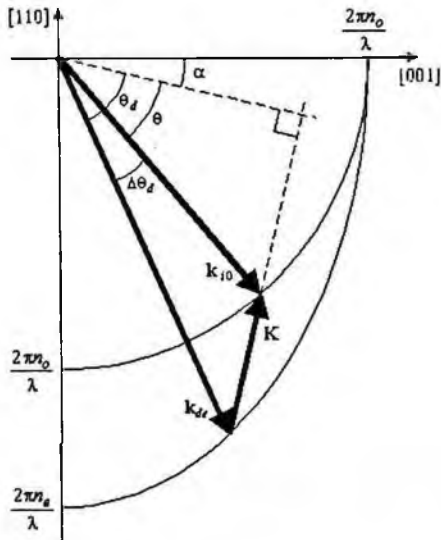


Figure 2. The vector diagram for noncollinear AO interaction in TeO₂.

The acoustic phase velocity in the direction defined by the tilt angle α in the crystal is given by

$$V(\alpha) = \sqrt{V_{110}^2 \cos^2 \alpha + V_{001}^2 \sin^2 \alpha} , \quad (3)$$

where V_{110} and V_{001} are the acoustic velocities in the [110] and [001] directions. The spectral tuning relationship for the filter is given by the following expression

$$\lambda_0 = \frac{\Delta n V \sin^2(\theta_d + \alpha)}{f \sin \theta_d} , \quad (4)$$

where Δn is the birefringence (difference of the two refractive indices). The angle θ_d is related to the Bragg angle of incidence θ as follows

$$\theta_d = \arccos\left[\frac{n_o \cos \theta}{n_d}\right] , \quad (5)$$

and the expression for n_d can be further simplified to give

$$n_d = n_o \left[1 + \frac{\Delta n}{n_o} \sin^2(\theta_d + \alpha)\right] . \quad (6)$$

Since the angle θ is known, the value of θ_d can be obtained by solving Eqs. (4) and (5). The angular field of view (the angular separation between the incident and the diffracted beams), $\Delta\theta = \theta - \theta_d$, in the crystal for small values of $\Delta\theta$ is given by the following expression

$$\Delta\theta = (\Delta n / n_o) \sin^2(\theta_d + \alpha) \cot \theta_d . \quad (7)$$

The optical passband $\Delta\lambda$ with an acoustic walk-off angle ψ is given by

$$\Delta\lambda \leq \frac{0.8\lambda_0^2 \cos \theta}{\Delta n L \sin^2(\theta_d + \alpha)} (1 + \tan \psi \tan \theta) , \quad (8)$$

where

$$\psi = \arctan[(V_{001}/V_{110})^2 \tan \alpha] - \alpha , \quad (9)$$

and L is the length of the transducer. The AO figure-of-merit M_2 is given as follows

$$M_2 = \frac{n_o^3 n_d^3 p^2}{\rho V^3} , \quad (10)$$

and the diffraction efficiency η of the filter is expressed as

$$\eta \approx \frac{\pi^2}{2\lambda_0^2} M_2 \left(\frac{P_a L}{H}\right) , \quad (11)$$

where p is the effective photoelastic coefficient, ρ is the density of the crystal material, P_a is the applied rf power, and H is the height of the transducer plate.

4. Filter parameters

The two AOTFs were fabricated with $\alpha = 6^\circ$, $\theta = 6.2^\circ$, a field-of-view of 4.2° in air, and with an aperture of $1.5 \times 1.5 \text{ cm}^2$. They were characterized for wide spectral imaging coverage in the combined spectral range from the visible to the midwave IR. The tuning range of the first cell is $0.43\text{--}2.1 \text{ }\mu\text{m}$, corresponding to an rf frequency range from 123 to 18 MHz, with transducer 1 operating over 45–123 MHz and transducer 2 operating over 18–46 MHz. The second cell can be tuned over $0.69\text{--}4.0 \text{ }\mu\text{m}$ with corresponding rf frequency range from 61 to 10 MHz, with transducer 1 operating from 24 to 61 MHz and transducer 2 operating from 10 to 25 MHz. The bandpass for each transducer is different for both these cells—for the first cell with transducer 1 it is 225 cm^{-1} , and with transducer 2 it is 70 cm^{-1} , while for the second cell, it is 125 cm^{-1} for transducer 1, and 68 cm^{-1} for transducer 2. The first cell operates with close to 96% efficiency for applied rf power of less than 1 W to either of the two transducers. The second cell has similarly high efficiency with transducer 1 in the shortwave IR region for 1.5 W of applied rf power but only half as much efficiency with 3.6 W rf power in the longer wavelength region because the applied rf power increases as the square of wavelength, because the AO figure-of-merit is smaller in this region, and because the transducer length used was shorter than what was required in the design due to the size of the available TeO_2 crystal. The operational parameters of the two AOTFs are listed in table 1. A schematic drawing of the cell design is shown in figure 3 and a photograph of the filter is shown in figure 4. Tuning curves showing the applied rf frequency dependence of the diffracted wavelength for both the filters were obtained by using both theoretical and measurement results, and are shown in figures 5 (a) and (b).

Table 1. AOTF parameters.

Parameters	AOTF1		AOTF2	
Material	TeO ₂		TeO ₂	
Linear aperture (cm ²)	1.5 × 1.5		1.5 × 1.5	
Angular aperture (°)	4.2		4.2	
	Transducer 1	Transducer 2	Transducer 1	Transducer 2
Spectral range (μm)	0.43–2.1		0.69–4.0	
	0.43–0.93	0.88–2.1	0.69–1.67	1.62–4.0
RF (MHz)	18 – 123		10–61	
	43–123	18–46	24–61	10–25
Spectral resolution (cm ⁻¹)	225	70	125	68
RF drive power (W)	0.7@0.63 μm	0.95@1.15 μm	1.5@1.15 μm	3.6@3.39 μm
Diffraction efficiency (%)	95@0.63 μm	96@1.15 μm	96@1.15 μm	50@3.39 μm

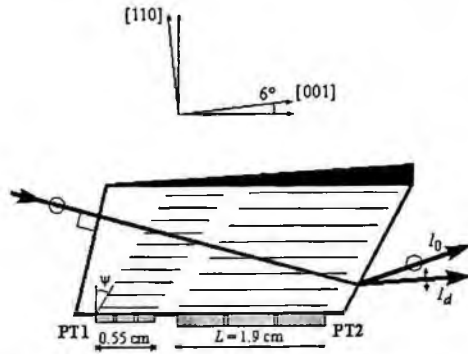


Figure 3. Schematic drawing for the design of one of the cells.



Figure 4. A two transducer AOTF.

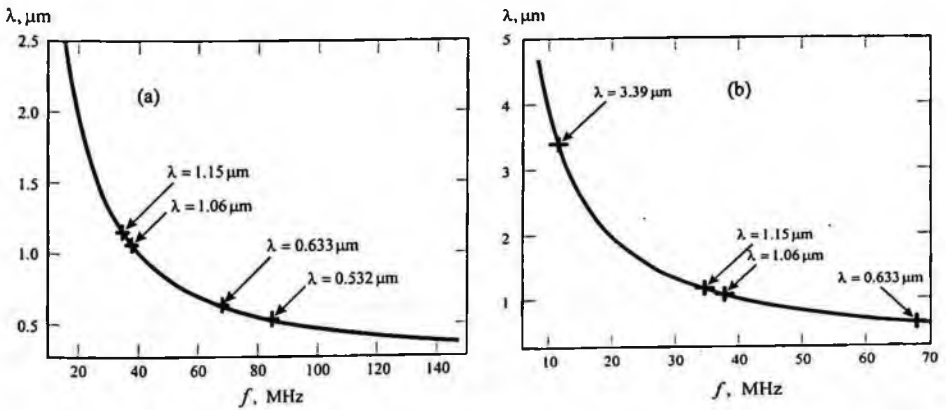


Figure 5. Tuning curves for the two AOTFs—(a) the first cell and (b) the second cell.

5. Imaging experiments and results

Three different imaging experiments were carried out using these filters. In one of the imaging experiments, we used three separate CCD cameras with AOTF1 operating from the visible to the shortwave IR—the middle camera images one of the undiffracted zero-order beams with a linear polarizer in the incidence direction, and the cameras on either side of it image the upshifted and the downshifted diffracted beams propagating at an angle relative to the incident optical beam. These two diffracted beams have orthogonal polarizations. This way we obtained both the spectral and the polarization information from the scene using the two side cameras and a broadband image of the scene by the middle camera. The wavelength of the diffracted beams was changed by changing the rf frequency applied to transducer 1. A schematic drawing of this experiment is shown in figure 6. The imaged object was a U.S. Air Force (AF) resolution chart on a glass plate. Four of the polarized spectral images at 500, 540, 580 and 620 nm obtained in this experiment are included in figure 7. In the second experiment we used both transducers of AOTF1 with three cameras. One InGaAs camera was used to image one of the diffracted beams corresponding to transducer 2, while the other diffracted beam due to transducer 1, and one of the undiffracted beams, was imaged using two separate CCD cameras. Again the AF resolution chart on glass was imaged, and a couple of polarized spectral images obtained in this experiment from the InGaAs camera are shown in figure 8. In the third experiment, we used an InSb camera with a Ge window cooled to 77K with AOTF2 operating from the near IR to midwave IR, and applied rf power to transducer 2 to image both the diffracted beams and subtracted out the undiffracted beams. The object imaged was a different resolution chart on a ZeSe plate. One of the polarized spectral images obtained in this experiment is shown in figure 9. The magnification in these three different sets of images is different due to use of different camera lenses.

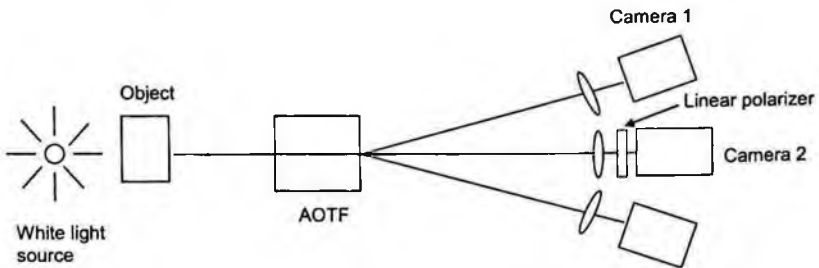


Figure 6. Schematic drawing of the spectral and polarization imaging experiment using one of the two-transducer AOTFs with three cameras to image the two diffracted and one of the undiffracted beams. The two cameras on the side, image orthogonally polarized spectral images from the two diffracted beams and the middle camera with a linear polarization filter in front, acquires a broadband image from one of the zero order beams.

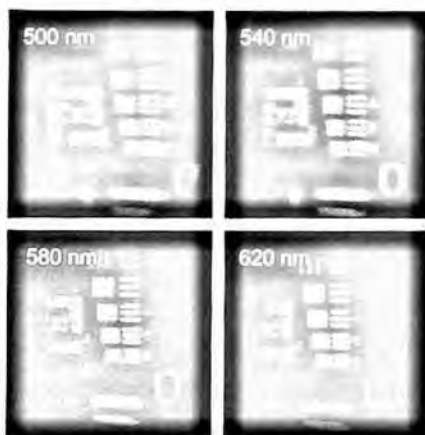


Figure 7. Polarized spectral images of a resolution chart using transducer 1 on AOTF 1.

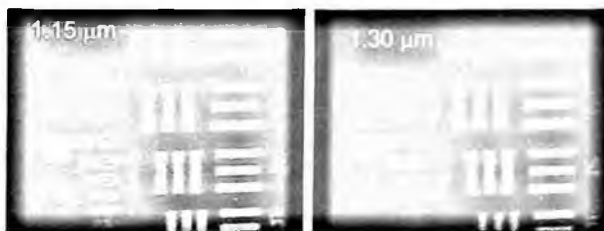


Figure 8. Polarized spectral images of a resolution chart using transducer 2 on AOTF 1.



Figure 9. Polarized spectral image of a resolution chart using transducer 2 on AOTF 2.

6. Discussion and conclusion

AOTFs designed for spectral imaging applications need to have large linear and angular apertures as well as a relatively wider bandpass to provide good signal-to-noise-ratio in remote sensing applications. Many spectral imaging applications require the use of AOTFs that cover spectral regions wider than one octave. In general, an AOTF cell with a single transducer covers only one octave or less in wavelength. We developed two double-transducer noncollinear AOTF cells, each with a $1.5 \times 1.5 \text{ cm}^2$ aperture and a 4.2° field-of-view, to operate over a more than two-octave spectral region. The first cell operates over $0.43\text{--}2.1 \text{ }\mu\text{m}$, corresponding to an rf frequency range from 123 to 18 MHz, 43–123 MHz with transducer 1 and 18–46 MHz with transducer 2. The second cell covers $0.69\text{--}4.0 \text{ }\mu\text{m}$, corresponding to an rf frequency range from 61 to 10 MHz, 24 to 61 MHz with transducer 1 and 10 to 25 MHz with transducer 2. The bandpass of the first cell is 225 cm^{-1} with transducer 1 and 70 cm^{-1} with transducer 2, and for the second cell, it is 125 cm^{-1} for transducer 1 and 68 cm^{-1} for transducer 2. Both cells operate with 96% efficiency from visible to shortwave IR for an applied rf power of less than 1.5 W, while the efficiency in the midwave IR is only 50% with 3.6 W due to the fact that applied power increases as the square of the wavelength. The unique feature of the AOTFs fabricated for this research is that both the electrical and acoustic impedances of the transducers were matched well such that most of the applied rf power coupled well to generate acoustic field inside the crystal.

We carried out three separate imaging experiments using CCD, InGaAs and cooled InSb cameras to collect polarized spectral images over the full range of coverage of the two cells. We obtained high quality images. The zero-order broadband image can be obtained with high spatial resolution. We had to use three separate cameras since each available camera has a limited spectral coverage. We could use a single cooled InSb camera with either a MgF_2 or CaF_2 window to cover the entire spectral range of operation.

References

1. N. Gupta, R. Dahmani, and S. Choy, "Acousto-optic tunable filter based visible-to-near-infrared spectropolarimetric imager," *Opt. Eng.* **41**, p. 1033–1038, 2002.
2. N. Gupta and V. Voloshinov, "Hyperspectral Imager from Ultraviolet to Visible Using KDP AOTF," *Appl. Opt.* **43**, p. 2752–2759 (2004).
3. N. Gupta, R. Dahmani, and K. Bennett, S. Simizu, D. R. Suhre, and N. B. Singh, "Progress in AOTF Hyperspectral Imagers," in *Automated Geo-Spatial Image and Data Exploitation*, W. E. Roper and M. K. Hamilton, eds, Proc. SPIE **4054**, p. 30–38 (2000).
4. N. Gupta, "Development of Agile Wide Spectral Range Hyperspectral/Polarization Imagers," Technical Digest CLEO/QELS/PhAST, PThA3 (2005).
5. N. Gupta, D. R. Suhre, and M. Gottlieb, "LWIR spectral imager with an 8-cm^{-1} passband acousto-optic tunable filter," *Opt. Eng.* **44**, p. 094601 1–7, (2005).
6. N. Gupta, "Acousto-optic tunable filters for Infrared Imaging," in *Acousto-optics and Photoacoustics*, R. Reibold, ed., Proc SPIE **5953**, 59530O 1–10 (2005).
7. N. Gupta and V. B. Voloshinov, "Hyperspectral imaging performance of a TeO_2 imaging acousto-optic tunable filter in the ultraviolet region," *Opt. Lett.* **30**, p. 985–987 (2005).

8. V. B. Voloshinov and N. Gupta, "Investigation of magnesium fluoride crystals for imaging acousto-optic tunable filter applications," *Appl. Opt.* **45**, p.3127–3135 (2006).
9. N. Gupta and V. B. Voloshinov, "Development and Characterization of Two-Transducer Imaging Acousto-Optic Tunable Filters with Extended Tuning Range," submitted to *Appl. Opt.* (July 2006).
10. Goutzoulis A. and Pape D. *Designing and Fabrication of Acousto-Optic Devices* (New York: Marcel Dekker, 1994).
11. J. Xu and R. Stroud, *Acousto-Optic Devices* (Wiley, New York, 1992).
12. N. Gupta, "Acousto-Optics," in *Optical Engineer's Desk Reference*, ed. W. Wolfe, (Optical Society of America, Washington D.C., 2003).
13. C. Chang, "Tunable acousto-optic filters: an overview," in *Acousto-Optics: Device Development/Instrumentation/Applications*, J. B. Houston, ed., Proc. SPIE **90**, p. 12–22 (1976).
14. B. Belikov, V. B. Voloshinov, A. B. Kasyanov, et. al., "Broadband matching of an Acousto-optical-cell transducer using Youla's complex normalization theory," *Radioelectronics and Communication Systems* **31**, p. 28–33 (1988).

DEEP-UV BASED ACOUSTO-OPTIC TUNABLE FILTER FOR SPECTRAL SENSING APPLICATIONS

NARASIMHA S. PRASAD

*Nasa Langley Research Center, 5 N. Dryden St., MS 468, Hampton, VA 23681
n.s.prasad@larc.nasa.gov*

In this paper, recent progress made in the development of quartz and KDP crystal based acousto-optic tunable filters (AOTF) are presented. These AOTFs are developed for operation over deep-UV to near-UV wavelengths of 190 nm to 400 nm. Preliminary output performance measurements of quartz AOTF and design specifications of KDP AOTF are presented. At 355 nm, the quartz AOTF device offered ~15% diffraction efficiency with a passband full-width-half-maximum (FWHM) of less than 0.0625 nm. Further characterization of quartz AOTF devices at deep-UV wavelengths is progressing. The hermetic packaging of KDP AOTF is nearing completion. The solid-state optical sources being used for excitation include nonlinear optics based high-energy tunable UV transmitters that operate around 320 nm and 308 nm wavelengths, and a tunable deep-UV laser operating over 193 nm to 210 nm. These AOTF devices have been developed as turn-key devices for primarily for space-based chemical and biological sensing applications using laser induced Fluorescence and resonance Raman techniques.

Keywords: Deep-UV wavelengths; AOTF; chem-bio sensing

1. Introduction

NASA began Laser Risk Reduction Program (LRRP) in 2002 to develop reliable, robust, and compact laser technologies for lidar applications from space based platforms. The goal is to advance 1 micron and 2 micron lasers and associated wavelength conversion technologies for reliable operation in space environments. LRRP is a joint operation of Langley Research Center and Goddard Space Flight Center. The LRRP effort focuses on four Lidar techniques namely altimetry, Doppler, Differential Absorption Lidar (DIAL), and backscatter lidar. The emphasis is on six priority Earth Science measurements: (i) surface and ice mapping, (ii) horizontal vector wind profiling, (iii) carbon dioxide (CO₂) profiling, (iv) ozone (O₃) profiling, (v) aerosol/clouds, and (vi) river current monitoring.

To explore ozone profiling using DIAL technique from space based platforms, an highly efficient, all-solid-state, diode pumped, conductively cooled, single longitudinal mode and high-energy 1-micron to UV wavelength conversion technology to primarily generate UV wavelengths of 308 nm and 320 nm has been developed. This effort is a collaborative effort among Sandia National Labs, Fibertek, Inc. and NASA Langley

Research Center¹. This tunable UV transmitter technology has chem-bio sensing applications for NASA, defense, and Homeland Security Agency (HSA) missions. In the case of NASA, this UV converter technology provides a path for the development of compact, electrically controllable, and rugged chem-bio sensors specifically designed for investigating geological and mineralogical compositions as well as exploring precursors of life signature on various bodies of our Solar system. Besides NASA applications, this technology is suited for several HSA applications. For e.g., Tryptophan (amino acid generated by protein digestion), absorbs at 280 nm and emits over a range of 300-400 nm wavelengths with a resonance peak at 330 nm.

Besides ozone related wavelengths, the LRRP UV transmitter technology is capable of generating tunable wavelengths ranging from <200 nm to 400 nm spanning over deep-UV to near-UV spectral region using nonlinear optical techniques such as three-wave mixing. Among these wavelengths, deep-UV wavelength band is considered to play an important role in characterizing chem-bio species for space exploration missions. The tunable deep-UV wavelengths can serve as excitation wavelengths for Laser Induced Fluorescence (LIF) and Raman schemes including resonant Raman, and surface enhanced Raman effects. On the basis of this technology, a tunable deep-UV based chem-bio sensor technology has been being conceived primarily for space exploration purposes. Both, imaging and direct detection schemes corresponding to Laser Induced Fluorescence (LIF) and Raman techniques are being explored.

Various elements required for developing a robust, high resolution and electronically controllable chem-bio sensor are being investigated. One of the critical elements of this proposed tunable chem-bio sensor scheme is a narrow band tunable filter required to analyze spectral signatures emanating from targets of interest. For our initial studies, the acousto-optic tunable filter (AOTF) for analyzing Raman emissions due to deep-UV excitations has been selected. Besides effectively blocking background fluorescence, use of a deep-UV AOTF will allow electronic tuning of the peak passband to a desired emission wavelength. In this paper, the development of AOTFs appropriate for deep-UV wavelength operation is discussed. In Section 2, a brief description of excitation at deep-UV wavelengths is discussed. In Section 3, a brief overview on the ongoing development of UV transmitters as excitation sources is presented. In Section 4, the proposed chem-bio scheme is illustrated. In Section 5, design and fabrication of quartz crystal based AOTF is presented. In Section 6, initial performance measurements of the quartz AOTF are discussed. Summary and Conclusions are given in Section 7.

2. Excitation at Deep-UV wavelengths

The excitation in the deep-UV results in more intense Raman scattering. Deep-UV laser excitation has the ability to avoid fluorescence background in the Raman spectra. Tunability of the UV source could allow exploitation of resonance Raman effect. Resonance Raman Effect enhances the intensity of Raman lines. Deep-UV LIF is

exceptionally sensitive to the presence of large biological molecules, bacteria, and organics. Recent work using UV LIF technique has shown that under suitable conditions, it is possible to detect as few as 5 to 10 bacteria on a surface, and zepta molar $\{10^{-21}$ mole $\}$ of poly-aromatic hydrocarbons (PAHs). The specificity of Raman spectral molecular finger-prints has made the technique of great value in a number of areas including biochemistry and biophysics, early detection of cancers, chem-bio agents and for identifying minerals on planetary surfaces. LIF and laser Raman spectroscopy are considered sensitive techniques for providing reliable information on characteristics of organic molecules and the molecules of life. Use of excitation wavelengths below 250 nm, endogenous and Raman emission of target molecules can be measured without mutual interference.

Abundant polycyclic aromatic hydrocarbons discovered in the fresh fracture surfaces of the Martian meteorite have stimulated new interest in the search for past life on Mars. These polyaromatic chains present in a Martian meteorite indicating probable past life signature. Deep-UV excitation of precursors such as kerogens and lignins could provide insight into biological precursors². These are insoluble macromolecular hydrocarbons with low hydrogen content containing aromatic and polyaromatic components and are considered precursors for origins of life. Efforts are underway to devise sensitive techniques for the detection of lignins and kerogens.

3. The UV Laser Transmitter

The performance specifications of the single mode, conductively cooled, all-solid-state UV transmitter built under LRRP effort for ozone sensing applications include (i) nominal wavelength of 320 nm and 308 nm, pulse energy of greater than 200 mJ, pulsewidth around 25 ns, and repetition frequency of 50 Hz. For ozone sensing, the need for high pulse energy is to obtain enhanced performance during strong day light conditions. The basic scheme of this UV transmitter consists of a Nd:YAG laser operating at 1064 nm that pumps a nonlinear optics arrangement consisting of an Optical Parametric Oscillator (OPO) and a Sum Frequency Generator (SFG). The pump laser consists of a Nd:YAG gain medium based ring oscillator and four zig-zag Brewster angle slab amplifiers. Greater than 1J/pulse at 1064 nm wavelength is obtained by amplifying the oscillator pulse energy of ~ 30 J with four zig-zag slab amplifiers. Using the Second Harmonic Generation (SHG), the 532 nm wavelength is generated from 1064 nm wavelength. With KTP as a SHG crystal, up to 80% conversion efficiency has been demonstrated. The general scheme to generate UV wavelengths involves splitting the 532 nm wavelength energy almost equally using a beam splitter. One portion pumps a seeded KTP crystal based OPO to generate near-IR signal wavelength of 803 nm. The other portion of 532 nm wavelength energy is mixed with 803 nm using a SFG crystal to generate 320 nm wavelength. In the case of 308 nm, the OPO is designed to generate 731.5 nm.

The OPO used for 532 nm pumped IR generation is known as Rotated Image Singly Resonant Twisted rectangle (RISTRA) type. The RISTRA OPO is a Four-mirror image-rotating non-planar ring resonator in a compact arrangement. The advantages of RISTRA OPO include no alignment, long term stability, and good beam quality. The wall plug efficiency of the Nd:YAG pump laser is equal to 7% and the optical slope efficiency is equal to 11%. The UV converter arrangement with 532 nm pump provides near-IR to UV wavelengths conversion efficiency of up to 24% conversion efficiency with extra cavity sum frequency generator arrangement. The semiconductor diode-pumped Nd:YAG pump laser with zigzag slab amplifiers allow robust and efficient design for use in space. The conduction cooled operation eliminates circulating liquids within laser cavity. Finally, space-qualifiable component designs establish a path to space missions. The details of this setup are given in reference 1.

The LRRP UV transmitter is a dual use technology that is suitable for many chem-bio applications. Besides ozone sensing, our objective is to develop an all solid-state deep (200-300 nm) UV laser transmitter and matched receiver unit that will be suitable for LIF and remote Raman techniques. The current tunable UV transmitter technology is anticipated to allow the generation of UV wavelengths ranging from <200 nm to 400 nm with pulse energies of greater than 50 mJ. The availability of large pump energy >1J/pulse at 1064 nm allows generation of sufficient energy at deep UV wavelengths for in-situ laser induced fluorescence experiments. The wavelength tunability of the UV source allows exploitation of resonance Raman effect that could enhance the intensity of Raman lines by 3 to 6 order of magnitude of aromatic amino acids, nucleic acids, and a variety of bacterial pigments. One of the methodologies to obtain wide range of UV wavelengths is by angle tuning one or more OPO and SFG crystals with appropriate seed wavelengths. Other efficient design possibilities are being explored. For our current experimental purposes, a deep-UV laser providing wavelengths from 193 nm to 210 nm has been procured and characterized from Coherent, Inc. The specifications of this deep-UV laser include PRF of 5 kHz, pulsewidth of ~20 ns and spatial mode of TEM₀₀. Deep UV wavelengths are obtained by fourth harmonic generation technique from a Nd:YLF laser pumped Ti:Sapphire laser arrangement. The modified transmitter coupled with a tunable filter is anticipated to be suitable for the following remote sensing techniques.

4. The Chem-Bio Sensing Scheme

The experimental arrangement for analyzing responses from chem-bio species due to deep-UV excitation under development is illustrated in Figure 1. In this scheme, a tunable deep UV laser transmitter is used for excitation. An acousto-optic tunable filter (AOTF) is utilized for UV LIF and resonance Raman Spectroscopic techniques. The fluorescence and Raman signals emanated from target species is collected using a telescope and selectively routed into an imaging spectrograph or an AOTF. The spectrally selective filtering is succeeded by detection and data processing tools that are commercially available. The proposed experimental scheme will be utilized for

investigating mineralogical samples and/or simulants pertaining to extraterrestrial environments and for nondestructive detection of biologically important organisms and molecules. A secondary objective is to examine potential experimental schemes for planetary sample acquisition and habitability assessment as well as for human protection.

The experimental set-up has the flexibility to be adapted for both in-situ (dotted lines) and short-to-medium range (up to few km) remote sensing. Furthermore, the proposed scheme allows for direct detection as well as imaging. Direct detection followed by imaging experiments based on LIF and resonance Raman provides species distribution features. In this setup, the digital CCD camera system is PIXIS 400B from Princeton Instruments and the UV detector is from Hamamatsu Corporation. A custom designed Cassegrain telescope for operation over deep-UV to visible region has been fabricated. For looking at visible returns, a commercially available AOTF operating in the visible region such as that made of TeO_2 will be utilized. Characterization of critical components of this arrangement is progressing. In the following section, the characterization of AOTF devices is discussed.

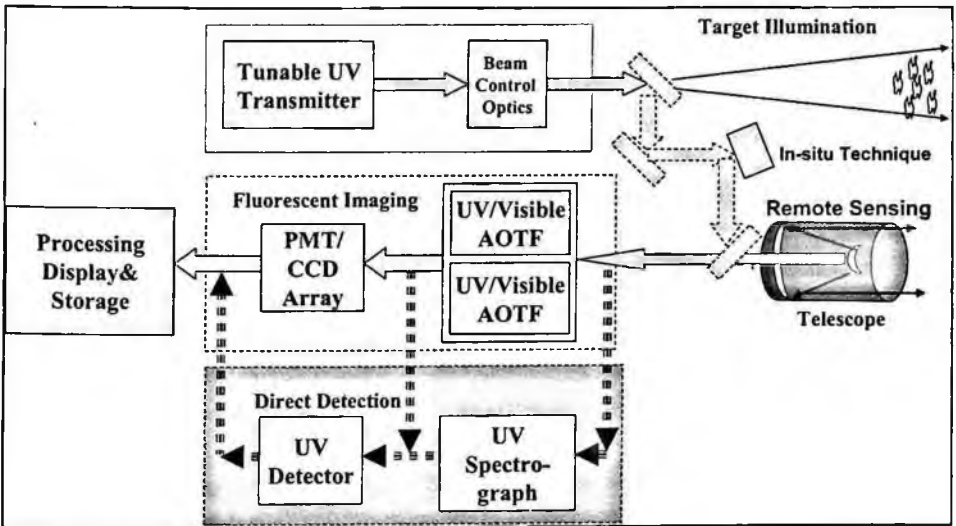


Fig. 1. Experimental setup for conducting chem.-bio sensing using single pixel detection as well as imaging schemes.

5. The Deep-UV AOTF device: Fabrication and Testing

Acousto-optic (AO) effect is the diffraction of light by sound generated phase grating via photo elastic effect (change of refractive index with stress). AO effect provides a convenient way of controlling the frequency, intensity and direction of an optical radiation³⁻⁷. The basic principle of an AOTF device is illustrated in Figure 2. The dependence of optical passband on RF drive frequency gives rise to spectral filtering. In

the case of an AOTF, an electronic tunable phase grating is established in an acousto-optic crystal due to the propagating ultrasonic wave⁸⁻¹⁰. Tuning is obtained by varying the RF drive frequency. AOTF devices are compact, solid-state devices with no moving parts. They provide good resolution (< 0.1 nm) wavelength agility and rapid tunability via electronic means. Another aspect is programmability and random wavelength selection that allows remote operation. Other attractive features include low repeatability errors ($< \pm 0.05$ nm), high signal to noise ratio, reliability, and repeatability. AOTF devices allow spectral imaging over a wide spectral band. As such, AOTFs are becoming attractive for a wide range of applications including sensing and imaging applications in place of conventional gratings. If power consumption of few watts of RF power required to operate these devices is not an issue in a given operational environment, then AOTF devices could play a significant role. This is particularly true in the case of space missions.

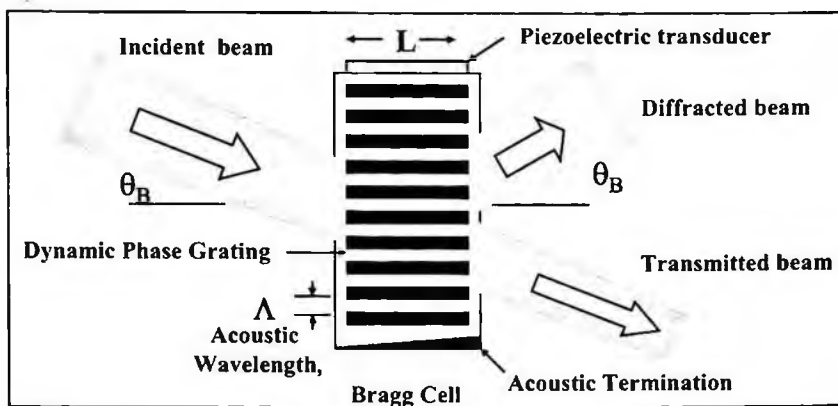


Fig. 2. The basic AOTF operational principle.

An AO device can be operated either in Bragg or Raman-Nath regime. Bragg condition generates only one diffraction order. As such, Bragg cells are widely used in various applications. AOTFs can be constructed using collinear or non-collinear type of interactions in an anisotropic medium. Non collinear interactions allow operational flexibility as well as the diffracted order is well separated from the zero-order beams. Collinear interactions provide large apertures. However, one of the main disadvantages of collinear case is that the diffracted and zero-order beams are very close to each other and hence need a polarizing beam splitter to separate them. AOTFs generally use shear-mode interactions to optimize diffraction efficiency and passband characteristics.

AOTFs operating in the visible and Near-IR spectral regions have been well established. Practical deep-UV, mid-wave IR (3-5 microns) and long wave IR (8-12 microns) AOTFs are under development. Materials growth concerns are hampering their progress. However, steady progress is made in the case of long-wave IR materials for

AOTF applications¹¹. Recently, a near-UV AOTF operation using KDP crystal was reported¹². In the following section, the design and measurements of a deep-UV AOTF is discussed.

For deep-UV operation, design configuration using quartz and KDP materials were carefully studied. The AOTF units were fabricated by Brimrose Corporation. In this paper, the preliminary results of the quartz AOTF is presented while the KDP AOTF characteristics that are being hermetically packaged will be presented soon. Hermetical packaging is being done to prevent degradation. Both collinear and non-collinear device configurations were theoretically investigated. Generally, non-collinear configuration is preferred for AOTF operation due to relatively high diffraction efficiency and the

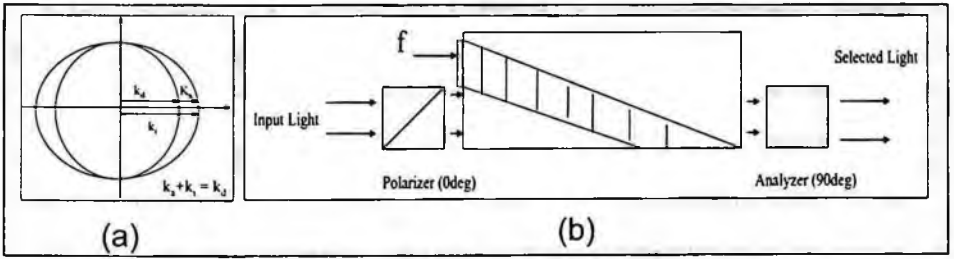


Fig. 3. Left: The collinear phasematching condition in relation to refractive index ellipsoids. Right: The collinear AOTF arrangement.

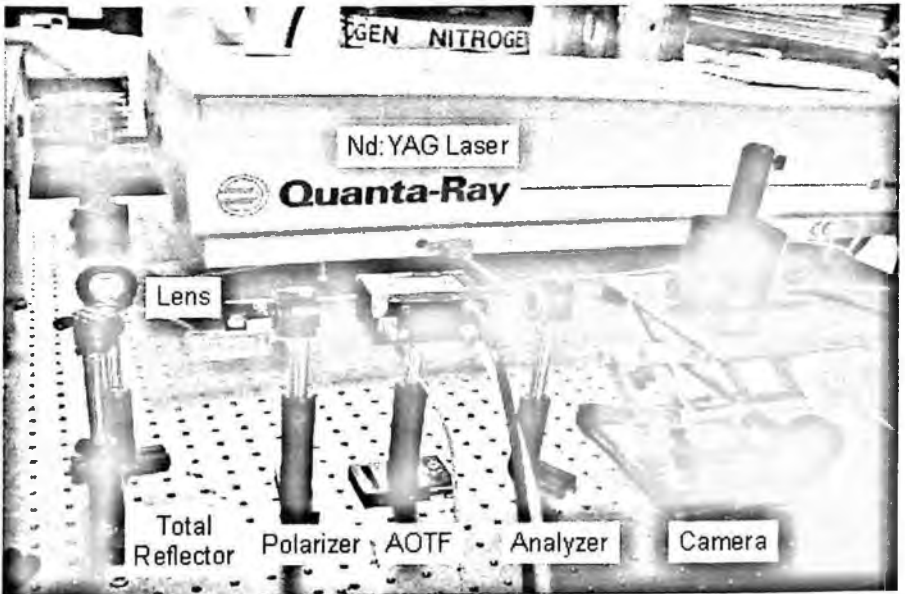


Figure 4. Experimental arrangement using a 355 nm wavelength obtained from a tripled Nd:YAG laser system.

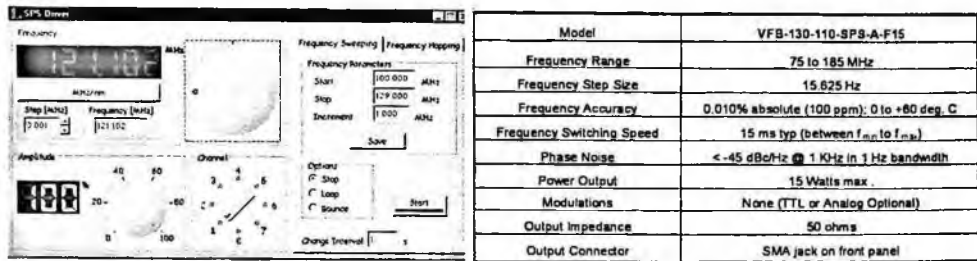


Figure 5. The AOTF control electronics box: Left: Graphical user interface. Right: Control system specifications. (Courtesy of Brimrose Corporation).

Table 1. Design specifications of fabricated quartz and KDP AOTF devices.

Parameter	Units	s	Specifications
Substrate	NA	Quartz	KDP
Spectral Range	nm	200-400	190-450
RF Frequency Range	MHz	82-164	75-185
Optical Aperture	mm	5.0 x 5.0	5.0 x 5.0
Acceptance Angle	Deg.	5.4-7.6	5.4-7.6
Separation Angle	Deg.	~0.5	~0.5
Drive Power	Watts	15	15
Spectral Resolution	nm	0.4 @ 300 nm 0.8 @ 400 nm	< 0.1
Diffraction Eff. At the peak	%	15-20	15-20
Polarization	NA	Linear	Linear
RF Connector	NA	SMA	SMA
Packaging	NA	Water Cooled	Hermetically sealed

advantage of having sufficiently separated output beams. In the case of non-collinear configuration design analysis, the angle between the zero-order and diffracted beam for quartz crystal was found to be around 0.5° . Since, no significant advantages with non-collinear design configuration was evident, collinear configuration was selected for quartz AOTF. The cut angle of the quartz AOTF crystal was 10° . Table 1 shows the test results along with design specifications. Although, the deep-UV laser was available for testing, no polarizers were available at the time of testing. Quartz AOTF characterization was carried out at 355nm due to the availability of a quarter wave plate for separation of orthogonally polarized zero order and the diffracted beams. Hence, tripled Nd:YAG laser providing 355 nm was utilized for our initial characterization. The PRF of the Nd:YAG laser was 10 Hz and the pulsewidth was less than 10 ns. A camera with spectral response ranging from 300 nm to 1 microns was utilized to capture output radiation from the AOTF device.

The AOTF was water-cooled for safety purposes. The first and zero order beams are almost collinear. Hence, a polarizer and an analyzer comprising of linear polarizers placed at right angles relative to each other, one before and one after the AOTF are required for beam separation. In the above table, spectral resolution is measured with collimated output. Diffraction efficiency refers to the laser line. The expected efficiency at the edges is around 10% of the peak value. Further characterization of this AOTF is planned with deep-UV wavelengths in coming days

6. Experimental Results

Measurements of AOTF performance were carried out at 355 nm. The throughput of the quartz AOTF device was >99%. Up to 400 images per wavelength were collected, processed, and analyzed. From the image response, the passband characteristics were determined. In the best case, full width at half maximum (FWHM) less than 0.0625 nm was achieved. From the second set of measurements, a FWHM of 0.125 nm obtained. These passband characteristics are shown in Figure 4 along with the RF drive frequency in the top axis. The measured diffraction efficiency of the device was ~ 15%.

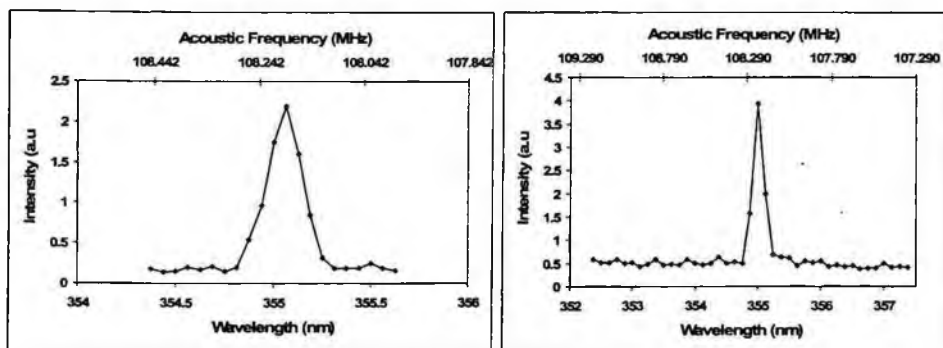


Fig. 6. Measured passband characteristics. Left Graph: Images per wavelength = 400 and FWHM passband = 0.0625 nm. Right Graph: Images per wavelength = 400 and resolution = 0.125 nm.

7. Summary and Conclusions

An all-solid-state Nd:YAG Pump laser has been developed with > 1J/pulse at 50 Hz PRF and a pulsewidth of 22 ns. Using this pump laser, efficient high-pulse-energy, tunable UV wavelength generation technology has been demonstrated. This pump laser has the potential to generate deep-UV wavelengths. A chem-bio sensor scheme is being developed around this UV transmitter technology. Among several critical components, quartz and KDP based AOTFs have been designed and fabricated. Quartz AOTF has provided a passband of less than 0.0625 nm (FWHM) with ~15% diffraction efficiency. Currently, KDP AOTF is being hermetically packaged. In-situ and remote chem-bio sensing experiments are planned in the near future. Potential NASA applications include mineralogical investigations on the Moon, asteroids and planets.

Acknowledgements

The author would like to thank Dr. Jolanta Soos, Dr. Sudhir Trivedi and Dr. Feng Jin of Brimrose Corporation and Prof. Uwe Hommerich of Hampton University for useful discussions and support.

Selected References

1. Narasimha S. Prasad, Upendra N. Sing, Floyd E. Hovis, and Darrell J. Armstrong, "High energy, single-mode, all-solid state and tunable UV laser transmitter," Proc. SPIE Vol. 6214, April 2006.
2. Narasimha S. Prasad, Richard E. Davis, Upendra N. Singh, "Sensing of kerogens and lignens using laser-induced fluorescence at deep-UV wavelengths," Proc. SPIE Vol. 5887, Lidar Remote Sensing for Environmental Monitoring VI; Upendra N. Singh; Ed., Aug 2005.
3. R.M. Dixon, "Acoustic diffraction of light in anisotropic media," IEEE J. Quantum Electron., QE-3, 85-93 (1967).
4. S.E. Harris and R.W. Wallace, "Acousto-optic Tunable Filter," J. Opt. Soc. Am., 59, 744-747 (1969).
5. A. Yariv and Yeh, "Optical Waves in Crystals," Wiley, New York 1984.
6. J. Xu and R. Stroud, *Acousto-Optic Devices*, Jon-Wiley & Sons, Inc., New York, 1992.
7. Neelam Gupta, "Acousto-optic Tunable Filters", optics and Photonics News, November 1997, pp 23-27.
8. I.C. Chang, "Noncollinear acousto-optic filter with large angular aperture," Appl. Phys. Lett. 25, 370-372 (1974).
9. I.C. Chang, "Collinear beam acousto-optic tunable filters," *Electronics Lett.* Vol 28, No 13 PP 1255-6 (1992).
10. I.C. Chang, Tunable Acousto-Optic Filter Utilizing Acoustic Beam Walkoff in Crystal Quartz," *Appl. Phys. Lett.*, 25, 323-4 (Sept. 1974).
11. Narasimha S. Prasad et. al, "Wide bandwidth long-wave IR acousto-optic filter for lidar applications," SPIE's Fourth International Asia-Pacific Environmental Remote Sensing Symposium, 8-11 November 2004, Honolulu, Hawaii, USA.
12. Vitaly Voloshinov and Neelam Gupta, "Ultraviolet-visible imaging acousto-optic tunable filters in KDP," Applied Optics, Vol. 43, No. 19, 1 Jul 2004, pp. 3901-3909.

DEVELOPMENT OF MINIATURE ACOUSTO OPTIC TUNABLE FILTER (AOTF) SPECTROMETER FOR DETECTION OF TOXIC INDUSTRIAL CHEMICALS (TICS)

PETER J. STOPA

*U.S. Army Edgewood Chemical Biological Center
AMSRD-ECB-ENP
5183 Blackhawk Road
Aberdeen Proving Ground, MD 21010-5424 USA
Peter.stopa@us.army.mil*

JOLANTA I. SOOS, RONALD G. ROSEMEIER, SUDHIR B. TRIVEDI and SUSAN W. KUTCHER

*Brimrose Corporation of America
17 Loveton Circle
Baltimore, MD 21152 USA
jsoos@brimrose.com*

This paper describes a remotely operable miniature spectrometer for the detection of Toxic Industrial Chemicals (TIC) using Near Infra-Red (NIR) spectroscopy. The spectrometer discussed here is compact, rugged and capable of operating on either as a handheld unit or on robotic platforms. The packaged unit consists of two modules: an optical head unit and an electronics unit. The optical head contains an infra-red light source, an Acousto-Optical Tunable Filter (AOTF) spectrometer, and associated electronics. The other module is contains the associated electronics assembly and cables connecting it with the optical head. This unit has fiber optic and free space detection capabilities and can be operated using 24 V battery power packs. We have developed a method to detect unknown chemical threat using measured signatures of known chemicals and chemometric techniques. At this time the system can detect chemicals in solid and liquid forms. We are currently optimizing the system for detection of chemicals in gaseous/vapor forms.

Keywords: Toxic Industrial Chemicals; AOTF spectrometer.

1. Introduction and Background

The threat of an accidental or intentional release of a toxic industrial chemical (TIC) is a serious concern. Chemical plants, pipelines, storage facilities, railroads, and trucks are all possible sources from which toxic industrial chemicals could be released. In the event such an incident occurs, there will be an immediate analytical need to determine the nature of the release, establish perimeters to ensure public safety, monitor decontamination efforts, and confirm effective remediation of the agent. It has therefore been a subject of great interest to scientists and engineers over the years to develop methods to detect such threats accurately. Although no single measurement technique is available to analyze for all known TICs, several technologies such as near and mid-infrared (NIR and MIR) spectroscopy,^{1,2} gas chromatography/mass spectrometry

(GC/MS)³, liquid chromatography/mass spectrometry (LC/MS)⁴, and inductively coupled plasma mass spectrometry (ICP-MS)⁵ evolved to address TICs of various forms and characteristics. We present an all optical single sensing instrument, based on NIR and MIR spectroscopy, which can be used for the detection and concentration measurement of a large number of TICs in either liquid or solid phases.

One of the more robust optical strategies to be employed for TIC sensing is absorption spectroscopy. However, most of the previous work based on the NIR and MIR absorption spectroscopy employed discrete optical sources such as laser diodes.^{1,2} Using a typical tunable diode laser absorption spectroscopy system for species concentration measurements, a large number of optical and electronic components are required for *each wavelength*, e.g. laser housings, controllers, fiber couplers, and a complex demultiplexing system. The integration of these components is a major obstacle in keeping multi-species sensing from being utilized in practical situations, as these systems are difficult to package and move when the number of wavelengths increases. In contrast we present a single optical instrument that is light weight, compact, and has a small size without bulky electronics for the detection and measurement of a variety of toxic industrial chemicals. This system utilizes a widely tunable near- and mid-infrared light source to make absorption measurements of the species of interest. The light is able to continuously tune from 0.8 to 2.7 microns while maintaining a narrow bandwidth of a few nm. The rapid tunability of this light source obviates the need for discrete optical sources at various absorption wavelengths of measurement species of interest. The instrument effectively operates as an ultra-miniature spectrometer, capable of quantifying species concentration, *in situ*, with a very small installation footprint that can be conveniently mounted on robots. Furthermore, the wide spectral bandwidth allows for the selection of the optimum absorption transition bands, without regard for the diode lasers at wavelengths difficult to find commercially. The instrumentation is environmentally rugged, with the ability to withstand extreme ranges of temperature, humidity, vibration and shock conditions.

The design and construction of the complete instrumentation is based on proprietary acousto-optic tunable filter (AOTF) technology platform. AOTF spectrometers have several advantages over traditional spectrometers. Traditional spectrometers require careful handling and frequent calibration. They also suffer lower scan speed and lower reliability. An AOTF is an all solid-state tunable filter with no moving parts and is therefore immune to orientation changes or even severe mechanical shock and vibrations. Moreover, the AOTF is a high throughput and high-speed programmable device capable of accessing wavelengths at rates of 100 kHz, making it an excellent tool for *in-situ* NIR spectroscopy. Some of the key advantages of the AOTF technology are summarized below:

- Fast speed. The AOTF wavelength can be electronically changed. The overall speed limit is the time needed for acoustic wave to fulfill the optical aperture,

which can be as fast as several microseconds (μs). This means that the wavelength can be modulated at > 100 kHz.

- Broad tuning range (0.4–5 μm). This results from the dynamic grating formed by the acoustic wave in the acousto-optic medium.
- High optical throughput. As a result of a high diffraction efficiency ($>90\%$).
- Large field of view angle (FOV). AOTFs, unlike conventional Bragg Cells, are capable of a large FOV (as large as a several degrees).
- Electronically programmable center wavelength. The center wavelength of the passband of an AOTF can be electronically programmed and tuned.
- Increased reliability and ruggedness. As a result of the AOTF being a compact solid-state device (5 cm^3) without moving parts, it is insensitive to mechanical vibration and shock, and can be battery powered. Temperature change induced shift of wavelength can be virtually eliminated by adjusting the acoustic frequency accordingly, and thus the device can operate over extreme temperature ranges.

2. AOTF Spectrometer

The fiber optic AOTF spectrometer is divided into several subsystems – collection optics, optical fiber, collimating optics, AOTF, detector system, driving electronics, and data acquisition.

Collection Optics: The AOTF spectrometer is an active detection system. The collection optics couples reflected source light by the surface materials into an optical fiber or fiber bundle, which convey the optical signal to the AOTF.

Optical Fiber: The optical fiber or fiber bundle is transparent across the wavelength range of from 0.5 μm to 4.0 μm for the NIR and MWIR systems. The widely used silica based fiber can only be used for wavelengths up to 2.5 μm . One has to use chalcogenide fiber for MWIR detection. Since the reflected sunlight by the surface material is non-coherent and can have any distribution, multimode fiber is preferred, which also improves the efficiency of collection optics.

Collimating Optics: Collimating optics is used to reduce the size of the beam inside the AOTF so that the diffraction efficiency of AOTF can be higher and the size of AOTF can be smaller.

AOTF: The AOTF operates in the wavelength range of 400nm – 2500nm. It is fabricated using TeO_2 crystal with 13.5°cut. This is the angle between the transducer surface and (001) direction of the crystal. This design provides high efficiency with moderate (1 to 1.5 watt) RF power and acceptable resolution.

Detector System: Because the spectrometer covers a very wide wavelength range (400nm – 2500nm for NIR and 2 μm to 4 μm for MWIR), different types of detectors are required such as Si, InGaAs, and InSb. This coincides with the multi-transducer approach of the AOTF.

The AOTF has two 1st order diffracted beams with orthogonal polarizations when unpolarized light is the input. Conventional AOTF uses only one of them by blocking the other one or by putting a polarizer at the input of the AOTF. In our AOTF spectrometer however, we use both of the diffracted outputs to increase the signal-to-noise ratio.

Driving Electronics: Operation of the AOTF requires an RF driver. In the spectrometer, more than one detector will be used to convert optical signals of different wavelength ranges to electrical ones. Therefore it is possible to allow more than one transducer to operate at the same time without interfering with each other, and the time required to obtain a whole spectrum can be reduced.

Data Acquisition: A multi-channel A/D converter to acquire the optical spectrum from detectors and combine the spectrum from each individual detector to a single spectrum was used. Simultaneous multi-channel A/D converters operating at high data rates allow multiple sampling at each detector for every data point. This allows us to modulate the RF driving signal in order to increase the effective SNR of the system.

3. Detection of Solids and Liquids

In this work, we developed two separate spectrometer designs, one design for solids and liquids and the other for addressing gases. The unit for the detection of solids and liquids mounted onto the robot is shown in Figure 1.

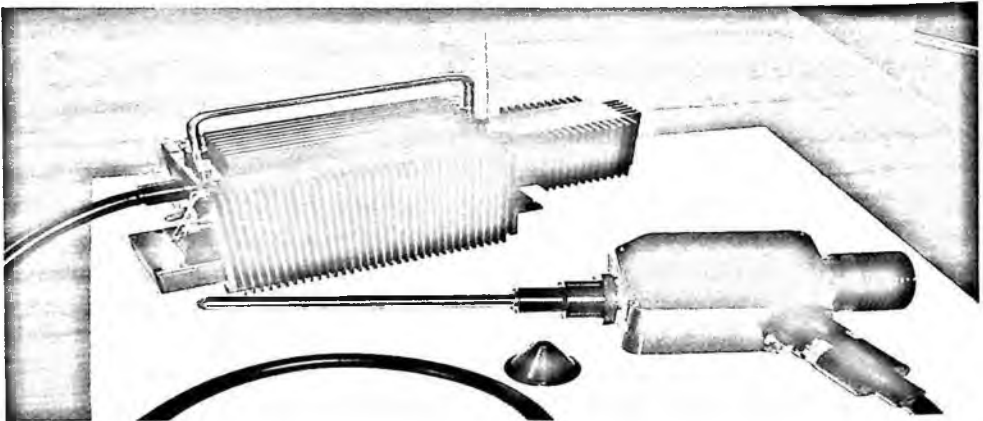


Figure 1. Electronics and optical head of the spectrometer for the detection of toxics in solid/liquid forms.

Using the solid/liquid unit we were able to successfully differentiate between eight different fragrant oils. In this study, a drop of oil was placed on a glass plate with a standard spectralon tile in the background. A wavelength range of 1100 nm to 2300 nm was used in increments of 2 nm. Two hundred scans were taken at a gain level of two and averaged into one spectrum. Three spectra per oil were taken and then processed into absorbance and first derivative. Figure 2 shows an expanded view of the first derivative spectra from 1600 to 1800 nm, which clearly indicates the separation between the different oils.

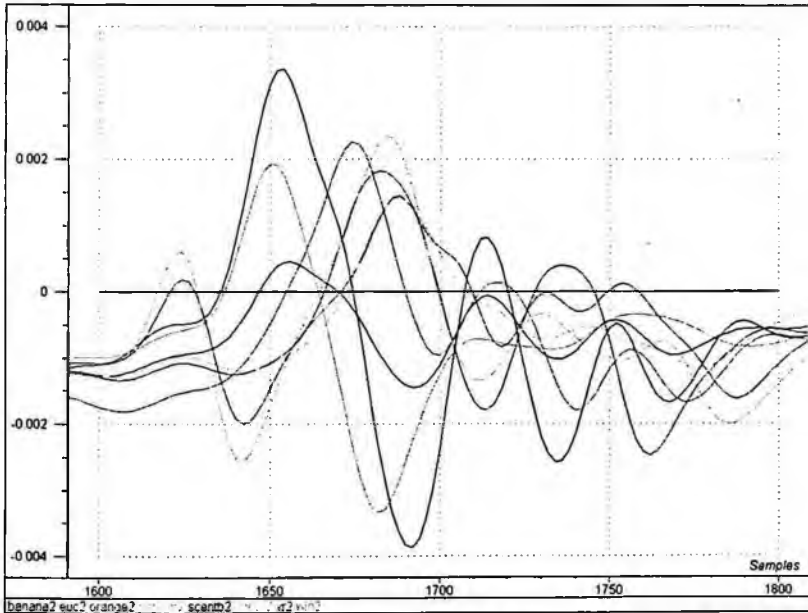


Figure 2. First derivative spectra from 1600 to 1800nm of oils of wintergreen, orange, banana, eucalyptus, peppermint, spearmint and rose.

Next, we performed experimental testing of TIC's on the following liquids: carbon disulfide, acrylonitrile, acrolein, nitric acid, propylene oxide, allyl alcohol, and phosphorus trichloride. The measurements were taken in vials containing three milliliters of liquid. Five spectra from 1100 to 2300 nm were taken of each liquid with 200 scans per spectra. The gain setting was two. The first derivative spectral graph for all of the liquids is shown in Figure 3.

The first derivative spectral data for each liquid was compiled into a prediction profile. The resulting spectra were compared to the compilation of the liquid spectra. A statistical analysis was performed using the Unscrambler 9.0 program, and is shown in Figure 4.

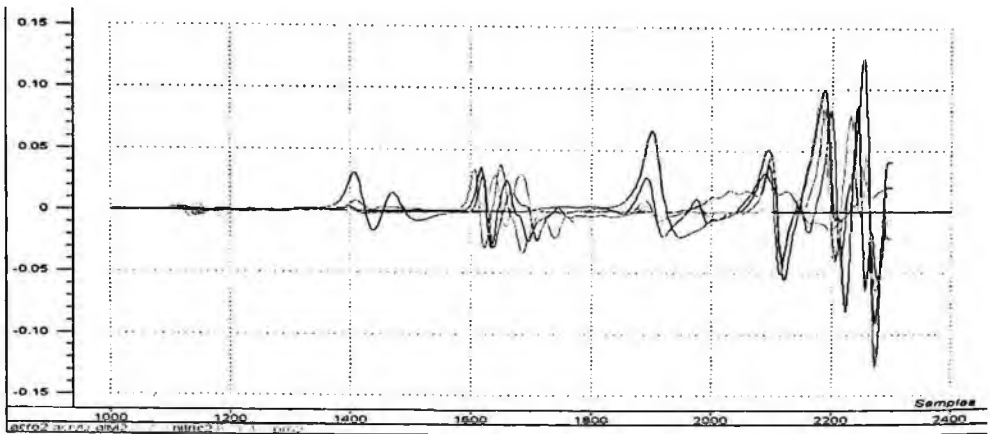


Figure 3. First derivative spectral graphs of the liquids.

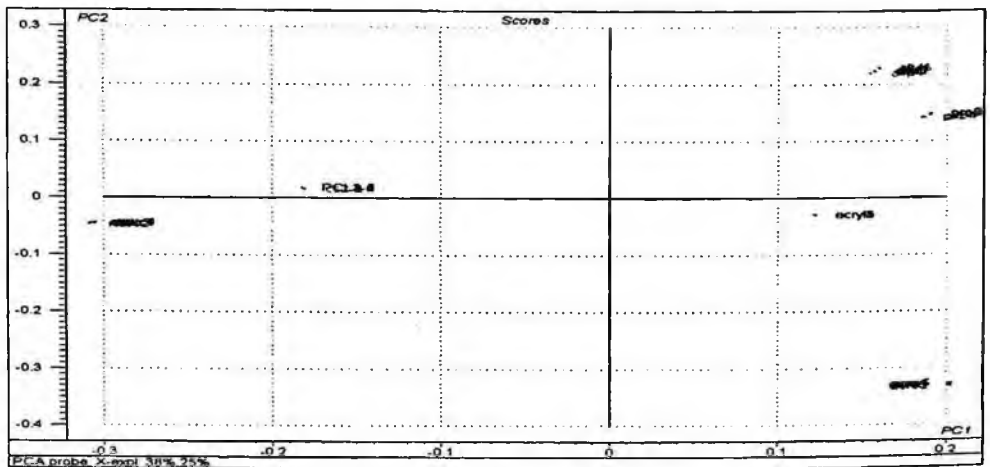


Figure 4. Statistical analysis using Unscrambler 9.0.

Following this first set of measurements, the liquids were then analyzed on butyl rubber gloves. Approximately 1 drop of liquid was placed on a section of the glove and the gain was increased from two to four so as to increase the signal. Five scans per liquid were performed using the cone attachment in direct contact with the glove and there was an average of 200 scans per spectra. The statistical analysis is shown in Figure 5.

Overall we determined that the prediction results obtained using the probes were better when compared to those obtained using the cone attachment. We found that it was difficult to distinguish between the carbon disulfide and the phosphorus trichloride, but the other five liquids were easily distinguishable.

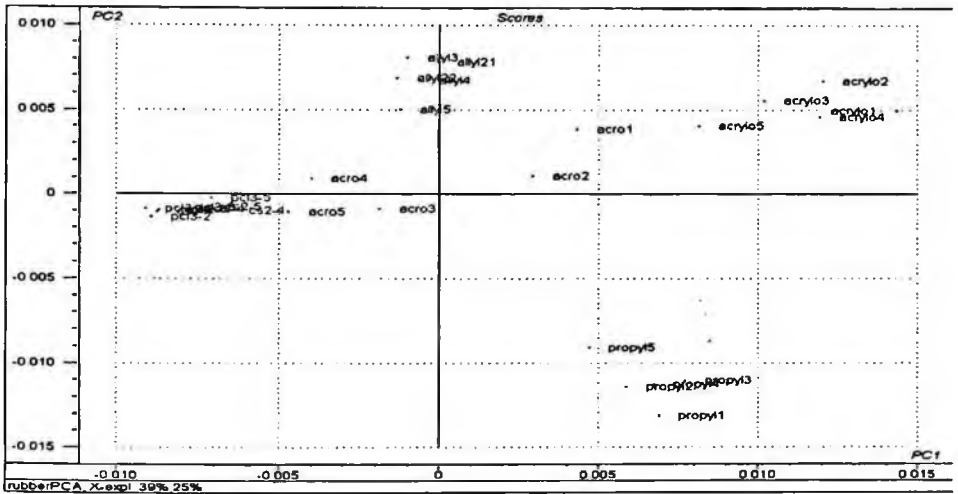


Figure 5. Statistical analysis of liquids on butyl rubber gloves.

4. Detection of Gases

We have also fabricated a spectrometer unit for the detection of gases. We designed both a one-stage and a three-stage gas cell for the collection of the gas samples and ran experiments using water vapor, carbon dioxide and duster gas as samples. Our results indicated that a one-stage gas cell had similar performance in revealing the presence of different gases. A photograph of the one-stage cell is shown in Figure 6. The unit that was fabricated for the detection of gases is shown in Figure 7.

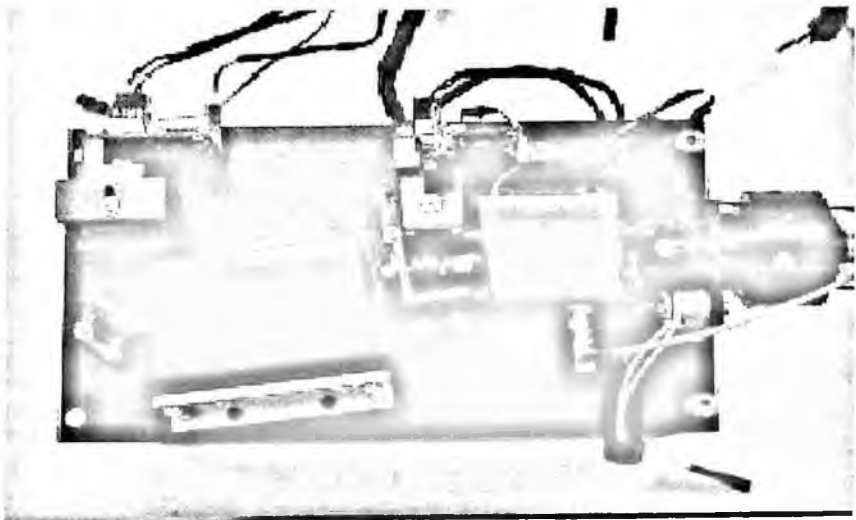


Figure 6. A 1-stage Gas Cell.

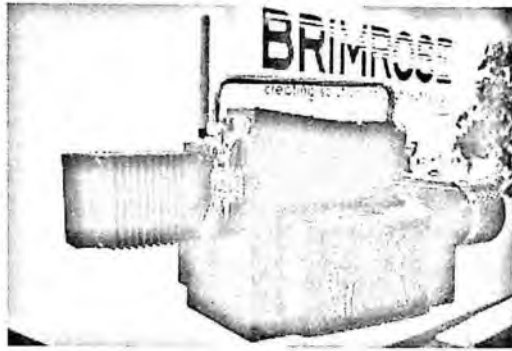


Figure 7. Photograph of spectrometer unit for detection of gases.

Using the spectrometer unit for the detection of gases, we measured the absorbance of six different gases: sulfur dioxide, arsine, bromomethane, chlorine, ethylene oxide, hydrogen chloride and ammonia. The absorbance and its first derivative spectra are shown in Figures 8 and 9.

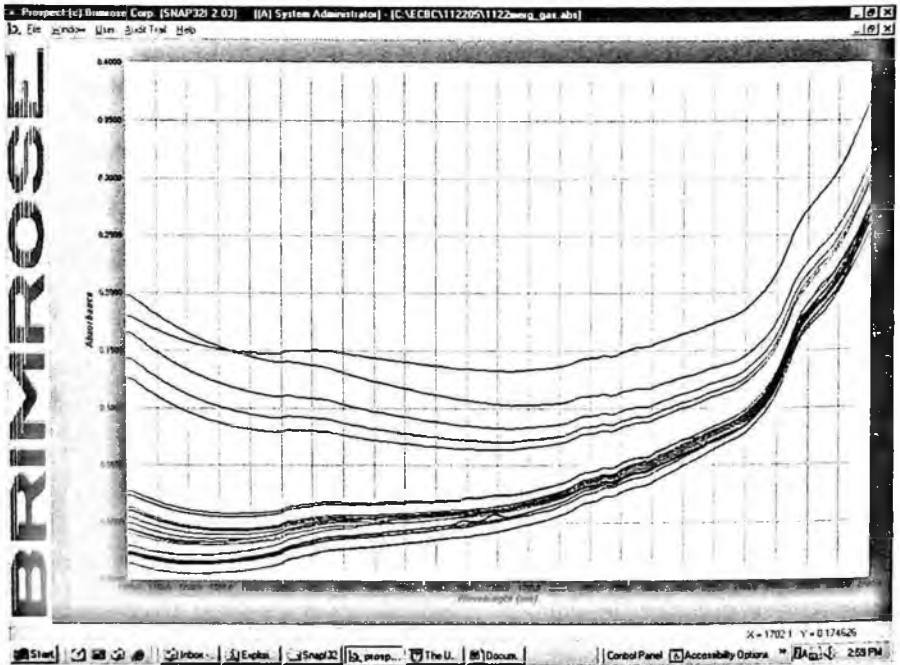


Figure 8. Absorbance spectra of sulfur dioxide, arsine, bromomethane, chlorine, ethylene oxide, hydrogen chloride and ammonia.

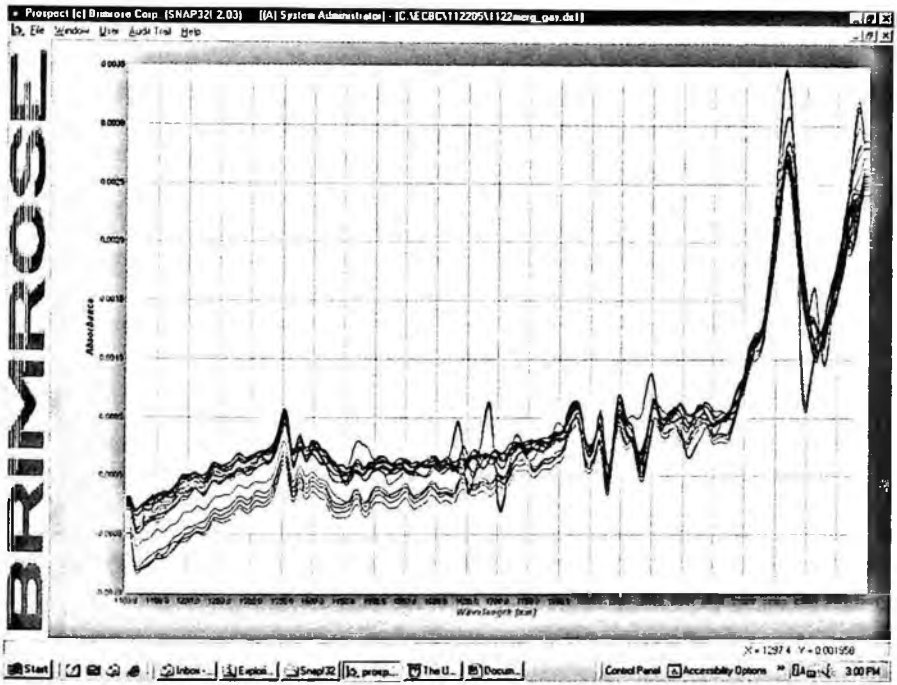


Figure 9. First derivative spectra of sulfur dioxide, arsine, bromomethane, chlorine, ethylene oxide, hydrogen chloride and ammonia.

5. Conclusions

We have successfully demonstrated a remotely operable miniature spectrometer for the detection of TICs using NIR spectroscopy. The spectrometer is compact, rugged and capable of working in manned as well unmanned operations on robotic platforms. This unit has fiber optic and free space detection capabilities and can be operated using 24 V battery power packs. Using measured signatures of known chemicals and chemometric techniques we have developed a method to detect unknown chemical threats. At this time this system can detect chemicals in solid and liquid forms. We are currently optimizing the detection of chemicals in gaseous/vapor form.

6. Acknowledgements

The authors wish to thank the Joint Science and Technology Office of the Defense Threat Reduction Agency for their support of this project.

References

1. P. Werle, Near-and mid-infrared laser-optical sensors for gas analysis, *Optics and Laser Eng.*, **37** 101-114 (2002).
2. D. Wiedmann, Mid-infrared trace-gas sensing with a quasi-continuous-wave Peltier-cooled distributed feedback quantum cascade laser, *Appl. Phys. B* (79), 907-913 (2004).
3. K. Ballschiter, The determination of chlorinated biphenyls, chlorinated dibenzodioxins and chlorinated dibenzofurans by GCMS, *J. High Res. Chromatography* **15**, 260-270 (1992).
4. Niessen WMA and Vander G., *Liquid Chromatography- Mass Spectroscopy*, (Marcel Dekker, NY, 1992).
5. Jackson, B.P., and P.M. Bertsch. Determination of arsenic speciation in poultry wastes by IC-ICP-MS. *Environmental Science & Technology* **35**,4868-4873 (2001).

SPIN DEPENDENT WIGNER FUNCTION SIMULATIONS OF DILUTED MAGNETIC SEMICONDUCTOR SUPERLATTICES – B FIELD TUNING

H. L. GRUBIN

*Department of Physics, University of Connecticut, 2152 Hillside Road,
Storrs, CT 06269-3046, USA
grubin@phys.uconn.edu*

Two terminal devices have traditionally provided band-structure based high frequency operation. Third terminal control often involves hybrid design approaches. The presence of diluted magnetic semiconductor layers in device fabrication should permit the magnetic field to function as a pseudo-third terminal. This is discussed for single barrier, double barrier and superlattice structures, where control is demonstrated. The limits of high frequency operation are discussed in general terms with application to barrier devices and superlattices containing DMS layers.

Keywords: Diluted magnetic semiconductors; RTDs, superlattices, Wigner functions, spin.

1. Introduction

Two terminal devices have traditionally provided band-structure based high frequency operation. For example, Gunn diodes are dependent upon electron transfer from high to low mobility portions of the conduction band [1] and resonant tunneling devices are dependent upon the tunneling characteristics at the interface of materials with different band edges. There has also been significant interest in exploiting two terminal characteristics in three terminal configurations, which in the case of Gunn diodes has resulted in gate triggered logic devices, and in the case of resonant tunneling diode, the presence of a third terminal to alter the IV NDR switching characteristics. But each of these configurations requires the presence of a physical third terminal contact. The introduction of diluted magnetic semiconductors (DMS) into the growth of semiconductor devices has the potential of creating a device technology that plays the role of a third terminal without the presence of a physical contact, and does so by altering, as a function of magnetic field, the band structure of the semiconductor device. This phenomenon arises from an exchange interaction of carriers with spin σ and magnetic atoms with spin S , leading to a splitting of the energy band with different spin up and spin down energies [2].

There are two significant consequences of the energy band splitting: First the populations of spin up and spin down carriers are different. Second if the 'g' factor of the materials is sufficiently high, then in modest magnetic fields there will be significant differ-

ences in the band edges seen by the carriers with different spin. Thus in the presence of a magnetic field, spin up carriers might see a quantum well, whereas spin down carriers would respond to a barrier. These differences form the basis for the types of device configurations one could conceive, which, while not limitless are broader than anything we have at the present time. For example, a double barrier diode with magnetic ions within one barrier could have its IV characteristic tuned by an external magnetic field [3]. A single barrier structure with an interior region containing magnetic atoms could be magnetically tuned to a double barrier structure for carriers with a specific spin [4]. Local regions of magnetic ions near one of the device boundaries could serve a magnetic field dependent nucleation sites for propagating space charge layers in, e.g. Gunn diodes, or superlattice structures [5]. It may be anticipated that interband tunneling phenomena would be impacted by the presence of magnetic ions.

While the origin of the magnetic field dependent transport is quantum mechanical for devices in which the magnetic ions are introduced with gradual spatial variations, "classical" devices can be envisioned where: gradients in the 'g factor can lead to current flow; spin-"up-down-up" transistors as well as spin-"up-down" junctions diodes can be designed. The types of materials that have been discussed within the framework of magnetic atom dependent structures include Gd_2S_3 , EuO , $CdCr_2Se_4$ which are both ferromagnetic and semiconducting, $Cd_{1-x}Mn_xTe$, which is antiferromagnetic, $Ga_{1-x}Mn_xAs$ which is ferromagnetic and semiconducting.

There are a number of existing models for the observed ferromagnetism in semiconductors. Near-field models consider ferromagnetism to be mediated by free holes in the p-type materials. The magnetic Mn ion provides a localized spin and acts as an acceptor in most III-V semiconductors. Ferromagnetism has also been observed in samples with very low hole concentrations and in n-type material. Models of these observations are starting to appear [6]. Much of the early work on DMS focused on GaAs and InAs systems doped with Mn. While this work is confined to p-type material a recent demonstration of room temperature hysteresis in n-GaMnN was reported, along with the effect of the total Mn concentration on the M vs. H and M vs. T behavior, suggesting the possibility of n-type DMS structures [7].

The discussion that follows uses Wigner function numerical simulations to illustrate the behavior of barrier type devices. We examine double barrier, single barrier, hetero-interfaces and superlattices, making the case for a DMS technology. We first present some background material, on the spin splitting, the development of the spin dependent transport, then illustrate with calculations, with one case devoted to increased functionality of the device.

2. Background

To describe this phenomenon, following Haas, we assume that the semiconductor contains one type of magnetic ion. We designate, respectively, the coordinates and spin

of the magnetic atoms as \mathbf{R}_i and \mathbf{S}_i , and their magnetic moments by $-g\mu\mathbf{S}_i$. The interaction of an electron at position \mathbf{r} , with spin σ , with the spins \mathbf{S}_i is:

$$H' = -\sum_i J(|\mathbf{r} - \mathbf{R}_i|) \sigma \cdot \mathbf{S}_i \tag{1}$$

where J is the exchange interaction. If an electron is in an unperturbed band state $f_{nk}(\mathbf{r})$ with energy e_{nk} , then for an energy sufficiently small to be treated by perturbation theory, the first order change in energy is given by:

$$\Delta e_{nk} = \sum_{\alpha} w_{\alpha} \langle \phi_{nk}(\mathbf{r}) X_{\alpha} | H' | \phi_{nk}(\mathbf{r}) X_{\alpha} \rangle \tag{2}$$

where X_{α} is the electron spin function. For collinear spins the average values $\sum_{\alpha} w_{\alpha} \langle \alpha | \mathbf{S}_i | \alpha \rangle$ all parallel to a common direction, the 'z' axis. This makes it possible to choose spin functions X^+ and X^- with the spin of the electron parallel and antiparallel to the z axis. Because all magnetic atoms are equivalent we can write:

$$\int d^3\mathbf{r} |u_{nk}(\mathbf{r})|^2 J(|\mathbf{r} - \mathbf{R}_i|) = J_{nk} / N \tag{3}$$

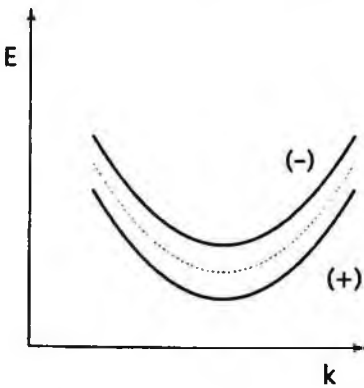


Figure 1. Splitting of the conduction band into sub-bands for spin (+) and spin (-).

where N represents the number of magnetic atoms per unit volume. If only one band is considered the energy change of electrons with spin parallel (+) and antiparallel (-) to the z axis is:

$$\Delta e_{nk}^{\pm} = m \frac{1}{2} J_k (M / M_0) \tag{4}$$

where $M = g\mu_B \sum_i \mathbf{S}_i$ is the magnetization due to the ions, and M_0 , it's value at $T = 0.0K$. The result of equation (4) is sketched in figure 1, where the band splits into a spin (+) band parallel and (-) antiparallel to the magnetization. A consequence of this split is

that the relative populations of spin (+) and spin (-) electrons are different.

In 1982 Ortenberg [8] pointed out that spin doping, achieved by substituting a small amount of crystal constituents with paramagnetic ions, affects the semiconductor electronic g factor. For the semiconductor $A_{1-x}M_xB$, with M indicating the paramagnetic substance, the effective g factor is of the form [9]: $g = g_0 + x \langle S_z \rangle / \mu_B B$. Here $\langle S_z \rangle$ is the thermal average of the localized spins of the paramagnetic dopant along the magnetic field axis. This leads to a modification of the effective Hamiltonian for motion parallel to the magnetic field given by

$$H' = \mu_B g(z) \sigma_z B \tag{5}$$

While the Hamiltonian discussed in [8] included Landau contributions for energy contributions perpendicular to the magnetic field, the heightened interest in the magnetic semiconductors lies, not necessarily in their modulation of the Landau energy levels, but in those materials displaying room temperature ferromagnetism [10] with very high g .

3. The spin dependent quantum transport equations

The Wigner spin dependent quantum transport equations are dependent upon the Hamiltonian and the model for dissipation. As discussed below, these equations are placed in a form separating the spin up, spin down and mixed components, which are coupled through scattering but most importantly through Poisson's equation. These equations may be developed from a set of symmetrized quantum distribution functions [11], from the density matrix in the coordinate representation, from nonequilibrium Greens functions, etc. We concentrate on the density matrix [12], from which we obtain the spin dependent Wigner functions and the corresponding Wigner integro-differential equation of motion. We begin with the spinless transport equations; introduce the necessary definitions and the *reduced coordinate system*.

Schrödinger's equations for conduction (ignoring the spatial dependence of the effective mass) is:

$$i\hbar \frac{\partial \psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m_c} \frac{\partial^2 \psi(x,t)}{\partial x^2} + E_c(x,t) \psi(x,t) \quad (6)$$

The time dependence of the mixed state density matrix for conduction band carriers, $\rho(x, x', t) = \sum_{i=1}^N p_i \psi_i(x, t) \psi_i^*(x', t)$, with, $\int \rho(x, x, t) dx = 1$, is governed by Schrödinger's equation:

$$i\hbar \frac{\partial \rho_c(x, x', t)}{\partial t} = \left\{ -\frac{\hbar^2}{2m_c} \left(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial x'^2} \right) + (E_c(x, t) - E_c(x', t)) \right\} \rho_c(x, x', t) \quad (7)$$

In the reduced coordinate system: $\eta \equiv (x + x')/2$; $\zeta \equiv (x - x')/2$, and equation (7) becomes:

$$i\hbar \left(\frac{\partial \rho_c(\eta, \zeta, t)}{\partial t} + \left(\frac{\partial \rho_c(\eta, \zeta, t)}{\partial t} \right)_{\text{scat}} \right) = \left\{ -\frac{\hbar^2}{2m_c} \frac{\partial^2}{\partial \eta \partial \zeta} + \begin{pmatrix} E_c(\eta + \zeta, t) \\ -E_c(\eta - \zeta, t) \end{pmatrix} \right\} \rho_c(\eta, \zeta, t) \quad (8)$$

In equation (8) we have included a conceptual dissipation term. We will discuss this later. For now we consider the spin dependent density matrix.

For pure states with ℓ representing $(0, X, Y, Z)$

$$\rho(x, x', t) \equiv (\psi^*(x', \uparrow), \psi^*(x', \downarrow)) \sigma^{(\ell)} \begin{pmatrix} \psi(x, \uparrow) \\ \psi(x, \downarrow) \end{pmatrix} \quad (9)$$

where $\sigma^{(i)}$, denotes the Pauli matrices in whose basis the 'z' component is diagonal. There are four components to the pure state spin density matrix, of which only two are of interest for the present discussion:

$$\begin{aligned} \rho(x, x', 0) &= \rho(x, x', \uparrow\uparrow) + \rho(x, x', \downarrow\downarrow) \\ \rho(x, x', Z) &= \rho(x, x', \uparrow\uparrow) - \rho(x, x', \downarrow\downarrow) \end{aligned} \tag{10}$$

In the above: $\rho(x, x', \uparrow\uparrow) \equiv \psi(x, \uparrow)\psi^*(x', \uparrow)$, with a similar definition for the spin down component. For mixed states $\rho(x, x', \uparrow\uparrow) \equiv \sum_{i=1}^{N\uparrow} p_{i,\uparrow} \psi_i(x, \uparrow)\psi_i^*(x', \uparrow)$. The term $N \uparrow$ in the sum represents the number of spin up states.

The Hamiltonian of the system governs the time dependent behavior of the spin dependent density matrix, and while we have already discussed this briefly, we revisit it below. As discussed in [13], when manganese ions Mn^{2+} are introduced into II-VI or III-V semiconductors a huge Zeeman splitting is observed for conduction and valence band electrons with $g^* \sim 100$; and as a lowest order approximation Landau level splitting can be ignored. The enhanced Zeeman splitting stems, as discussed earlier, from the exchange interaction between the extended electronic states in the conduction and valence bands and the magnetic moments of the localized Mn $3d^5$ electrons. The interaction is described as earlier, equation (1), where the sum runs over all Mn atoms n , and S_n is the operator for the total spin of an Mn $3d$ shell at position R_n . Within the framework of first order perturbation theory the matrix elements of H' must be evaluated with respect to the band edge wave functions. These wave functions extend over the whole crystal so that we assume they feel the mean magnetic moment of all Mn $3d$ shells (*the mean-field approximation*) oriented parallel to the magnetic field. The intraband matrix elements are given approximately by, $\langle \nu | J | \nu \rangle N_0 x \langle S \rangle \langle \sigma | \sigma' \rangle$. In this expression the exchange integral is material specific, N_0 is the number of cells per unit volume, and x is the mole fraction of Mn^{2+} ions. The expectation value $\langle S \rangle$ of the spin of the Mn $3d$ shells is described by a modified Brillouin function for spin $j=5/2$:

$$\langle S \rangle = \hat{e} S_0 B_{5/2} (5g_0 \mu_B B / 2k_B T) \tag{11}$$

In the above $g_0 = 2$, \hat{e} is a unit vector parallel to the external magnetic field B . The effective spin S_0 and the effective temperature $T_{eff} = T + T_0$ take into account the anti-ferromagnetic interaction between the Mn^{2+} ions. In the limit in which the argument of the Brillouin function much is less than unity $\langle S \rangle \approx (175/6) \times (\mu_B B / k_B T_{eff})$

Furdyna [14], in discussing the exchange contribution to the l^{th} Landau level, writes the energies of the two spin sub-states as a combination

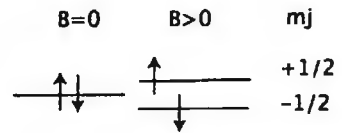


Figure 2. Conduction band Zeeman splitting for a typical diluted magnetic semiconductor

of the 'band' g factor of the electrons, and that arising from an exchange interaction. Within this framework there is an effective g factor given by

$$g_{\text{eff}} = g_0 \left(1 + N_0 x \langle \nu | J | \nu \rangle (175/6) (\mu_B B / k_B T_{\text{eff}}) \right) \quad (12)$$

For the material selected by Furdyna [14], the exchange integral was negative and so the (+) states sustain increases in energy with increasing magnetic field. As discussed in [15] this leads to the picture shown in figure 2 for Zeeman splitting in a typical dilute magnetic semiconductor.

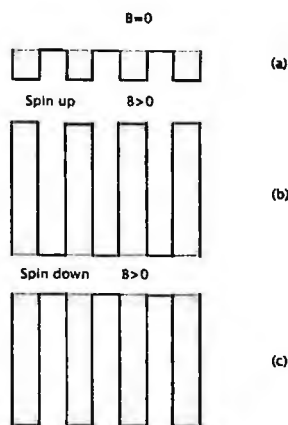


Figure 3. Sketch of a section of a DMS superlattice subject to a B field oriented parallel to the growth direction. Part (a) is for $B=0$. Parts (b) and (c) are for $B>0$

With minor modifications to the exchange energy expression of equation (18), Chang, et al., [16] examined DMS superlattices. A schematic of this structure is shown in figure 3. Lee, et al., [17] examined $\text{Zn}_{1-x-y}\text{Cd}_x\text{Mn}_y\text{Se}/\text{ZnSe}$ multiple quantum well structures. Egues [18] examined perpendicular transport through a ZnMnSe heterostructure, again with the exchange energy expressed as above.

The contributions of the spatial variations in the 'g' factor to the Hamiltonian are expressed as in Eq. (5) with

$$H' = \mu g(z) \sigma \cdot \mathbf{B} \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \phi_c(z) \quad (13)$$

In the above $\phi_c(z) = \mu g(z) B / 2$. With this definition the time dependent equation of motion of the density matrix for the component $\rho(x, x', 0)$, expressed in reduced coordinates, is that given by Eq. (8) with the addition, to the right hand side, of the term $(\phi_c(\eta + \zeta) - \phi_c(\eta - \zeta)) \rho_c(\eta, \zeta, Z, t)$. Thus there is coupling of the two spin contributions. Rather than deal with the coupled components we introduce:

$$\rho(\eta, \zeta, \pm) = (\rho(\eta, \zeta, 0) \pm \rho(\eta, \zeta, Z)) / 2 \quad (14)$$

In light of the above discussion we express the Hamiltonian governing the time dependence of the density matrix, in terms of an effective position dependent g factor:

$$i\hbar \left(\frac{\partial \rho_c(\eta, \zeta, \pm)}{\partial t} + \left(\frac{\partial \rho_c(\eta, \zeta, \pm)}{\partial t} \right)_{\text{exch}} \right) = \left\{ -\frac{\hbar^2}{2m_c} \frac{\partial^2}{\partial \eta \partial \zeta} + \left(\pm \phi_c(\eta + \zeta) \right) - \left(\pm \phi_c(\eta - \zeta) \right) \right\} \rho_c(\eta, \zeta, \pm) \quad (15)$$

Additionally for the simulations presented below, a simple relaxation time approximation was used for the spin up and spin down carriers.

The spin dependent Wigner functions can be obtained either from the spin density matrix or directly from a spin constructed Wigner function as discussed in [10], which we have used in the past. We deal with a modified version of the Wyle [19] transformation:

$$\rho(\eta, \zeta, t) = \frac{1}{(2\pi)^3} \int d^3k f(\mathbf{k}, \eta, t) e^{2i\mathbf{k}\cdot\zeta} \quad (16)$$

The inverse transformation coupled to the density matrix equations yield the time dependent Wigner equations, shown below, which are solved self-consistently with Poisson's equation using procedures outlined in [8].

$$\frac{\partial f(z, \mathbf{k}, \pm)}{\partial t} + \frac{f(z, \mathbf{k}, \pm) - f_0(z, \mathbf{k}, \pm)}{\tau} + \frac{\hbar k_z}{m} \frac{\partial f(z, \mathbf{k}, \pm)}{\partial z} = \frac{1}{i\hbar\pi} \left(\int d\zeta dk'_z e^{2i(k_z - k'_z)\zeta} f(z, k_x, k_y, k'_z, \pm) \begin{pmatrix} [E_c(z - \zeta) \pm \phi(z - \zeta)] \\ -[E_c(z + \zeta) \pm \phi(z + \zeta)] \end{pmatrix} \right) \quad (17)$$

In the Wigner spin dependent simulations we deal with a simple relaxation time approximation and treat transport only by electrons.

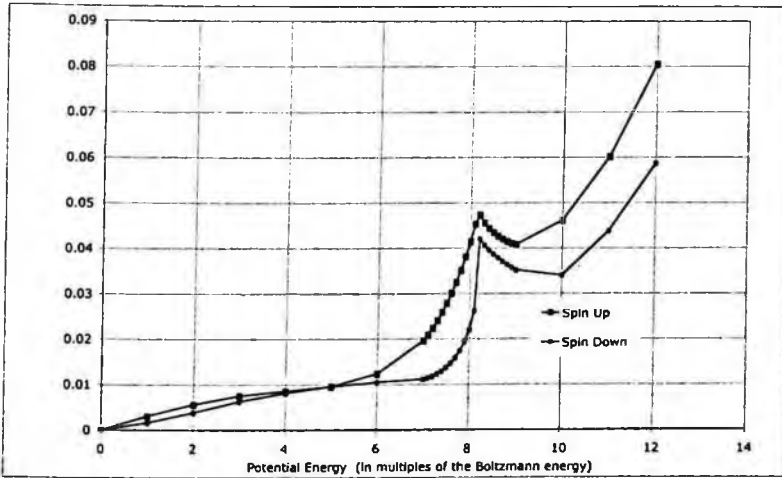


Figure 4. Spin up and spin down IV characteristics for an RTD with DMS emitter barriers.

4. The results

The simulations are designed to demonstrate that the magnetic field coupled to the DMS layers can provide effective third terminal control of an otherwise two terminal structure. In doing so it is direct develop two terminal devices with increased functionality. The devices require the presence of high 'g' factor materials. But as we are looking for devices that sustain negative differential resistance, tunneling times will constrain the upper frequency of operation of these devices. In practice negative differential resistance

is a dynamic phenomenon, as we have discussed in the past, and revisit below. As may be expected the presence of DMS layers in one or both of the barriers, and in the quantum well of a resonant tunneling device will alter the IV characteristics of the device in finite B fields. This is illustrated in figure 4 for a 70 nm structure, with a nominal donor doping concentration of $10^{24}/\text{m}^3$, a central undoped region 30 nm wide and 300 meV-5nm barriers separated by a 5nm. The emitter barrier contains DMS material and combination of g factor and magnetic field decrease (increase) the barrier seen by the spin up (down) carriers by 100 meV. The resulting IV characteristics of each are shown in figure 4. The modest change seen in the figure becomes more pronounced for increased magnetic field. For the simple case of coupling the net IV characteristic with a dc load line, modest changes in magnetic field will alter the switching states of the structure.

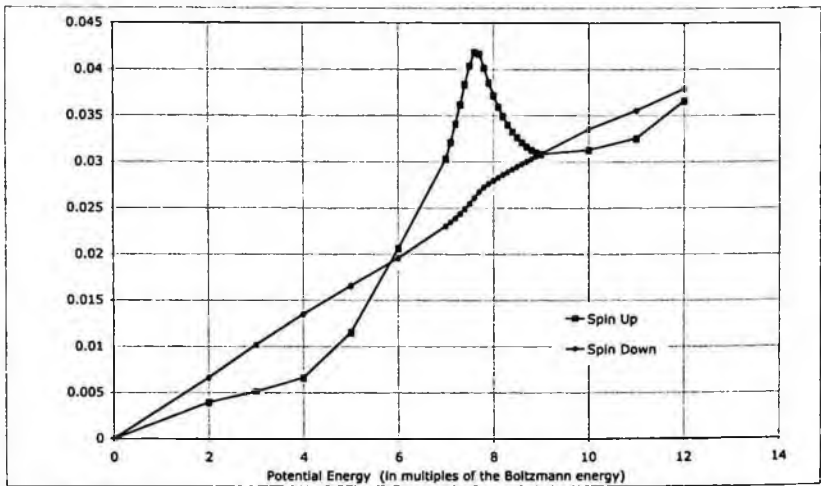


Figure 5. Spin up and spin down IV characteristics for a single barrier diode with a central DMS layer.

Perhaps the most dramatic illustration of this is shown in figure 5 which displays the IV characteristic of a broad 15nm-300 meV barrier embedded again in a 70nm- $10^{24}/\text{m}^3$, a central undoped region 30 nm wide. In this case a central 5nm, with a combined g factor and magnetic field yielding a change in magnetic energy of 300 meV, provides an RTD configuration for the spin up carriers, and an enhanced barrier for the spin down carriers. The result is an IV characteristic that displays negative differential resistance for one species of carrier and a nonlinear positive resistance for the other specie. Thus there is a prospect for designing a device, which in the absence of a magnetic field does not sustain any switching characteristics, to one that can sustain switching and perhaps gain as a function of magnetic field.

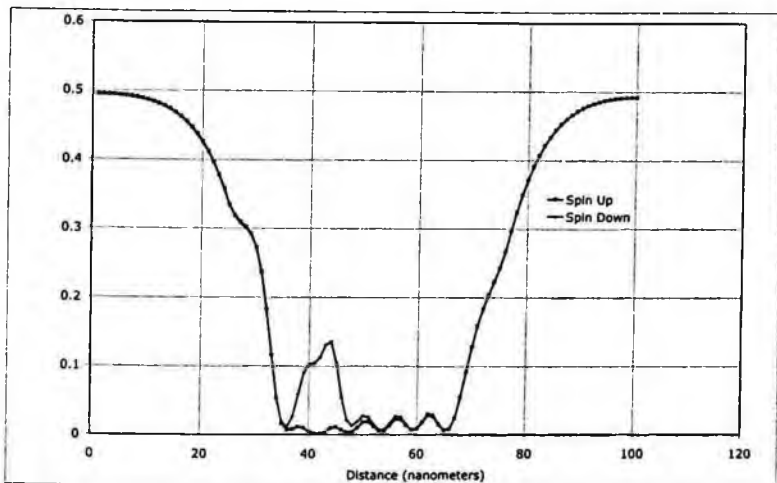


Figure 4. Equilibrium spin up and spin down density distribution for a six barrier – five well superlattice structure with a DMS second barrier.

Oscillations in superlattices are often accompanied by negative differential resistance IV curves, and propagating high field domains. While the device structures discussed in this paper are too small to sustain propagating domains, one of the characteristics of these propagating domains is that they lead to transit time oscillation, and these domains are commonly triggered at a nucleation site either deliberately incorporated into the design of the device, or is randomly triggered at a nonuniformity that occurs during device fabrication. We are proposing that such nonuniformity can be provided by a DMS barrier configuration, and if length dependent transit time oscillations is a sought-after feature, that the DMS barrier configuration be confined to one end of the superlattice device structure. To display some of the features we might expect to find in such a structure, we examined a 100 nm structure, again smaller than those of standard superlattice structures. For this structure there were six barriers and five quantum wells. The barriers were each 3nm wide, 300 meV high and separated by 3 nm quantum wells. The second barrier was a DMS barrier, where the spin up carrier band edge increased with increasing magnetic field, which is the reverse of the first two studies. The combination of g factor and magnetic field yields a 100 meV band edge shift. Figure 6 displays the distribution of spin up and spin down carriers, with an expected population enhancement of the spin down carriers in the vicinity of the DMS barrier. The space charge modulation downstream of the DMS barrier is a consequence of the modulation provided by the superlattice. Small barrier heights do not provide such a dramatic modulation. In the absence of a magnetic field the spin up and spin down populations are equal; here the inequality, which originates at the second barrier spreads to the adjacent two barriers. Figure 7 displays the spin down distribution as a function of bias. There is an increased presence of charge in the vicinity of the DMS barrier. While the decrease in charge for multiple barrier devices, in going from the emitter side quantum well to the collector side quantum well is not un-

common in multiple barrier devices, the enhanced charge distribution in the vicinity of the emitter boundary has as a consequence nonuniformity in the potential energy distribution that could yield a space charge oscillation.

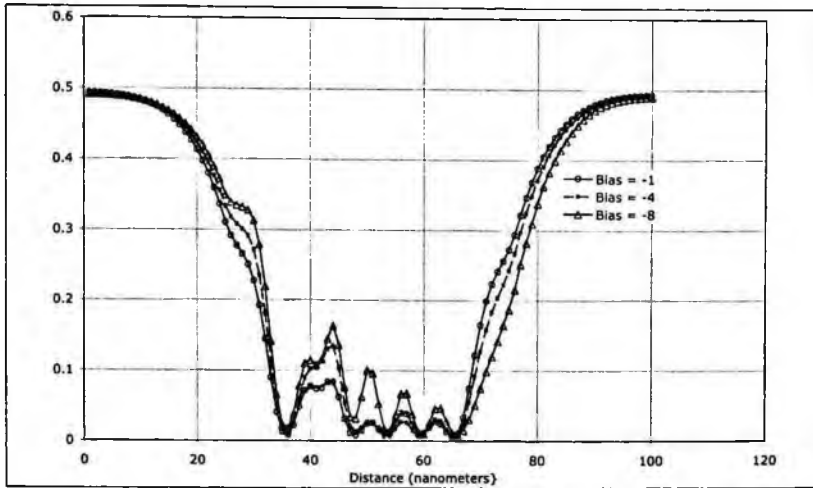


Figure 5. Bias dependent spin down density distribution for a six barrier – five well superlattice structure with a DMS second barrier. Bias is in multiples of kBT .

5. Transient behavior

As indicated earlier there are two parts of device design that we have examined. One is the role of the magnetic field as a controlling pseudo-third terminal, and the second is the ability of the device to sustain negative conductance. Traditional IV curves of negative conductance devices are steady state curves, in which the time scale of the transient calculation is long (effectively infinite) compared to the tunneling times. When the transient time scale is of the order of the tunneling times, the IV curves, which are necessarily dynamic, should differ from the steady state, and we may anticipate the absence of NDR for sufficiently short oscillation periods. There are several ways to examine this. One is to treat the negative differential conductance device as a relaxation oscillator with van der Pol oscillator characteristics. In this case, as long as the device retains some negative conductance characteristics, it will oscillate. In this case we have obtained sustained oscillations for frequencies near 120 GHz, above which the oscillations damped as the negative differential conductance curve softens. An alternative approach subjects the device to a sinusoidal bias at a given frequency. While the transients associated with the sinusoidal oscillator are not the same as that associated with the relaxation oscillator, for in the latter there is usually a very slow rise in potential energy followed by a sharp drop in potential energy, we have found that the steady state IV characteristic is not compromised at frequencies near 1 GHz as displayed in figure 8, but is compromised at frequencies in excess of 40 GHz, when operated in this configuration. This latter frequency be-

ing approximately thirty percent of the maximum frequency associated with relaxation oscillations. Calculations have also been performed with DMS RTD structures. One calculation with a period of 125 ps showed a net particle current density intermediate between the results. At least for the calculations we have performed, the transient character of DMS structures is qualitatively similar to that of non-DMS structures.

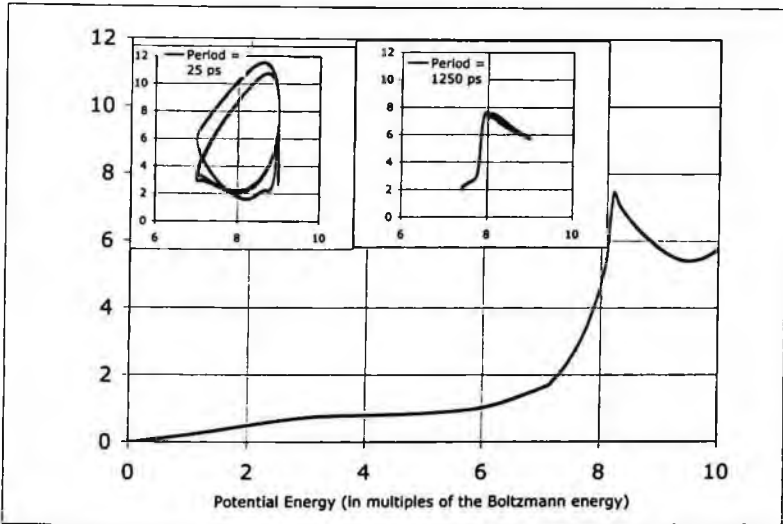


Figure 8. The steady state IV for a double barrier RTD without any DMS layers. The insets are computations for particle current density subject to an ac signal of indicated period, varying from 7.5 kBT to 9.5 kBT.

6. Conclusion

We have demonstrated, at least for n type material DMS layers embedded in barrier devices, that magnetic field control of two terminal devices with increased functionality is possible. Transient appear to be governed by constraints similar to devices that do not contain these layers, indicating that design constraints for the DMS based devices; i.e., frequency limitations, power considerations, can properly evolve from non DMS studies. The band structure richness of the DMS device is only beginning to evolve. Important issues that need to be addressed include the role of the DMS layer for interband devices. How will p-type Mn delta doping affect the operation of n-type devices. But of course, an n-type device technology requires room temperature n-type material. Here we are encouraged by the recent studies of [7] and [6].

7. Acknowledgement

This work was supported by the U. S. Army Research Office, to whom the author is grateful

References

1. Shaw, M.P., H.L. Grubin, and P.R. Solomon, *The Gunn-Hilsum Effect*. 1979, New York: Academic Press. 250.
2. Haas, C., *Spin-Disorder Scattering and Magnetoresistance of Magnetic Semiconductors*. Phys. Rev., 1968. **168**(2): p. 531-538.
3. Grubin, H.L., *Transient Simulations of dilute magnetic semiconductor RTDs* Proc. SPIE, 2005. **5790**: p. 231.
4. Grubin, H.L., *Wigner simulation of the transition of a 'single' to 'double' barrier DMS device*. Proceedings of IEEE-Nano 2005, 2005. **2**: p. 808-810.
5. Grubin, H.L., *Diluted Magnetic Semiconductor Superlattices*. Proc. SPIE, 2006. **To be published**.
6. Sato, K. and H. Katayama-Yoshida, *Material design of GaN-based ferromagnetic diluted magnetic semiconductors*. Japanese Journal of Applied Physics Part 2 - Letters, 2001. **40**(5B): p. L485-L487.
7. Thaler, G.T., et al., *Magnetic Properties of n-GaMnN thin films*. Applied Physics Letters, 2002. **80**(21): p. 3964-3966.
8. Ortenberg, M.v., *Spin Superlattice with Tunable Minigap*. Phys. Rev. Letts, 1982. **49**(14): p. 1041-1043.
9. Kossut, J., *The dependence of the quantum oscillation amplitude on spin splitting*. Solid State Communications, 1978. **27**(11): p. 1237-1240.
10. Pearton, S.J., et al., *Room temperature ferromagnetism in GaMnN and GaMnP*. Physica Status Solidi (a), 2003. **195**(1): p. 222-227.
11. O'Connell, R.F. and E.P. Wigner, *Manifestations of Bose and Fermi statistics on the quantum distribution function for systems of spin-0 and spin-1/2 particles*. Physical Review A, 1984. **30**(5): p. 2613-2618.
12. Grubin, H.L., *Density Matrix Simulations of Semiconductor Devices*, in *Quantum Transport in Ultrasmall Devices*, D.K. Ferry, et al., Editors. 1995, Plenum Press: New York. p. 241-280.
13. Winkler, R., *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems*. Springer Tracts in Modern Physics, ed. J. Kuhn, et al. 2003, Berlin: Springer. 228+xii.
14. Furdyna, J.K., *Diluted magnetic semiconductors*. Journal of Applied Physics, 1988. **64**(4): p. R29-R64.
15. Schmidt, G. and L.W. Molenkamp, *Electrical Spin Injection: Spin-Polarized Transport from Magnetic into Non-Magnetic Semiconductors*, in *Semiconductor Spintronics and Quantum Computation*, D.D. Awschalom, D. Loss, and N. Samarth, Editors. 2002, Springer: Berlin. p. 93-106.
16. Chang, K., J.B. Xia, and F.M. Peeters, *Longitudinal spin transport in diluted magnetic semiconductor superlattices: the effect of the giant Zeeman splitting*. Phys. Rev. B, 2002. **65**: p. 155211-1--8.
17. Lee, S., et al., *Wave-function mapping in multiple quantum wells using diluted magnetic semiconductors*. Phys. Rev. B, 1999. **59**(15): p. 10302-10308.
18. Egues, J.C., *Spin-Dependent Perpendicular Magnetotransport through a Tunable ZnSe/Zn(1-x)Mn(x)Se Heterostructure: A Possible Spin Filter?* Phys. Rev. Letts., 1998. **80**(20): p. 4578-4581.

INDIUM NITRIDE: A NEW MATERIAL FOR HIGH EFFICIENCY, COMPACT, 1550nm LASER-BASED TERAHERTZ SOURCES IN CHEMICAL AND BIOLOGICAL DETECTION

MICHAEL WRABACK, GRACE D. CHERN, ERIC D. READINGER, PAUL H. SHEN

*U.S. Army Research Laboratory, Sensors and Electron Devices Directorate,
2800 Powder Mill Road, Adelphi, MD 20783
mwraback@arl.army.mil*

GREGOR KOBLMÜLLER, CHAD GALLINAT, JAMES S. SPECK

Materials Department, University of California, Santa Barbara, CA 93106

Indium nitride (InN) is identified as a promising terahertz (THz) emitter based on the optical and electronic properties of high quality In- and N-face samples. Time domain THz spectroscopy has been employed to measure the pump wavelength and background carrier concentration dependence of THz emission from InN. There is no discernable difference between the In- and N-face InN samples, as expected for the improved crystalline quality and concomitant low background electron density and high mobility for both polarities. While there is only a weak dependence of THz signal on pump wavelength from 800 nm to 1500 nm, there is a strong dependence on background electron density. Modeling shows that the dominant mechanism for THz generation in bulk InN is the current associated with the diffusion of the photo-generated electrons at elevated electron temperature (photo-Dember effect) and the redistribution of the background electrons under drift, with larger screening from the higher mobility electrons as compared to holes. Compensation or p-type doping in conjunction with manipulation of the large internal electric fields in InN/InGaN nanostructures should lead to significant improvements in THz emitters.

Keywords: InN; terahertz; femtosecond; wavelength dependence; photo-Dember effect.

1. Introduction

The terahertz region of the electromagnetic spectrum, lying between microwave frequencies (100 GHz) and photonic frequencies (30 THz), is a potentially important region for chemical and biological detection¹⁻⁷. In spite of its potential, the THz regime remains one of the least explored portions of the electromagnetic spectrum, in part because of the difficulty in efficiently generating and detecting terahertz radiation. There are two approaches to generating terahertz radiation that have found some success. One approach uses optical pulses of ~ 100 fs duration to produce broadband radiation from semiconductors via acceleration of photogenerated charge, dynamic field screening, optical rectification, or photo-Dember effect⁸⁻¹⁰. Narrow band radiation can also be

created through the mixing of quasi-CW lasers separated in frequency by the desired terahertz difference frequency in a semiconductor photomixer¹¹. Traditionally, optically generated THz systems utilize femtosecond mode-locked Ti-sapphire lasers or CW lasers with a wavelength near 800 nm. Significant advantages associated with the use of telecommunications grade optoelectronics and compact fiber lasers could be gained in cost, size, weight, efficiency, and field deployability by shifting the wavelength to 1550 nm if an appropriate narrow bandgap semiconductor THz source were available.

With a bandgap below 0.7 eV¹², InN is an attractive material amenable to 1550 nm photoexcitation that has not been extensively explored for THz applications. Manipulation of the large internal electric fields in InN/InGaN nanostructures may enable one to far surpass the state-of-the-art in THz emitters. Unlike conventional III-V semiconductor compounds, III-V nitride semiconductors have a wurtzite crystal structure. Wurtzite is a polar structure that can support a significant spontaneous polarization even in the absence of any external field or strain. In InN/GaN multiple quantum wells (MQWs), the termination of the spontaneous polarization at the well/barrier interfaces can lead to a large built-in electric field in InN that points in the growth direction. This internal field can be further enhanced by the piezoelectric field in the strained heterostructure, and the combination of these effects is predicted to create fields about an order of magnitude larger than conventional III-V materials¹³. For example, we have calculated that an InN/Ga_{0.2}In_{0.8}N MQW with 2 nm well width will have a field greater than 2 MV/cm. While the THz radiation field is proportional to the number of photoexcited carriers under low excitation conditions, one can readily gauge the importance of the enhanced internal electric field by considering the MQW structure as a nanoscale capacitor, with the energy transfer from the excitation pulse to the THz pulse mediated by the partial or complete discharge of the nanocapacitor. Therefore the maximum THz pulse energy is limited by the electrostatic energy stored in the nanocapacitor, $U=1/2 \epsilon\epsilon_0 AdF^2$, where $\epsilon\epsilon_0$ is the static permittivity, A is the area, d is the effective width of the capacitor, and F is the electric field inside the capacitor. This simple relation therefore shows that the large internal electric field in InN/GaN MQWs should yield a significant enhancement in the efficiency of THz generation when the sample is excited by short optical pulses in the fs fiber laser spectral region (1550 nm, 1050 nm). In addition, for CW THz generation obtained from the mixing of two frequency offset CW lasers, 1550 nm operation is advantageous because the photomixing conversion factor $(e/h\nu)^2$ is about four times higher than it is at GaAs wavelengths ($\lambda < 850$ nm). This result combined with the high saturation velocity ($> 1.5 \times 10^7$ cm/s)¹⁴, large intervalley spacing (2.7 eV) to confine carriers within the high mobility gamma valley, short carrier lifetime, and potentially high breakdown voltage in InN should lead to up to an order of magnitude higher efficiency InN-based sources compared with the commonly employed arsenides.

In this paper, we report on the development of InN as a potential narrow bandgap semiconductor source of THz radiation under 1550 nm fs or CW laser excitation. The

growth, optical, and electronic properties of optimized In- and N-face InN are presented. Femtosecond pulses tunable between 800 nm and 1500 nm are used to generate THz radiation from this optimized InN. These results are evaluated within the context of p-InAs, the best known THz surface emitter, using drift-diffusion equations incorporating momentum conservation and relaxation.

2. Growth and Properties of High Quality InN

High quality In-face and N-face InN samples have been grown by plasma-assisted molecular beam epitaxy. The In-face InN samples were grown on Ga-polar GaN templates using GaN buffer layers grown in the Ga droplet regime, which significantly improved the surface morphology of the InN layers¹⁵. The N-face InN samples were grown on free-standing N-face GaN or C-face 6H-SiC substrates¹⁶. Growth on the N-face exhibits higher thermal stability, with growth temperatures up to $\sim 100^\circ\text{C}$ higher than for In-face growth¹⁷.

Unintentionally doped In-polar InN films have been realized with room temperature electron mobilities as high as $2900\text{ cm}^2/\text{Vs}$ and free electron concentrations as low as $2.3 \times 10^{17}\text{ cm}^{-3}$. Figure 1 shows the optical properties of an unintentionally doped InN thickness series evaluated at room temperature by CW reflectance, transmission, and photoluminescence (PL) measurements. The absorption coefficient was calculated from the reflectance and transmission data, and its square is plotted versus photon energy to determine the direct bandgap of the InN. Figure 1a shows representative results for a $2.7\text{ }\mu\text{m}$ thick sample. The intercept of the linear fit with the energy axis provides a bandgap of $0.654 \pm 0.002\text{ eV}$. The PL peak was Stokes-shifted by 20 meV to 0.634 eV , implying that the emission emanates substantially from the bandtail states. Figure 1b compares the bandgap energy, peak PL emission energy and residual compressive strain as a function of InN thickness; the PL peak energy, and the residual

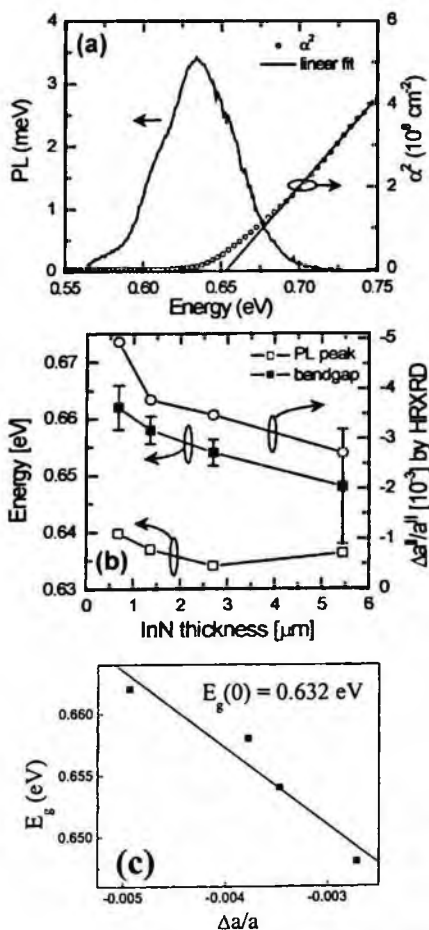


Fig. 1. (a) PL spectra and bandgap from fit to absorption data in InN; (b) Comparison of bandgap energy, peak PL emission energy and residual compressive strain as a function of InN thickness; (c) Extrapolation of bandgap to zero strain.

compressive strain (as measured by HRXRD) for the InN film thickness series. The bandgap decreased from 0.662 ± 0.004 eV to 0.648 ± 0.01 eV as the film thickness increased from 0.68 to 5.45 μm , most likely due to the reduction in residual compressive strain. A similar shift in PL peak energy, from 0.64 eV to ~ 0.635 eV, is observed with increasing thickness. This shift may be less pronounced because the absorption depth of the 900 nm excitation laser is much smaller ($\sim 200\text{--}300$ nm) than the thickness of most of the InN films. Figure 1c shows a linear fit to the relation between the bandgap and the strain. Extrapolation of the bandgap data to zero strain using the deformation potential for InN (4.1 eV) obtained from the slope of the line provides an unstrained bandgap of ~ 0.63 eV. Similar data has also been obtained on optimized N-face InN films. The electron mobility as high as 2370 cm^2/Vs is almost a factor of two higher and the corresponding bulk carrier concentration of 2.8×10^{17} cm^{-3} is nearly one order of magnitude lower than those reported for the best MBE grown N-face InN¹⁷. Optical measurements provide a bandgap of 0.651 eV, with PL peak emission at 0.626 eV, both among the lowest reported values for N-face InN.

3. THz Emission from High Quality InN

The measurements were performed using a Coherent regenerative amplifier (RegA) system. The output of the RegA at 800 nm is split into two beams. The stronger beam is frequency doubled to serve as the pump source for an optical parametric amplifier, which generates a tunable infrared pulse (0.9–2.6 μm) in the idler beam. The infrared beam, after compression with a prism pair to typical pulse widths of approximately 150 fs, is then incident on the InN sample at 45° to the surface normal. The subsequent terahertz emission is collected with a pair of parabolic mirrors onto a ZnTe crystal for electro-optic sampling. The weaker RegA split-off beam is used to probe the terahertz emission for all excitation wavelengths, as well as to pump the InN samples at 800 nm.

Preliminary measurements of terahertz emission from InN thin films deposited on GaN buffers on sapphire using fs

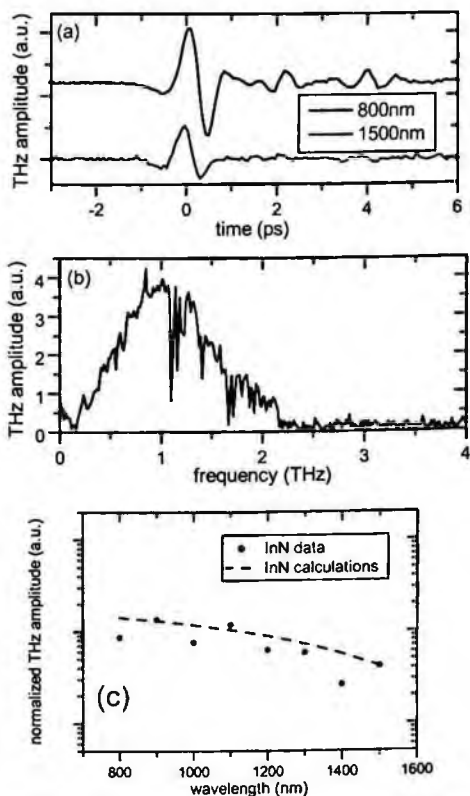


Fig. 2. (a) Time-domain THz signal from InN excited by 800 nm and 1500 nm fs pulses; (b) amplitude spectrum after FFT of the temporal waveform using 800 nm pulses; (c) Pump wavelength dependence of THz amplitude.

pulse excitation at 800 nm have been performed¹⁸. The THz signal from the InN surface was of the same order of magnitude as that from p-type InAs, one of the strongest semiconductor surface emitters of THz radiation. We have extended these measurements to include wavelength- (800 to 1500 nm) and background carrier-dependent studies of the THz emission from InN thin films. Figure 2a shows the typical time-domain waveforms of the THz emission normalized to pump and probe power for bulk *n*-InN excited with 800 nm and 1500 nm fs pulses. The amplitude spectrum in Fig. 2b corresponds to the temporal waveform using 800 nm in Fig. 2a. The spectrum peaks at ~1 THz, with sharp dips due to atmospheric water vapor. A plot of systematic measurements of THz emission from *n*-InN at excitation wavelengths from 800 nm to 1500 nm is displayed in Fig. 2c. Only a weak dependence on pump wavelength is observed, with the signal dropping by less than a factor of 4 over this spectral range.

Further insight into this data can be obtained by modeling using a one-dimensional momentum conservation and relaxation equation

$$\frac{\partial J_n}{\partial t} = e \left(n \frac{eE}{m^*} + \frac{kT}{m^*} \frac{\partial n}{\partial x} \right) - \frac{J_n}{\tau}, \quad (1)$$

with a *k*-dependent effective mass¹⁹ m^* is used to simulate the sub-ps electron current density J_n . The parameters e , n , E , T , and τ are the elementary electric charge, carrier concentration, electric field, carrier temperature, and momentum relaxation time, respectively. Photo-generated and background electron currents are calculated separately due to their different carrier temperatures. Both the photo-generated and background hole currents are calculated using drift-diffusion equations assuming their temperatures are invariant at room temperature. The emitted THz signal is then calculated as the volume integration of the time derivative of the total current density.

From this model we have determined that the dominant mechanism for THz generation in bulk InN is the current associated with the diffusion of the photo-generated electrons at elevated electron temperature (photo-Dember effect) and the redistribution of the background electrons under drift. As the excitation wavelength increases, the photoexcited electron temperature decreases by a factor of 5, thereby lowering the THz amplitude. However, the photon number increases by a factor of 1.9 as the excitation wavelength increases, enhancing the THz amplitude. The combination of these effects leads to the weak dependence on excitation wavelength observed in Fig. 2. In addition, the measured THz amplitude from *n*-InN is smaller than that from *p*-InAs due to larger screening from the higher mobility electrons as compared to holes. For example, at the fixed bulk carrier concentration of 10^{18} cm^{-3} , the calculations in Fig. 3 show that the THz amplitude from *p*-InN would be more than an order of magnitude larger than from *n*-InN. Moreover, the measured normalized THz amplitude from InN shown in Fig. 3 decreases by about one order of magnitude as the bulk carrier concentration n_{bulk} increases by one

order of magnitude. Although there are several other effects which contribute to the THz amplitude, including photo-excited electron temperature, mobility, absorption, and carrier lifetime, screening from carriers appears to be the dominant effect. There is no discernable difference between the In- and N-face InN samples, as expected for the improved crystalline quality and concomitant low background electron density and high mobility for both polarities.

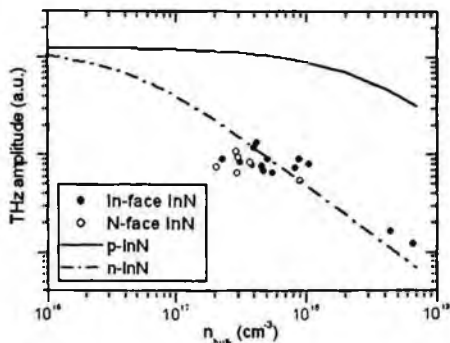


Fig. 3. Background carrier density dependence of THz emission from InN.

4. Conclusions

InN and InN/InGaN nanostructures are promising materials for THz sources pumped by low cost 1550 nm fs fiber and CW lasers because of their combination of small bandgap, large saturation velocity and intervalley spacing, and ability to support MV internal electric fields based on spontaneous polarization. Time domain THz spectroscopy shows only a weak dependence of the THz signal on pump wavelength, dropping by less than a factor of 4 in moving from 800 nm to 1500 nm excitation. Modeling shows that the dominant mechanism for THz generation in bulk InN is the current associated with the diffusion of the photo-generated electrons at elevated electron temperature (photo-Dember effect) and the redistribution of the background electrons under drift, with larger screening from the higher mobility electrons as compared to holes. While these studies indicate that InN can approach the performance of the best bulk semiconductor THz emitter, p-InAs, through compensation/p-type doping, manipulation of the large internal electric fields in InN/InGaN nanostructures should enable significant improvements in THz emitters.

References

1. A.C. Samuels, J.O. Jensen, R.D. Suenram, A.H. Walker, and D.L. Woolard, *SPIE Proc.* **3703**, 121 (1999).
2. A.G. Markelz, A. Roitberg, and E.J. Heilweil, *Chem. Phys. Lett.* **320**, 42 (2000).
3. D.L. Woolard, T. R. Globus, B. L. Gelmont, M. Bykhovskaia, A. C. Samuels, D. Cookmeyer, J. L. Hesler, T. W. Crowe, J. O. Jensen, J. L. Jensen and W.R. Loerop, *Phys. Rev E.* **65**, 051903 (2002).
4. T. R. Globus, D. L. Woolard, T. Khromova, T. W. Crowe, M. Bykhovskaia, B. L. Gelmont, J. Hesler, and A. C. Samuels, *J. Biologic. Phys.* **29**, 89 (2003).
5. H. Liu, Y. Chen, T. Yuan, F. Al-Douseri, J. Xu, and X.-C. Zhang, *SPIE Proc.* **5268**, 43 (2004).
6. J. Barber, D.E. Hooks, D.J. Funk, R.D. Averitt, A.J. Taylor, and D. Babikov, *J. Phys. Chem.* **A109**, 3501 (2005).
7. Shen, Y.C., T. Lo, P.F. Taday, B.E. Cole, W.R. Tribe, and M.C. Kemp, *Appl. Phys. Lett.* **86**, 241116 (2005).

8. X.-C. Zhang, B. B. Hu, J. T. Darrow and D. H. Auston, *Appl. Phys. Lett.* **56**, 1011 (1990).
9. D. Turchinovich, P. Uhd Jepsen, B.S. Monozon, M. Koch, S. Lahmann, U. Rossow, and A. Hangleiter, *Phys. Rev.* **B68**, 241307(R) (2003).
10. K. Liu, J. Xu, T. Yuan, and X.-C. Zhang, *Phys. Rev.* **B73**, 155330 (2006).
11. J.E. Bjarnason, T.L.J. Chan, A.W.M. Lee, E.R. Brown, D.C. Driscoll, M. Hanson, A.C. Gossard, and R.E. Muller, *Appl. Phys. Lett.*, Vol. 85, pp. 3983-3985, 2004.
12. J. Wu, W. Walukiewicz, K. M. Yu, J. W. Ager III, E. E. Haller, H. Lu, W. J. Schaff, Y. Saito, and Y. Nanishi, *Appl. Phys. Lett.* **80**, 3967 (2002).
13. Bernardini, F., V. Fiorentini and D. Vanderbilt, *Phys. Rev. B* **56**, 10024, (1997).
14. W. Liang, K.T. Tsen, D.K. Ferry, H. Lu, and W.J. Schaff, *Appl. Phys. Lett.* **84**, 3681 (2004).
15. Gallinat, C.S., G. Koblmüller, J.S. Brown, S. Bernardis, J.S. Speck, G.D. Chern, E.D. Readinger, H. Shen, and M. Wraback, *Appl. Phys. Lett.* **89**, 032109 (2006).
16. Koblmüller, G., C.S. Gallinat, S. Bernardis, J.S. Speck, G.D. Chern, E.D. Readinger, H. Shen, and M. Wraback, to appear in *Appl. Phys. Lett.* (2006).
17. K. Xu and A. Yoshikawa, *Appl. Phys. Lett.* **83**, 251 (2003).
18. R. Ascazubi, I. Wilke, K. Denniston, H. Lu, and W.J. Schaff, *Appl. Phys. Lett.* **84**, 4810 (2004).
19. E. O. Kane, *J. Phys. Chem. Solids* **1**, 249 (1957).

FLUCTUATION-ENHANCED CHEMICAL/BIOLOGICAL SENSING AND PROMPT IDENTIFICATION OF BACTERIA BY SENSING OF PHAGE-TRIGGERED ION CASCADE (SEPTIC)

L.B. KISH¹, G. SCHMERA², M.D. KING³, M. CHENG¹, R. YOUNG⁴, C.G. GRANQVIST⁵

¹*Department of Electrical and Computer Engineering, Texas A&M University
College Station, TX 77843
Laszlo.Kish@ee.tamu.edu*

²*Space and Naval Warfare Systems Center
San Diego, CA 92152-5001*

³*Department of Mechanical Engineering, Texas A&M University
College Station, TX 77843*

⁴*Department of Biochemistry and Biophysics, Texas A&M University
College Station, TX 77843*

⁵*Angstrom Lab, Uppsala University, P.O.B. 534
Uppsala, SE-75121, Sweden*

Both selectivity and sensitivity of chemical sensors can be significantly improved by exploiting the information contained in microfluctuations present in the sensor system. We call our collection of methods to extract information from these microfluctuations Fluctuation-Enhanced Sensing (FES). In this review paper we discuss general FES principles and two types of applications; gas sensing with commercial solid state sensors and the Sensing of Phage-Triggered Ion Cascade (SEPTIC) technique to detect and identify bacteria.

Keywords: Fluctuation-Enhanced, gas sensing, FES, noise, phage-triggered, ion cascade, SEPTIC.

1. Introduction

The interaction between a chemical sensor and the molecules it detects is a dynamic stochastic process. The resultant fluctuations carry a “stochastic fingerprint”. Conventional sensing methods measure average values and thereby ignore the stochastic component of the signal; these sensors are several orders of magnitude less sensitive than the nose of a canine or even a human.

Why do biological noses perform so much better? They contain a large array of olfactory neurons which communicate stochastic voltage spikes to the brain. When odor molecules are adsorbed by a number of neurons, the statistical properties of these stochastic spikes change. The brain decodes the changes in statistics and matches the result with an odor database in memory.

Recently, a new method, Fluctuation-Enhanced Sensing¹⁻¹² (FES) has been discovered and developed which mimics the biological way of smelling.

2. Fluctuation-Enhanced Chemical Sensing (FES)

The first step is similar to the classical sensing method: the value of a physical quantity in the sensing medium is measured, for example the output voltage of a chemical sensor. Then the microscopic spontaneous fluctuations of these measurements are strongly amplified (typically 1-100 million times) and the statistical properties of these fluctuations are analyzed. These fluctuations are due to the dynamically changing molecular-level interactions between the odor molecules and the sensing media, thus they contain the chemical signature of the odor. The results are compared with a statistical pattern database to identify the odor.

Fluctuation-enhanced sensing has outstanding properties compared to traditional methods:

- (i) It provides significantly more chemical information about the odor. Even a single sensor produces a complete pattern corresponding to the chemical and its concentration (Fig. 1). It is important to note that these results were obtained by a Commercial Off the Shelf (COTS) sensor designed for Nitric Oxide detection. These patterns can be stored in a pattern database and may be used to identify the chemicals belonging to subsequently measured patterns.
- (ii) The relative mean-square fluctuations scale inversely with the active surface area of the sensor. Therefore, the smaller the sensor, the greater the sensitivity of this method. Fig. 1 was obtained by a large sensor with an active surface of a few mm². Using the same kind of sensor material with a submicron size, the estimated sensitivity is in the sub-ppb concentration range.

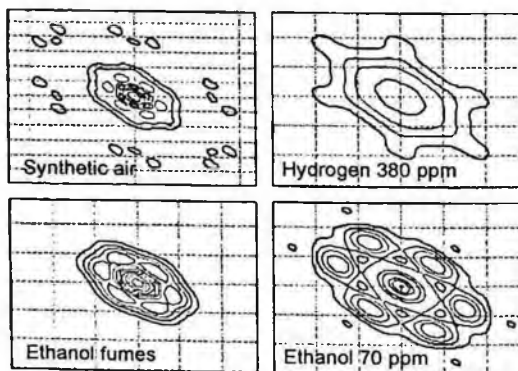


Figure 1. Bi-spectra signatures recorded with a commercial nitric-oxide sensor of surface area of the order of square millimeters. Because the fluctuations scale linearly with the surface area, a micron-sized sensor would allow sub parts-per-billion sensitivity.

Molecule Counting

The amplitude density function of Surface Acoustic Wave (SAW) resonators or MOSFET sensor signals contains information about the number of adsorbed molecules. By measuring the amplitude density function and comparing it with the theoretical form, the concentration of the detected gas can be estimated. If sensor size is in the nanometer range, then the exact number of adsorbed molecules can be determined. Figure 2 shows the result of computer simulations of the output signal of a nanoscale MOSFET sensor.

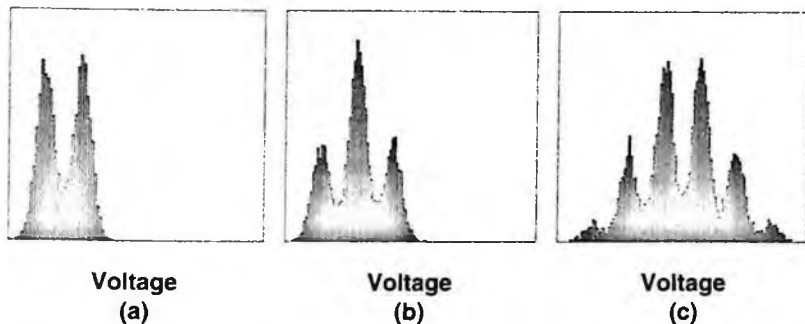


Figure 2. Fluctuation-enhanced sensing (molecular counting) simulation results by using the amplitude density function

Though fluctuation-enhanced sensing has been proven to be superior to classical methods many important questions have to be solved before the method can be industrialized, such as signal separation of simultaneously present chemicals, selection of the most appropriate sensor materials and optimal sensor geometries, etc.

3. FES by using Power Spectra

There are various ways of characterizing the statistical properties of the underlying stochastic processes, such as amplitude density, power spectra, higher order cumulants etc. In this section we illustrate the use of power spectra to extract information from the underlying stochastic processes by utilizing a simple mathematical model. Let us suppose a resistive sensor exposed to a chemical environment, and then use the *stochastic resistance fluctuation* (noise) $dR(t)$ of the sensor resistance R instead of using the induced static change dR of its mean value. The *phenomenological* origin of agent-induced noise in chemical sensors is obvious, as the noise is due to the chemical fragments which originate from an absorbed ambient gas or liquid. After equilibration of the sensor with its environment, the capturing and releasing of molecules becomes a stationary stochastic process. Hence, $dR(t)$ is also a stationary stochastic process.

Three important noise sources exist in the systems of present interest¹⁴: dynamic adsorption-desorption, dynamical percolation¹² of the chemical molecules, and charge

carrier trapping noise¹⁰. These effects provide a noise spectrum characteristic of the chemical species, the applied sensor material, and its microstructure. It is well known that $S(f)$, the power spectrum of the noise, or its higher-order version the bispectrum, is a much more sensitive indicator of changes in the material structure and in the environment than the change of the mean resistance.

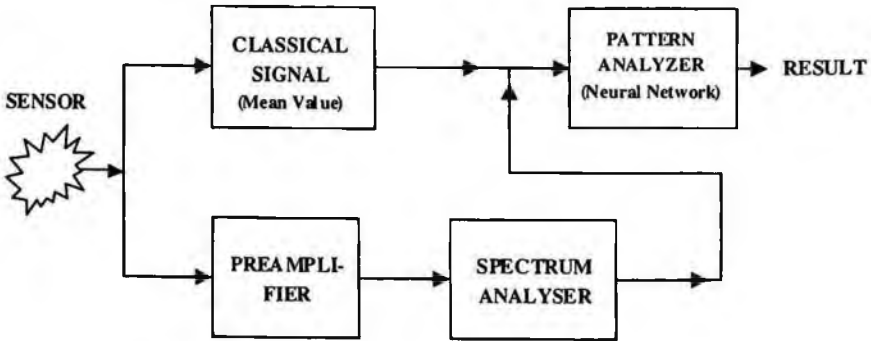


Figure 3. Fluctuation-Enhanced Sensing using power density spectra.

We consider a system with M sensors and N different chemical species. The resistance response of the sensors can be described by a system of equations according to

$$\begin{aligned}
 dR_1 &= A_{1,1}C_1 + A_{1,2}C_2 + \dots + A_{1,N}C_N \\
 &\vdots \\
 dR_M &= A_{M,1}C_1 + A_{M,2}C_2 + \dots + A_{M,N}C_N
 \end{aligned}
 \tag{1}$$

where dR_i is the change of the resistance of the i -th sensor, C_j is the concentration of the j -th chemical species, and A_{ij} is a set of calibration functions. The task is to determine the concentrations C_1, \dots, C_N by the measurement of dR_1, \dots, dR_N . Generally, the equations are nonlinear, and the A_{ij} quantities are functions of all of the concentrations, i.e., $A_{ij} = A_{ij}(C_1, \dots, C_N)$. This non-linearity is the reason why practical sensor systems need a neural network that "learns" these functions during a calibration process.

In order to illustrate the technical problem with the help of the simplest mathematical formalism, and for the sake of simplicity, we first assume that the sensors are linear and that their responses are independent for each investigated chemical species. Therefore the A_{ij} quantities are only calibration constants. From the basic algebra of linear equation systems, it then follows that one needs N independent equations to solve the equation system (1). Therefore, the number M of different sensors has to be larger than or equal to the number N of chemical species, i.e.,

$$M \geq N \quad (2)$$

In a practical case, when the equations are nonlinear, the limit given by Rel. (2) still holds. However, the situation becomes much more complex, with several possible solutions being allowed, and the application of a neural network is required. The validity of Rel. (2) makes electronic noses and tongues cumbersome and expensive because all of the sensors have to exhibit different natures of their response.

We now turn to the case of noise spectroscopy and in particular consider the important issue of the number of sensors needed to analyze a mixture of different chemical species. It is first assumed that only one sensor is present. If the power density spectrum of the resistance fluctuations in this sensor has K different frequency ranges, in which the dependence of the response on the concentration of the chemical species is different from the response in the other ranges, one can write

$$\begin{aligned} dS(f_1) &= B_{1,1} C_1 + B_{1,2} C_2 + \dots + B_{1,N} C_N \\ &\vdots \\ dS(f_K) &= B_{K,1} C_1 + B_{K,2} C_2 + \dots + B_{M,N} C_N \end{aligned} \quad (3)$$

where $dS(f_i)$ is the change of the power density spectrum of resistance fluctuations at the i -th characteristic frequency (or frequency range), and the B_{ij} quantities are calibration constants analogous to the A_{ij} quantities in the linear response limit. Thus, a single sensor is able to provide a set of independent equations to determine the gas composition around the sensor. Considerations corresponding to those that led to Rel. (2) then imply that the number K of different applicable frequency ranges has to be greater than or equal to the number N of chemical species, i.e.,

$$K \geq N \quad (4)$$

In principle, even one sensor can be enough to analyze the composition of a number of different gases. If we have P different detectors, and if we can use the same characteristic frequency ranges for all detectors, then we have

$$\begin{aligned} dS^{(1)}(f_1) &= B^{(1)}_{1,1} C_1 + B^{(1)}_{1,2} C_2 + \dots + B^{(1)}_{1,N} C_N \\ &\vdots \\ dS^{(P)}(f_K) &= B^{(P)}_{K,1} C_1 + B^{(P)}_{K,2} C_2 + \dots + B^{(P)}_{M,N} C_N \end{aligned} \quad (5)$$

where $dS^{(k)}(f_i)$ is the change of the power density spectrum of resistance fluctuations at the i -th characteristic frequency range in the k -th sensor. In the best case, the number of independent equations is PK , so the relation

$$PK \geq N \quad (6)$$

holds.

One should note that the mean resistance variation $dR^{(k)}$ of the different sensors can also be used for detection, and it can provide an additional independent equation. The relevant relation is then

$$P(K + 1) \geq N \quad (7)$$

4. Fluctuation-Enhanced Biological Sensing: SEPTIC

Fluctuation-Enhanced methods can be used to detect biological agents. Existing methods are slow; require culturing, laborious assaying by trained personnel and bulky equipment (except some special narrowly applicable methods for quick tests which can identify only certain types of bacteria).

On the other hand, the revolutionary SEPTIC (SEnsing of Phage-Triggered Ion Cascade) technique¹³⁻¹⁵, if fully developed, has the following outstanding characteristics:

- Can test a library of bacteria within 5 minutes
- Can detect a wide variety of bacteria
- Detection limit without extra concentrator: 10 bacteria/microliter
- No false alarms
- Easily used by non-technical personnel, portable, ready for cycle testing and field applications
- Small, lightweight, expandable to automatic mode
- Ten years shelf lifetime of bio sensing agent in the sensor
- Superior to other methods because utilizing identification technique by the natural enemy (bacteriophages) of the bacterium

The SEPTIC method is a combination of virology and microelectronics, it is a technique for identifying small quantities of bacteria in minutes.¹³⁻¹⁵ Phages are viruses that attack bacteria with the same verve that some other viral species attack people. But phages are very specific about their prey. Most species can parasitize only a single sort of bacterium. Phages start their attack by injecting DNA into their victims. When this happens, there is a short-lived flow of ions (electrically charged atoms and molecules) out of the bacterium. This phenomenon, whimsically described as the bacterium "bleeding" or "screaming", is the first indication that the phage has found its target. By detecting the

"scream", the exact type of bacteria can be determined. The microelectronics is doing the "hearing" the scream which is picked up by two microscopic antennas on the biochip. Initial tests of SEPTIC with *E. coli* showed excellent. SEPTIC is applicable to detect a wide variety of bacteria, such as anthrax, plague, botulism and shigella.

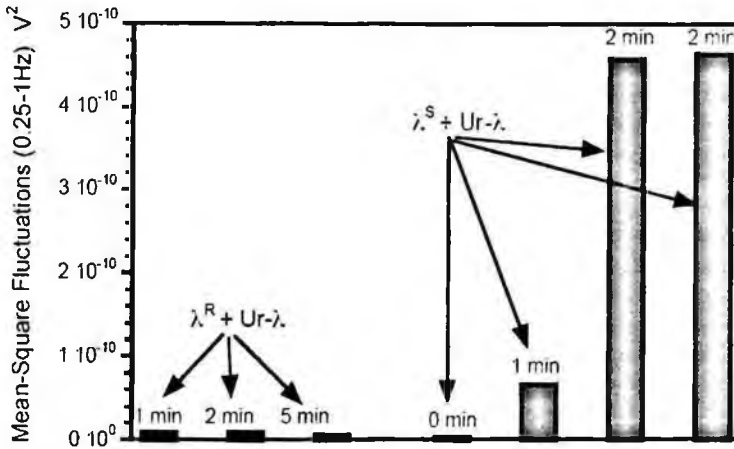


Figure 4. SEPTIC results. Red: positive experiments with *E. coli* and relevant phages at different waiting times. Blue: control experiments with *E. coli* and irrelevant phages and different waiting times.

The SEPTIC method has the potential of applications in biological agent detection and medicine. For example, an industrialized SEPTIC sensor would be able to identify the presence of anthrax in minutes.

5. Conclusion

We have introduced the basic principles of Fluctuation Enhanced Sensing and demonstrated its superior sensitivity and selectivity for gases and bacteria detection. It is clear from the presented results that Fluctuation Enhanced Sensing (FES) and the SPETIC method in particular, offer significant advantages when compared with traditional methods.

References

1. L.B. Kish, Y. Li, J.L. Solis, W.H. Marlow, R. Vajtai, C.G. Granqvist, V. Lantto, J.M. Smulko, G. Schmera, "Detecting Harmful Gases Using Fluctuation-Enhanced Sensing", *IEEE Sensors Journal* 5 (2005) 671-676.
2. J.L. Solis, G.E. Seeton, Y. Li, L.B. Kish, "Fluctuation-Enhanced Multiple-Gas Sensing", *IEEE Sensors Journal* 5 (2005) 1338-1345.
3. J. Ederth, J.M. Smulko, L.B. Kish, P. Heszler and C.G. Granqvist, "Comparison of classical and fluctuation-enhanced gas sensing with PdxWO₃ nanoparticle films", *Sensors and Actuators B* 113 (2006) 310-315.

4. J. Smulko, J. Ederth, L.B. Kish, P. Heszler, C.G. Granqvist, "Higher-Order Spectra in Nanoparticle Gas Sensors", *Fluctuation and Noise Letters* 4 (2004) L597-L603.
5. J.M. Smulko, L.B. Kish, "Higher-Order Statistics for Fluctuation-Enhanced Gas-Sensing", *Sensors and Materials* 16 (2004) 291-299.
6. J.L. Solis, G. Seeton, Y. Li, L.B. Kish, "Fluctuation-Enhanced Sensing with Commercial Gas Sensors", *Sensors & Transducers Magazine* 38 (2003) 59-66.
7. G. Schmera, L.B. Kish, "Surface diffusion enhanced chemical sensing by surface acoustic waves", *Sensors and Actuators B* 93 (2003) 159-163.
8. G. Schmera, L.B. Kish, "Fluctuation Enhanced Chemical Sensing by Surface Acoustic Wave Devices", *Fluct. Noise Lett.*, 2 (2002) L117-L123.
9. A. Hoel, J. Ederth, J. Kopniczky, P. Heszler, L.B. Kish, E. Olsson and C.G. Granqvist, "Conduction invasion noise in nanoparticle WO_3/Au thin-film devices for gas sensing application", *Smart Mater. Struct.* 11 (2002) 640-644.
10. J. Smulko, C.G. Granqvist, L.B. Kish, "On the statistical analysis of noise in chemical sensors and its application for sensing", *Fluct. Noise Lett.*, 1 (2001) L147.
11. J.L. Solis, L.B. Kish, R. Vajtai, C.G. Granqvist, J. Olsson, J. Schnurer, V. Lantto, "Identifying natural and artificial odors through noise analysis with a sampling-and-hold electronic nose", *Sensors and Actuators B* 77 (2001) 312.
12. L.B. Kish, R. Vajtai, C.-G. Granqvist, "Extracting information from noise spectra of chemical sensors: single sensor electronic noses and tongues", *Sensors and Actuators B* 71 (2000) 55.
13. J.R. Biard and L.B. Kish, "Enhancing the sensitivity of the SEPTIC bacterium detection method by concentrating the phage-infected bacteria via DC electrical current", *Fluct. Noise Lett.* 5 (2005) L153-L158.
14. L.B. Kish, M. Cheng, J.U. Kim, S. Seo, M.D. King, R. Young, A. Der, G. Schmera, "Estimation of Detection Limits of the Phage-Invasion Based Identification of Bacteria", *Fluct. Noise Lett.*, 5 (2005) L105-L108.
15. M. Dobozi-King, S. Seo, J.U. Kim, R. Young, M. Cheng, and L.B. Kish, (corresponding author) "Rapid Detection and Identification of Bacteria: SENSING of Phage-Triggered Ion Cascade (SEPTIC)", *Journal of Biological Physics and Chemistry*, 5 (2005) 3-7.

THE PHENOMENOLOGY OF HIGH EXPLOSIVE FIREBALLS FROM FIELDED SPECTROSCOPIC AND IMAGING SENSORS FOR EVENT CLASSIFICATION

KEVIN C. GROSS

*Riverside Research Institute, 2681 Commons Blvd
Beavercreek, Ohio 45431 USA
kevin.gross@afit.edu*

GLEN P. PERRAM

*Department of Engineering Physics, Air Force Institute of Technology, 2950 Hobson Way
Wright-Patterson Air Force Base, Ohio 45433-7765 USA
glen.perram@afit.edu*

Conventional munitions emit intense radiation upon detonation which spans much of the electromagnetic spectrum. The phenomenology of time-resolved visible, near- and mid-IR spectra from these fast transient events is poorly understood. The observed spectrum is driven by many factors including the type, size and age of the chemical explosive, method of detonation, interaction with the environment, and the casing used to enclose the explosive. Midwave infrared emissions ($1800\text{--}6000\text{ cm}^{-1}$, $1.67\text{--}5.56\text{ }\mu\text{m}$) from a variety of conventional military munitions were collected with a Fourier transform spectrometer (16 cm^{-1} , 21 Hz) to assess the possibility of event classification via remotely sensed spectra. Conventional munitions fireballs appear to be graybodies in the midwave. Modeling the spectra as a single-temperature Planckian (appropriately modified by atmospheric transmittance) provided key features for classification and substantially reduced the dimensionality of the data. The temperature cools from $\sim 1800\text{ K}$ to ambient conditions in $3\text{--}5\text{ s}$, often following an exponential decay with a rate near 1 s^{-1} second. A systematic, large residual spanning $2050\text{--}2250\text{ cm}^{-1}$ was consistently observed shortly after detonation and may be attributable to hot CO_2 emission at the periphery of the fireball. For two different explosive types detonated under similar conditions, features based on the temperature, area and fit residuals could be used to distinguish between them. This paper will present the phenomenology of detonation fireballs and explore the utility of physics-based features for explosive classification.

Keywords: bomb phenomenology, detonation fireball spectra, Fourier-transform spectroscopy, classification, feature extraction.

1. Introduction

Most commercial remote sensing endeavors focus on the observation of gradual change (hours, days or months) to enhance meteorological predictions, understand climate change, and improve crop management. To effectively characterize the battle space, the military is often interested in dynamic events occurring on much shorter timescales, from

minutes to sub-seconds. Time-resolved infrared profiles collected from a variety of platforms have been studied for rocket plumes,¹⁻⁴ missile and aircraft emissions,⁵⁻⁶ muzzle flashes,⁷⁻⁸ and to a limited extent detonation fireballs.⁹⁻¹⁴ The temporal response of band integrated intensity contains information useful in distinguishing broad classes of events. However, the munitions classification problem is particularly challenging for several reasons: (1) several factors complicate and decrease the reproducibility of detonation signatures, e.g. casing and mixture tolerances, detonation method, target interaction, and atmospheric conditions; (2) the variability in band-integrated temporal signatures within a munitions class is comparable to the variability between different classes; (3) the available data is sparse and spans multiple degrees-of-freedom with limited repetitions. To date, we are unaware of features derived from infrared temporal signatures being successfully used to discriminate between different explosives. Capturing spectral variations in the infrared may provide the information needed to distinguish one explosive type from another since absorption and emission features in this wavelength regime can often be associated with molecular species. However, the transient nature of detonation fireballs poses instrumentation challenges and trade-offs between temporal, spectral and spatial resolution common to hyperspectral platforms must be carefully considered. Moreover, extracting reproducible yet distinguishing features from time-resolved spectral data requires an understanding of the physics governing the infrared emission from high-explosives detonations. Although bomb detonation pressure waves¹⁵ and condensed matter detonation and shock kinetics are well studied,¹⁶⁻²⁰ the conversion of reaction exothermicity to infrared and visible emissions is poorly documented. This research group is developing a phenomenological model for infrared emission from the detonation of high explosives to assist the munitions classification problem. We are also exploring the challenges associated with collecting time-resolved spectra of rapid infrared events via interferometry. To aid this effort, a collection of spectrometers, radiometers, and visible and infrared imagers have been deployed in several field tests to collect optical signatures from the detonation of conventional military munitions (CMMs), enhanced novel explosives, and improvised explosive devices. The model currently reproduces the CMM spectral and radiometric data with high fidelity, substantially reduces data dimensionality, provides key features for classification, and offers physical insight. Distilling the time-resolved hyperspectral data to a minimal collection of descriptive features enables pattern recognition techniques to be applied to the classification problem. The physical insight gained aids the discovery of key features that are reproducible for a munitions class and distinct from other munitions classes. This paper summarizes our spectral modeling of CMMs and the utility of derived features in discriminating two classes of ordnance.

2. Methodology

In the Fall of 1999, a set of field tests were conducted at Fallon Naval Air Station to study the optical infrared signatures arising from the detonation of conventional munitions. Three types of explosives (A, B and C) spanning several weights (extra small,

small, medium, and large) were used for a total of 56 events. About half the events were air delivered by an F-18 aircraft and the rest were statically detonated on the ground. The detonation fireballs were observed using a Fourier-transform spectrometer (FTS) operating a 16 cm^{-1} resolution. Multiple spectra were collected for each event at a rate of 21 Hz. The FTS was equipped with an InSb ($1800\text{--}6000 \text{ cm}^{-1}$) and HgCdTe ($500\text{--}6000 \text{ cm}^{-1}$) detector for coverage of the mid-wave infrared (MWIR) spectrum. The background noise level for the InSb detector was about five times smaller than that of the HgCdTe detector. Calibration was performed with a blackbody at multiple temperatures resulting in data collected in intensity units ($\text{W}/\text{sr}\cdot\text{cm}^{-1}$). Absolute radiometry was verified with four 200 Hz radiometers, each with a different MWIR bandpass filter. The test range, Bravo 20, was a dry lakebed about 4000 ft above sea level. The geographical layout is detailed in Fig 1. Further details about the test setup and calibration procedure can be found in Orson's thesis.¹³

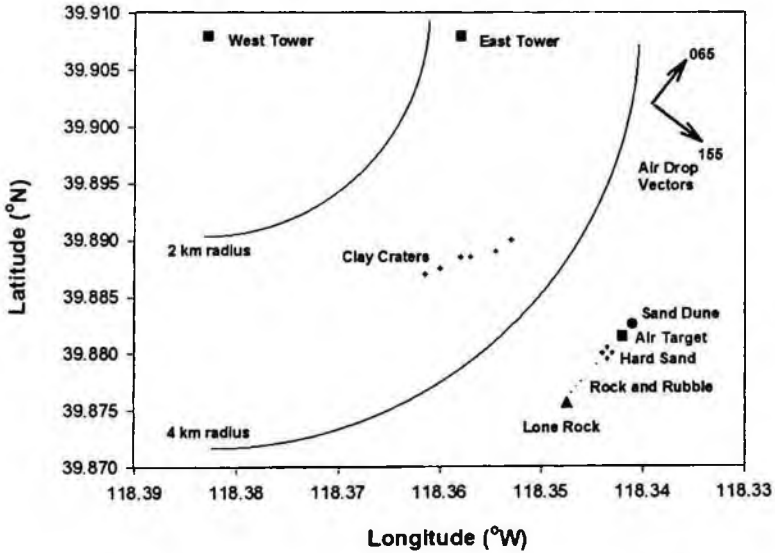


Fig 1. Layout of the Radiant Brass set of field experiments conducted at the Bravo 20 test range, Fallon Naval Air Station, Nevada. An F-18 delivered munitions along three different approach vectors. Static detonations occurred at Clay Craters, Hard Sand and Rock/Rubble.

A simple phenomenological description of the fireball was developed to reduce the dimensionality of the data and extract features for classification. The observed fireball intensity I_{obs} was modeled by a Planckian distribution, modified by the transmittance of the atmosphere (τ), with temperature T and emissive area ϵA as the fit parameters, i.e.

$$I_{obs}(\bar{\nu}, t) = \tau(\bar{\nu}) \epsilon A(t) \frac{2hc^2 \bar{\nu}^3}{e^{hc\bar{\nu}/k_B T(t)} - 1} \tag{1}$$

Since the Planckian source is a slowly-varying function of frequency, the sharp atmospheric absorption features in the observed data reveal information about the concentration of the absorbers. An iterative method based on the on-resonance off-resonance ratio of intensities at many absorption features was used to compute the concentration of atmospheric H₂O, CO₂, N₂O, and CH₄. This method does not require a model for the source intensity term and only assumes that its spectral variation with frequency is slow relative to most atmospheric absorption features. Accurate meteorological data was not available for the field test; however, reasonable agreement was found between the calculated H₂O concentration and that reported by the nearest weather station. Experimental CO₂, N₂O and CH₄ concentrations were not available, but calculated values were consistent with historical averages. Atmospheric transmission modeling was performed using the Line-by-Line Radiative-Transfer-Model.²¹

Fourier-transform spectroscopy is based on the Michelson interferometer and assumes the scene under analysis does not change appreciably during each mirror scan. Spectral resolution improves with total distance traveled by the mirror, but at the expense of temporal resolution. The transient nature of detonation fireballs requires that a judicious choice of spectral and temporal resolution be made to avoid corrupting the data with temporal aliasing artifacts. Assuming the Planckian fireball model above, we quantified the error associated with the Fourier transform of an interferogram acquired when the temperature decayed exponentially with a varying rate k_T . The spectrum computed from an interferometer scanning at rate (k_i) was compared to the true spectrum² at zero-path difference and defined the temporal aliasing error E. The largest error at any frequency was computed for several values of k_i/k_T and is presented in Fig 2. As discussed below, the choice of 16 cm⁻¹ resolution provided an acquisition rate of 21 Hz and minimized the effects of temporal aliasing.

A classification technique based on Fisher linear discrimination (FLD) was applied to features extracted from the phenomenological model just described. To illustrate how this method works, Fig 3 presents a scatter plot of two features (F2 versus F1) for several fictional events, each a member from one of three distinct classes (A, B and C). While neither feature alone provides ample separation between all three classes, a combination of these features can be found which clearly separates all classes. FLD finds the line in this feature space that maximizes between-class scatter and minimizes within-class scatter upon projection. The mean and standard deviation of each class projected onto this Fisher line can be used to define a probability distribution function (PDF). When projecting a new event onto the Fisher line, the PDF assigns a probability that it belongs to a particular class. Feature saliency and PDF stability are critical in assessing the confidence level associated with a prediction, and are thoroughly examined in Dills' dissertation.²²

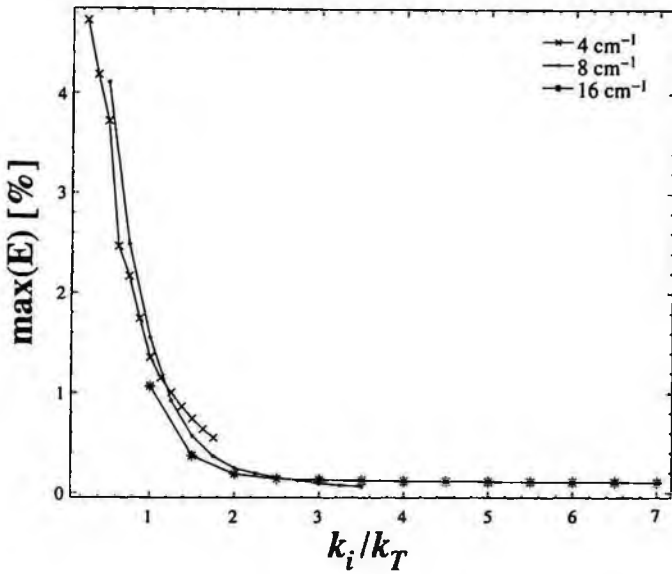


Fig 2. Maximum spectral error (E) attributed to temporal aliasing for an interferogram scanned at a rate k , that collects an atmospheric-attenuated Planckian source that changes in time according to an exponentially decaying temperature (k_T).

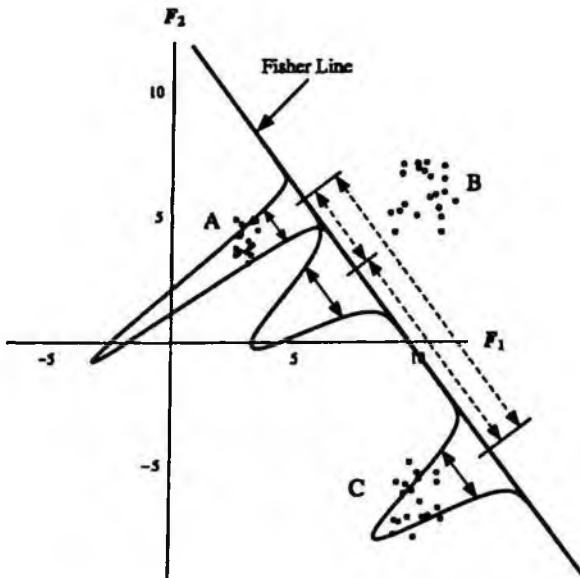


Fig 3. Example data consisting of three classes (A, B & C) with two descriptors (F_1 & F_2) illustrates the Fisher linear discrimination technique. When the data is projected onto the Fisher line, between-class scatter (dashed arrows) is maximized and within-class scatter (solid arrows) is minimized. Gaussian curves for each class, defined by the mean and standard deviation upon projection onto the Fisher line, are also provided.

3. Results and Discussion

A typical fireball spectrum from the InSb detector is provided in Fig 4. In regions where the atmosphere is transparent, the spectra of detonation fireballs arising from both the airdropped and statically-detonated ordnance were fairly broadband in nature. Absorption from atmospheric gases such as H₂O and CO₂ is readily identified by characteristic spectral features. Modeling the observed spectra I_{obs} by a single-temperature Planckian distribution modified by the atmospheric transmittance profile accounted for most of the variation in the data. This fit is presented with the data in Fig 4 along with the residual ΔI (data minus the fit) at two different time steps. In the transmission bands ($\tau > 0.4$), the average relative error is less than 5%. The small residuals indicate that the fireball is well-described by a single temperature and that our estimation of the atmospheric transmittance function was successful. However, a systematic underestimation of the intensity was observed near 2000–2200 cm⁻¹ shortly after detonation and changes with time. The cause is currently being investigated and could be the result of emission from hot CO₂ in the fireball. Accurate removal of the atmospheric effects is critical to interpret the fit residuals as non-Planckian behavior of the fireball. The time-dependent nature of the residual at 2152 cm⁻¹ indicates that it is a feature of the dynamic fireball and not an artifact of incorrectly modeling atmospheric transmission.

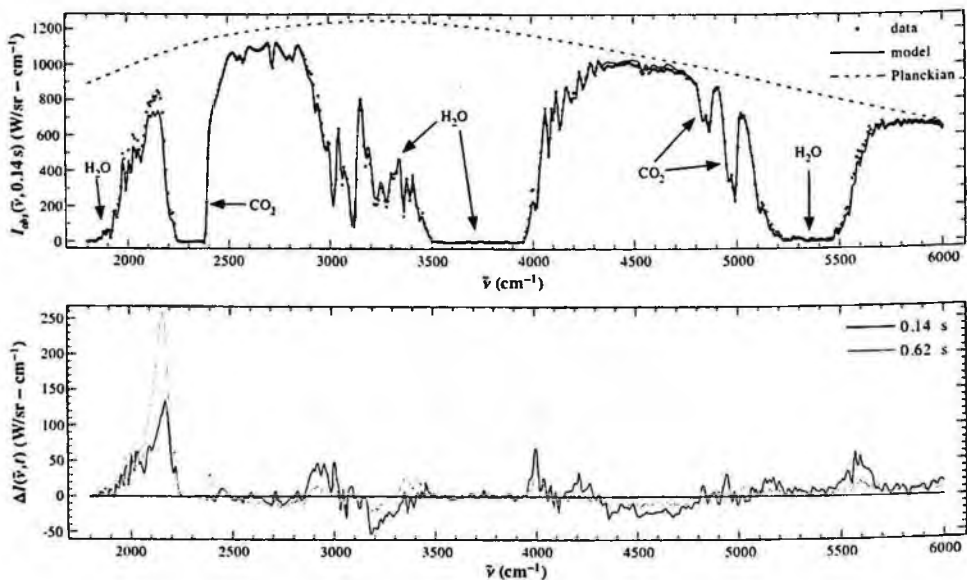


Fig 4. Typical fit results for the static detonation of a small Type A bomb. In the top plot, the third spectrum after detonation ($t = 0.14$ s) is compared with the attenuation-modified Planckian fit. The difference between the observed data and model ΔI is shown at both $t = 0.14$ s (—) and $t = 0.62$ s (---) in the bottom plot.

Fitting the data to this simple phenomenological model affords a substantial dimensionality reduction while preserving most of the original fidelity. Each data matrix (intensity vs. frequency and time) is reduced to two vectors, namely temperature and area versus time. A representative example of the temperature and area curves are presented in Fig 5. The standard error in the area was about ten times greater than temperature, and is periodically displayed in Fig 5. Initial temperatures were 1700–2000 K and often followed an exponential decay with rates between 0.91–1.24 s^{-1} . Examining Fig 2, we note that errors in the residuals due to temporal aliasing contribute less than 0.5% since the interferometer scanned at least 16 times faster than the temperature decay rate.

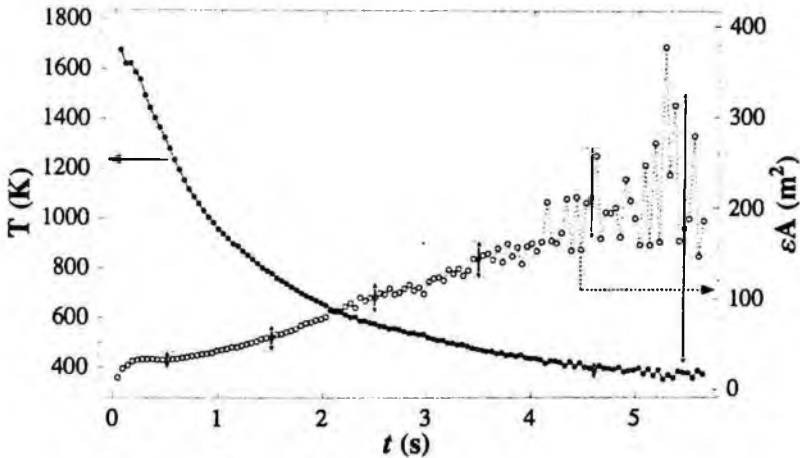


Fig 5. Temperature and area curves extracted from time-resolved spectra of statically-detonated Type A Small munitions. The periodic error bars reflect the standard error in the fit parameter.

Features based on these fit parameters and residuals can now be explored for their potential use in classifying munitions type. The most useful feature for graphically separating the events into their respective classes was the time dependent fit residual at 2152 cm^{-1} . Figure 6a compares this feature for a set of Type A Small (AS) explosives with a set of Type B Large (BL) explosives. Both sets of munitions were air-delivered. The temperature and area curves provided a more subtle but still discernible separation between these two classes. A variety of features based on the temperature, area and fit residuals were subject to Fisher linear discrimination. The most salient features were consistently based on the fit residuals, i.e., the non-Planckian character of the fireball emissions. As an example, applying FLD to the AS, BL subset using three values taken from the 2152 cm^{-1} residual—the initial, maximal, and $t = 0.5 \text{ s}$ values of ΔI —indicates its utility to the classification problem. The PDFs shown in Fig 6b are well separated for the two classes. The limited number of repetitions available in the Radiant Brass field data makes assessing confidence in the PDFs' predictive capability difficult. Dills' work addresses this issue using bootstrapping and cross-validation techniques.¹⁴ It was found

that the initial and maximum values of ΔI were salient features, and were also useful indiscriminating static detonations from dynamic ones and identifying the explosive weight if the type was known.

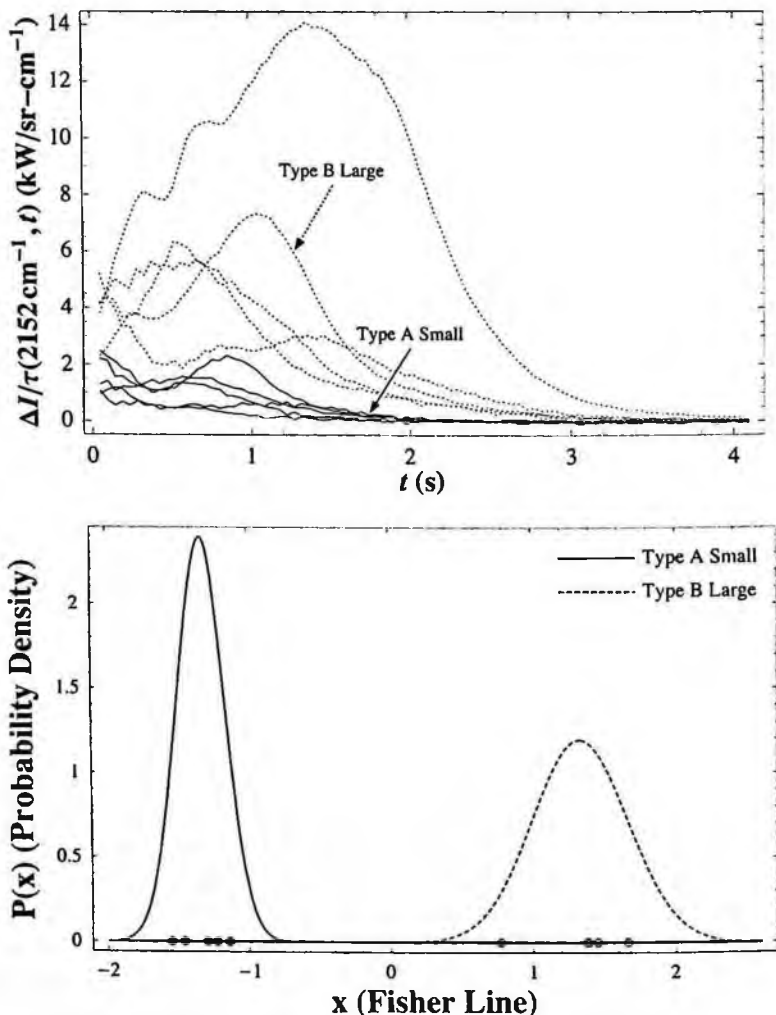


Fig 6 (a) The temporal evolution of the fit residual at 2152 cm^{-1} of (—) air dropped Type A Small and (···) Type B large ordnance, (b) Probability density functions generated using FLD using three features taken from residual, namely the initial, maximal and $t = 0.5$ values of ΔI , with weights 0.522, 0.548, and 0.654, respectively.

4. Conclusions

Our efforts to collect, characterize, model, and derive classifiers from time-resolved infrared spectra of conventional explosives was summarized in this document. Detailed

accounts of this research can be found in the References and in several forthcoming publications. We now conclude with several important findings of these efforts.

Fourier-transform spectroscopy can be used to study fast infrared events provided the selection of temporal resolution is commensurate with event timescales. For an exponentially cooling Planckian radiator, temporal aliasing of the spectra introduces a maximum error of less than 0.5% when the FTS scan rate is at least twice the temperature decay rate. Cased munitions are broadband across most of the MWIR and are well-described by a single-temperature Planckian distribution. A systematic underestimation of the observed intensity by the Planckian model is found between 2000–2200 cm^{-1} and is likely the result of hot CO_2 emissions in the fireball. The phenomenological model provides high fidelity dimensionality reduction and offers several key, discriminating features. The time-dependent fit residual at 2152 cm^{-1} was the best feature for graphically separating two different explosive types, suggesting that the non-Planckian nature of detonation fireball is key to the classification problem. Accurately capturing the non-Planckian features by the fit residuals requires an accurate atmospheric transmission function. With even moderate spectral resolution (16 cm^{-1}), the sharp absorption features in the MWIR due to atmospheric water and carbon dioxide can be used to estimate their respective concentrations assuming only that the source emission is broadband in nature. Inverting concentrations of other trace gases, namely N_2O and CH_4 , is also possible. The substantial dimensionality reduction afforded by the current fireball model enables pattern-recognition methodologies to be applied to the classification problem. Probability density functions based on FLD were developed using non-Planckian fireball emission features and indicate strong classification potential. Additional field tests that both expand the diversity and number of repetitions will help define this potential and reduce the amount of *a priori* information needed to classify events. These pattern recognition techniques have also been applied to features derived from fireball imagery with promising results.¹⁰⁻¹¹

A synergistic relationship exists between the modeling and classification efforts. The development of a reasonably accurate model is essential to uncovering good classifiers since the parameters offer physical interpretation and are more likely to capture real fireball characteristics. (Although dimensionality reduction is possible with an arbitrary parametric representation, the loss of feature interpretability is often accompanied by degraded classification potential.) On the other hand, the Fisher discrimination techniques unveil the most salient features and thereby guide our efforts in improving the model. We will soon include a CO_2 emission term to directly model the non-Planckian behavior observed near 2150 cm^{-1} . Physical models capable of explaining the temperature and area dynamics are also being developed, and the empirical model¹² to describe time-resolved intensity curves seems well-suited to this problem. With these developments, a faithful representation of the spectral data cube will soon be expressed by a collection of 10-15 physically-interpretable fit parameters.

5. Acknowledgments

We gratefully acknowledge practical insight and constructive comments from Jim Engle and Randy Bostick. We thank Sean Miller, Tom Fitzgerald and their team for excellent work in test design and data collection under challenging field conditions.

References

1. A. Blanc, L. Deimling, N. Eisenreich, 2002. UV- and IR signatures of rocket plumes. *Propellants, Explosives, Pyrotechnics* 27(3), 185-189 (2002).
2. P.F. Bythrow, D.A. Oursler, Detection of transient optical events at narrowband visible wavelengths, Johns Hopkins APL Technical Digest 20(2), 155-161 (1999).
3. F.S. Simmons, *Rocket exhaust plume phenomenology* (Aerospace Press, 2000).
4. D.E. Siskind, M.H. Stevens, J.T. Emmert, D.P. Drob, Signatures of shuttle and rocket exhaust plumes in TIMED/SABER radiance data, *Geophysical Research Letters* 30(15): 1-4 (2003).
5. D.R. Crow, C.F. Coker, High-Fidelity Phenomenology Modeling of Infrared Emissions from Missile and Aircraft Exhaust Plumes, *Proceedings of SPIE* 2741, 242-250 (1996).
6. H.M. Schleijsen, M.W. Craje, and S.M. Eisses, High resolution, spectroscopy in the field. *Proceedings of SPIE* 2020, 225-233 (1993)
7. G. Klingenberg, Experimental diagnostics in reacting muzzle flows, gun propulsion technology. Stiefel, L. (ed.) *Progress in Astronautics and Aeronautics* 109 183-259 (1988).
8. G. Klingenberg, J.M. Heimerl, Gun muzzle blast and flash. *Progress in Astronautics and Aeronautics* 139. 3-14 (1992).
9. W.F. Bagby, Collection of detonation signatures and characterization of spectral signatures, M.S. thesis, Air Force Institute of Technology, AFIT/GAP/ENP/01M-01 (2001).
10. A.N. Dills, K.C. Gross, G.P. Perram, Detonation discrimination techniques using a fourier transform infrared spectrometer system and a near-infrared focal plane array, *Proceedings of the SPIE*, 5075, 208-216 (2003).
11. A.N. Dills, G.P. Perram, S.C. Gustafson, Detonation discrimination and feature saliency using a near-infrared focal plane array and a visible CCD camera. *Proceedings of the SPIE* 5431 77-86 (2004).
12. K.C. Gross, A.N. Dills, G.P. Perram, Phenomenology of Exploding Ordnance Using Spectrally and Temporally Resolved Infrared Emissions, *Proceedings of the SPIE*. 5075, 217-227 (2003).
13. J.A. Orson, Collection of detonation signatures and characterization of spectral signatures, M.S. thesis, Air Force Institute of Technology, AFIT/GSO/ENP/00M-01 (2000).
14. J.A. Orson, W.F. Bagby, G.P. Perram, Infrared signatures from bomb detonations. *Infrared Physics and Technologies* 44 101-107 (2003).
15. P.W. Cooper, *Explosives Engineering* (Wiley-VCH, 1996).
16. A.L. Kuhl, J.-C. Leyer, A.A. Borisov, W.A. Sirignano, W.A. editors. *Dynamic Aspects of Detonations*. Progress in Astronautics and Aeronautics Series. (AIAA 1993).
17. A.L. Kuhl, J.-C. Leyer, A.A. Borisov, W.A. Sirignano, W.A. editors.. *Dynamic Aspects of Explosion Phenomena* Progress in Astronautics and Aeronautics Series. (AIAA 1993)
18. C.L. Mader, *Explosives and Propellants* (CRC Press 1998).
19. R.E. Setchell, Optical studies of chemical energy release during shock initiation of granular explosives, *Progress in Astronautics and Aeronautics* 106, 607-628 (1986)
20. W.G. Von Holle, C.M. Traver, Temperature measurements of shocked explosives by time resolved infrared radiometry—a new technique to measure shock-induced reaction, Seventh Symposium on Detonation, Annapolis, Maryland, 993-1003.

21. S.A. Clough, M.W. Shephard, E.J. Mlawer, J.S. Delamere, M.J. Iacono, K. Cady-Pereira, S. Boukabara, P.D. Brown, Atmospheric radiative transfer modeling: A summary of the AER codes, *Journal of Quantitative Spectroscopy and Radiative Transfer* **91**(2): 233-244 (2005).
22. A.N. Dills, Temporal and spectral classification of battlespace detonations, Ph.D. dissertation, Air Force Institute of Technology, AFIT/DS/ENP/04-2 (2005).

CHIRPED-PULSE FOURIER TRANSFORM MICROWAVE SPECTROSCOPY: A NEW TECHNIQUE FOR RAPID IDENTIFICATION OF CHEMICAL AGENTS

JASON J. PAJSKI

*Department of Chemistry, University of Virginia, McCormick Rd,
Charlottesville, Virginia 22904-4319, USA
jjp7h@virginia.edu*

MATTHEW D. LOGAN

*Department of Chemistry, University of Virginia, McCormick Rd,
Charlottesville, Virginia 22904-4319, USA
mdl4b@virginia.edu*

KEVIN O. DOUGLASS

*Department of Chemistry, University of Virginia, McCormick Rd,
Charlottesville, Virginia 22904-4319, USA
kodouglass@gmail.com*

GORDON G. BROWN

*Department of Chemistry, University of Virginia, McCormick Rd,
Charlottesville, Virginia 22904-4319, USA
ggb4r@virginia.edu*

BRIAN C. DIAN*

*Department of Chemistry, University of Virginia, McCormick Rd,
Charlottesville, Virginia 22904-4319, USA
dianb@purdue.edu*

BROOKS H. PATE[†]

*Department of Chemistry, University of Virginia, McCormick Rd,
Charlottesville, Virginia 22904-4319, USA
brookspate@virginia.edu*

RICHARD D. SUENRAM

*Department of Chemistry, University of Virginia, McCormick Rd,
Charlottesville, Virginia 22904-4319, USA
suenram@earthlink.net*

[†]To whom correspondence should be addressed.

We have developed a new broadband Chirped-Pulse Fourier Transform Microwave (CP-FTMW) spectrometer that allows the microwave spectrum in the 7.5-18.5 GHz range to be measured in a single data event. This technique produces a pure rotational spectrum that can be used for unambiguous identification of any species having a permanent electric dipole moment. CP-FTMW is a gas phase technique that is ideally suited for the detection of airborne chemical warfare agents (CWA) which must be detected in trace amounts (< 10 ppm in air). The high resolution of the technique allows the identification of complex mixtures without the need for a preliminary separation step, such as gas chromatography, which significantly reduces analysis time. The technique is "blind" to major atmospheric components (N₂, O₂, CO₂, H₂O) as they either do not possess a permanent dipole moment or do not absorb in the range of the spectrometer, thereby eliminating large background signals. In this paper we will present preliminary results that are focused on early detection of airborne CWA, including acquisition time, sensitivity limits, and sample handling requirements for several of these species.

Keywords: microwave spectroscopy; chemical warfare agent detection; spectral sensing; molecular fingerprint.

1. Introduction

Detection of airborne chemical warfare agents (CWA's) places stringent requirements on the techniques used for positive identification of these species. The high toxicity of these agents requires highly sensitive methodologies that can detect trace species at $\mu\text{mol/mol}$ (part per million, ppm) levels or lower to minimize human casualties.¹ In addition, the technique must be able to discriminate against atmospheric background species of reduced toxicity, such as fuels and pesticides, and minimize false positives and negative identification of target species. Furthermore, sensing devices that can maintain remote surveillance as stand-alone devices that are operational for weeks or months before routine service is needed are desirable for early warning systems. Ideally these remote devices would continuously monitor for airborne contaminants with sampling times of minutes or less to allow for immediate transmission to first responders. Traditional techniques such as GC/MS may excel in sensitivity and selectivity, but are hindered by other operational aspects such as long analysis times and maintenance intensive operation.² More recently microwave based techniques have been investigated which have distinct advantages over more conventional methods.³⁻¹²

Microwave technology is robust. Highly stable solid state devices are commercially available that operate from 1-18 GHz. These devices can be locked to atomic clocks or global positioning systems to achieve frequency stability that far exceeds the natural linewidth of the measured rotational transitions. The result is a frequency generation system that requires little to no maintenance over long time periods (months to years), a feature that is currently impossible to achieve with broadly tunable laser systems. These stand-alone microwave systems, which can be remotely controlled, are technically feasible and can be operated with push-button control.¹³⁻¹⁶

Rotational spectroscopy in the microwave frequency region provides precise determination of molecular structure. The power of the technique lies in its ability to discriminate between distinct molecular shapes, allowing easy differentiation between isomeric species of the same mass. The introduction of high finesse Fabry-Perot

cavities^{17,18} and phase sensitive detection¹⁹ pushed the detection limits of the technique to nmol/mol (part per billion, ppb) levels or lower. The spectrometer design by Balle and Flygare in 1981 (referred to as the Balle-Flygare design, BF) incorporated a pulsed molecular beam source that improved the sensitivity of the technique by cooling the sample in a free jet source.¹⁸ The BF spectrometer has been applied to the detection of exotic molecular species such as large biomolecules²⁰, weakly bound clusters²¹, and radicals.²² Furthermore, it is a general method, requiring only that the species of interest have a permanent dipole moment for detection.^{23,24}

Despite the general sensing capabilities and high sensitivity of the BF design, the spectrometer suffers from a technical limitation, restricting its use as a high-throughput analytical sensing device. Although the Fabry-Perot cavity acts as a passive amplifier for the incident microwave field and the detected molecular emission, it limits the active bandwidth of the spectrometer by the following way:

$$\Delta f = \frac{f}{Q} \quad (1)$$

Where f is the frequency of the incident microwave radiation and Q is the dimensionless cavity quality factor. Thus for a typical cavity Q of 20,000 at 10 GHz the operational bandwidth is on the order of 0.5 MHz. Therefore a typical rotational spectrum that covers 7.5-18.5 GHz must be stepped in 500 kHz step sizes over that spectral range. The result is a recording device that must take 22,000 steps to record an 11 GHz spectral region leading to data acquisition times that can take upwards of 14 hours. This analysis time can be reduced to minutes if "spot checks" are performed by tuning the cavity only to the rotational transition of a known species, however, in this mode of operation, only the species of interest will be detected, i.e., molecular species with transitions outside the spectral window being monitored will not be detected.

We have developed a spectrometer that significantly reduces the data acquisition time over the BF design.^{15,16} By removing the Fabry-Perot cavity (thereby removing the passive bandwidth filter) our operational bandwidth is limited by the digital electronics used in frequency generation and detection, as well as the bandwidth of our amplification sources. In the past few years, technical advances in microwave technology have developed digital frequency generation and detection devices that span into the microwave regime. Our ultrabroadband spectrometer design can record 11 GHz of rotational spectrum in a single shot. In this paper we demonstrate the analytical capabilities of the ultrabroadband spectrometer, and the inherent strengths towards using rotational spectroscopy for CWA detection and identification.

2. Experimental

The details of spectrometer operation and performance are given in separate publications and only relevant details will be given here.^{15,16} The chirped-pulse, Fourier transform

microwave (CP-FTMW) spectrometer consists of three basic components: 1) chirped microwave pulse generation, 2) molecular beam chamber (which includes the horn antennas used to couple the microwave radiation), and 3) free induction decay (FID) detection. These three components are generalized in Figure 1. The spectrometer design has no moving parts or frequency components that require active tuning (e.g. synthesizers). The result is a spectrometer with a push-button operation that can be operated remotely for weeks before routine maintenance is performed.

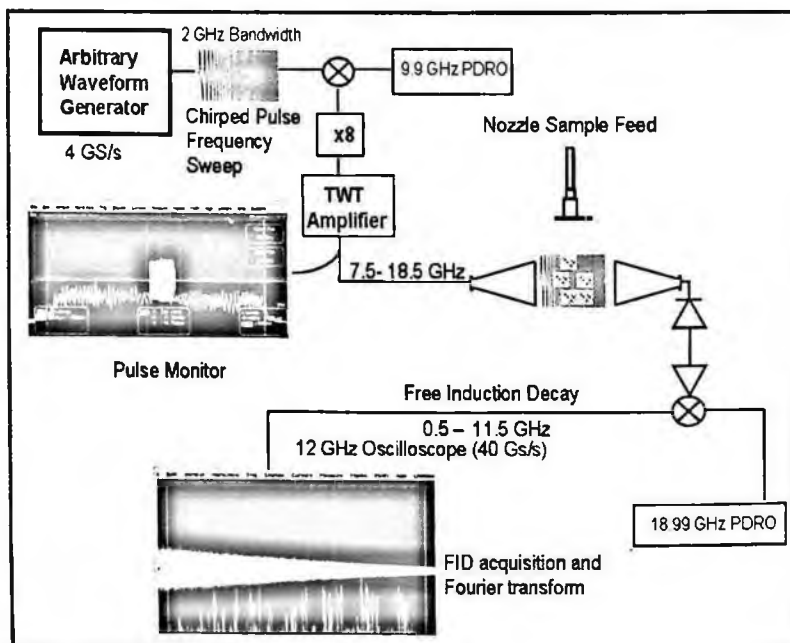


Figure 1: Diagram of the chirped-pulse, Fourier transform microwave spectrometer showing the generation of the the linear frequency sweep which is bandwidth multiplied and amplified. The inset shows the pulse in both time and frequency space. The pulse is broadcast into the chamber through a double-ridge microwave horn and the free induction decay (FID) is collected with another microwave horn . The FID is downconverted to the 0.5 - 11.5 GHz range before being recorded by the oscilloscope.

2.1 Chirped microwave pulse generation

To perform broadband FTMW spectroscopy we require a microwave source that can produce phase-locked linear frequency sweeps over an 11 GHz frequency range in times ranging from 100 ns to 1 μ s (sweep rates of 10^{10} – 10^{11} MHz/s). The short sweep durations are required so that the sample is polarized on a time scale faster than the pure dephasing of the rotational free induction decay (FID). Traditional microwave synthesized sweepers are typically limited to sweep rates of about 10^6 MHz/s. We have developed a microwave source based on a 4 Gs/s arbitrary waveform generator that can produce the required excitation pulses. A schematic of the microwave source is shown in Figure 1.

The base linear sweep chirped pulse is created using the arbitrary waveform generator (AWG, Tektronix AWG 710B). However, the frequency sweep from this source only produces frequencies from DC to 2 GHz. A microwave circuit is used to convert the arbitrary waveform generator sweep in two ways: expand the sweep bandwidth by a factor of eight and to up convert the frequencies into the 7.5 – 18.5 GHz frequency range. Bandwidth extension is performed using active frequency multipliers. These devices linearly increase the bandwidth when a single frequency linear sweep is supplied to the input. Frequency up conversion is accomplished using a phase-locked dielectric resonator oscillator (PDRO) source and a broadband microwave mixer. The microwave circuit is designed to use only a single external microwave frequency (the 9.9 GHz PDRO) and to permit simple filtering of the output of the frequency multipliers to improve the spectral purity of the final chirped pulse output. The details of this design are described in a separate publication.¹⁵

All frequency sources in the experiment are phase locked to a Rb-disciplined crystal oscillator operating at 10 MHz (Stanford Research Systems FS 725). The specified frequency accuracy of the 10 MHz reference is $\pm 5 \times 10^{-10}$ and ultimately determines the accuracy of our microwave frequency measurements. An important feature of this reference source is the low phase noise (< -130 dBc/Hz at 10 Hz). The ability to average FID signal in the time domain requires high phase stability of the microwave source. In practice, we can average 10,000 successive FID measurements in the time domain while retaining 90% of the single measurement signal amplitude.

Chirped pulse amplification is achieved using a pulsed traveling wave tube amplifier (TWTA, Amplifier Research 1000TP8G18) with peak power of 2 kW (7 – 18 GHz). The final pulse is shown as an inset in Figure 1. The spurious signals in the pulses we create using this technique are at least 20 dB lower in power than the instantaneous sweep frequency across the full 11 GHz range of the pulse.

2.2 Molecular beam chamber

The molecular beam chamber contains the pulsed valve source and two horn antennas to broadcast the linear sweep polarizing pulse and to receive the rotational FID. The chamber consists of a 13.25" conflat flange six-way cross and a tee. Two 10" diffusion pumps (Varian VHS 10), backed by a mechanical pump – roots blower combination (Edwards E2M40 and E2H250) are used to maintain vacuum in the chamber. The pulsed valve (General Valve Series 9) is mounted to the top flange of the six way cross and its height relative to the microwave interaction region can be manually adjusted. Samples are approximately 0.1% in first run neon (70% Ne/30% He) in a gas cylinder unless otherwise stated. The samples are expanded into the vacuum chamber with 1 – 2 bar backing pressure, which causes the molecules to cool in the resulting supersonic expansion to approximately 1 K.

The high peak power polarizing pulse is coupled into the chamber using double-ridge waveguide with a Kapton window to maintain vacuum. The polarizing pulse is broadcast into the chamber using a double-ridge standard gain horn (20 dB gain over isotropic). A

second double ridge rectangular horn is used to collect the rotational FID signal. The microwave power collected by the receiver horn is coupled out of the molecular beam chamber using a SMA coupler and flexible microwave power cables. All interior surfaces of the molecular beam chamber are covered by microwave foam absorber that is secured to metal foil backing. The microwave absorber is needed to prevent long lived resonances in the molecular beam chamber that produce spurious microwave contributions to the collected microwave signal.

2.3 FID detection

The first two components in the detection assembly, a high-power PIN diode and a single-pole switch are required to protect later circuit elements from the high-power microwave polarizing pulse. The leakage peak power from the diode is still too high for the low noise amplifier used to amplify the rotational FID, so the microwave switch is placed after the diode. This switch is controlled by the same pulse that activates the TWTA pulse output.

The rotational FID signal that is collected in the spectrometer chamber using a horn antenna is amplified by a high gain (38 dB), broadband (8-18 GHz), low noise amplifier (noise figure 1.4 dB, Miteq AMF-5F-0800-1800-4-10P-1). The amplified FID signal is subsequently down converted to the 500 MHz – 11500 MHz frequency range using an ultra broadband triple balanced mixer. The local oscillator in this stage is a single frequency PDRO source at 18.990 MHz. The down converted FID signal passes through a low pass filter (12.2 GHz cut off) that removes the local oscillator signal that leaks through the mixer. A DC block (250 MHz – 18.5 GHz pass band) follows the low pass filter to remove the low frequency (1/f) noise prior to digitization.

The FID is digitized using a digital oscilloscope (Tektronix TDS6124C) operating with a 40 GS/s sample rate. The hardware bandwidth limit (-3 dB) of the oscilloscope is 12 GHz. The record length of the FID can be set independently. Signal averaging long record lengths is computationally intensive and, therefore, the repetition rate of the signal collection depends on chosen gate length. For a 4 μ s gate the spectrometer can acquire and signal average the signal at 10 Hz. For higher spectral resolution we use a 20 μ s gate (800,000 points) which reduces the experiment repetition rate to about 3.5 Hz.

3. Results and Discussion

3.1 Pure rotational spectra

We have chosen four molecules that are either simulants or pre-cursors of CWA's in order to characterize the capabilities of the CP-FTMW spectrometer for the detection of airborne CWA's. The rotational spectrum of each compound serves as a basis for molecular recognition allowing positive identification of the species of interest. The resulting spectrum contains both frequency and intensity information that can be used to discriminate against other CWA or background pollutants.

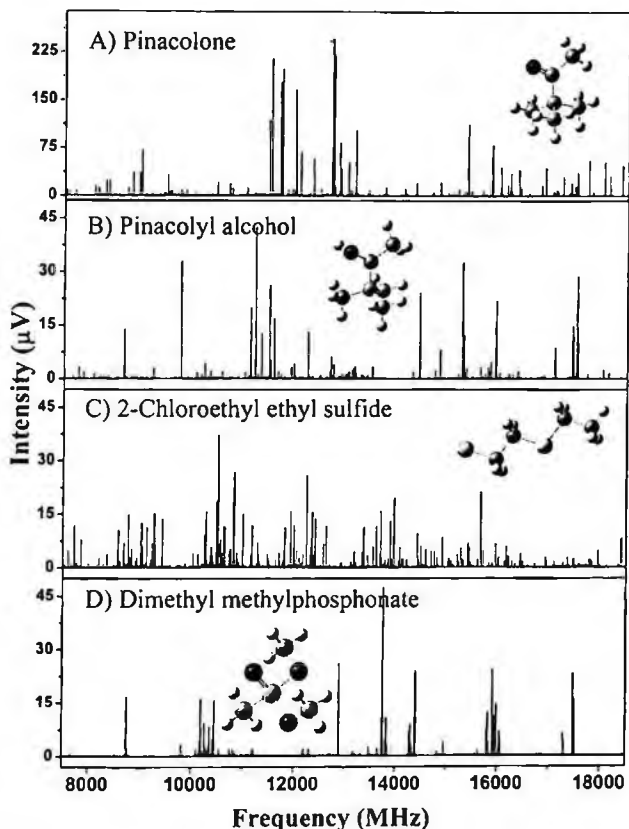


Figure 2: Pure rotational spectra for the chemical weapon simulants pinacolone (A), pinacolyl alcohol (B), 2-chloroethyl ethyl sulfide (C), and dimethyl methylphosphonate (D).

Table 1: Rotational constants and electric dipole moments of the chemical weapon simulants pinacolone, pinacolyl alcohol, 2-chloroethyl ethyl sulfide, and dimethyl methylphosphonate.

	A (MHz)	B (MHz)	C (MHz)	μ_a (D)	μ_b (D)	μ_c (D)
Pinacolone ^a	3210	2333	1935	-1.1	2.4	0.0
Pinacolyl alcohol ^b	3095	2251	1874	0.9	1.2	-0.7
2-Chloroethyl ethyl sulfide ^c						
Trans-gauche-trans conformer	6762.821	798.253	747.993	1.8 ^d	1.1 ^d	0.2 ^d
Gauche-gauche-trans conformer	4550.833	925.748	835.201	1.6 ^d	1.1 ^d	0.7 ^d
Dimethyl methylphosphonate ^e	2828.753	1972.359	1614.268			
Conformer I ^f				0.5	0.7	2.2
Conformer II ^f				1.5	1.6	0.0
Conformer III ^f				1.5	2.7	3.5

^aCalculated at MP2/cc-pVTZ

^bCalculated at MP2/6-311++G(d,p)

^cFrom reference 27

^dCalculated at MP2/6-311++G**

^eFrom reference 28

^fCalculated at HF 6-31G*

The rotational spectra of the four CWA simulants are shown in Figure 2. The upper two panels show the rotational spectrum of pinacolone (PC) and pinacolyl alcohol (PCA), respectively and are precursors to the nerve agent soman. The rotational constants, distortion terms, and dipole moment components (determined by *ab initio* calculations²⁵) have been reported elsewhere.²⁶ The rotational constants and projection of the permanent dipole moment onto the inertial axes are reported in Table 1.

Each of the spectra shown in Figure 2 was recorded by signal averaging 10,000 time domain spectra and performing a fast Fourier transform to recover the frequency domain spectrum. The acquisition gate of the Fourier transform function digitized 4 μ s of the FID, setting the resolution of the frequency domain spectrum. We also employ a digital filter to the time domain signal prior to fast Fourier transform for side lobe suppression. We have chosen a Kaiser-Bessel type windowing function (69 dB side-lobe suppression) as a compromise between increased resolution and decrease in signal intensity. The frequency resolution (FWHM) is approximately 420 kHz using a 4 μ s gate. These conditions were chosen as a trade-off between acquisition speed and resolution. The time to acquire 10,000 time domain spectra at 4 μ s resolution is approximately 22 min. resulting in 10,000 S/N for the strongest transition in the PC spectrum (Fig. 2A).

Although the molecular structure of the two pinacolone and pinacolyl alcohol differs only by the addition of two hydrogen atoms to convert the ketone (PC) to the alcohol (PCA), the qualitative appearance of the rotational spectra is quite different. The addition of hydrogen atoms to PC shifts the center of mass of the molecule serving to reorient the inertia frame of PCA relative to PC. Although this reorientation is small, it effectively shifts the rotational energy levels and, therefore, the measured transition frequencies. The addition of the hydrogen atoms serves to reorient the direction of the permanent dipole moment of PCA relative to PC. PC belongs (approximately) to the low symmetry C_s point group having a plane of reflection that bisects the t-butyl group along the C-C(=O)-CH₃ tail group. The result is that there is no projection of the permanent dipole moment along the c-inertial axis. The addition of the hydrogen atom in PCA breaks this symmetry resulting in a projection of the permanent dipole moment along the c-inertial axis (μ_c). Since the intensity of an allowed rotational transition is proportional to the dipole moment component,^{23,24} new transitions appear in the PCA spectrum that were absent in the PC spectrum.

The third panel of Figure 2 shows the rotational spectrum of 2-chloroethyl ethyl sulfide (2CEES) a mustard gas simulant. The rotational constants and inertial projection of the permanent dipole moment have been previously reported and are reproduced in Table 1.²⁷ The rotational spectrum of 2CEES is highly congested by the addition of the chlorine atom. The 2CEES spectrum is fragmented by the appearance of two nuclear isotopes (³⁵Cl and ³⁷Cl) and their associated nuclear hyperfine splitting ($I=3/2$).²⁴ In addition, the 2CEES molecule has two conformational isomers (same atom connectivity) with significant population yielding two sets of independent rotational spectra. Finally, the bottom panel of Figure 2 shows the rotational spectrum of dimethyl methylphosphonate (DMMP). The rotational constant and projections of the permanent

dipole moment on the inertial axes have been reported previously and are reproduced in Table 1.²⁸

3.2 Data acquisition

The real advantage of adopting a broadband spectrometer is demonstrated in Figure 3. The pure rotational spectrum of PC is shown after different numbers of averages. A few of the transitions attributed to PC are marked with an asterisk, the strong lines in the spectrum are due to acetone seeded into the sample reservoir. The first panel shows the single-shot rotational spectrum recorded without any signal averaging. PC peaks are clearly discernable in the spectrum with signal-to-noise ratio of approximately 3:1 (50 ppm). For strong molecular signals, such as acetone, the signal-to-noise ratio is 100:1 (~2000 ppm). The acquisition time under single shot conditions is limited only by the digitization time of the FID, 4 μ s in this case.

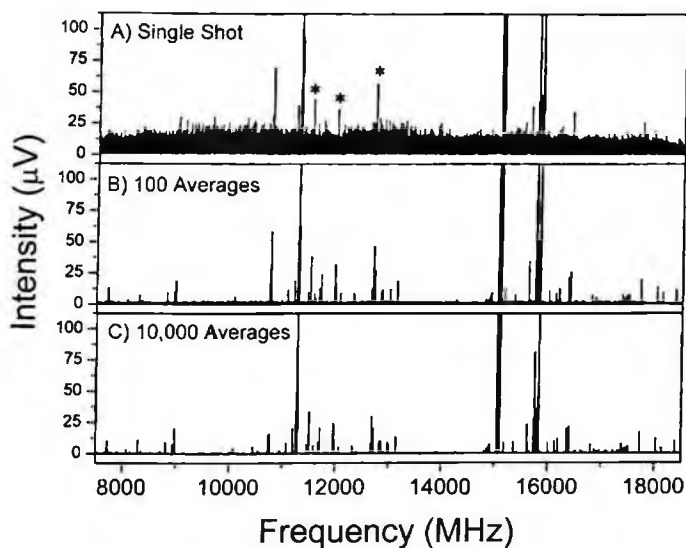


Figure 3: Signal-to-noise comparison for a single shot (A) versus 100 averages (B) and 10,000 averages (C) on the spectrum of a mixture of pinacolone and acetone. The peaks labeled with an asterisk indicate the pinacolone transitions. The remaining peaks are from acetone.

The signal-to-noise ratio can be improved by signal averaging. The middle panel of Figure 3 shows the same spectrum after 100 time domain averages. The signal-to-noise ratio has improved to better 30:1 and the molecular fingerprint of PC is now apparent in the rotational spectrum. The acquisition time is now limited by the repetition rate of the experiment (~10 Hz), which is limited by data manipulation of the FID. The spectrum in Figure 3B was acquired in 13 seconds. The bottom panel of Figure 3 shows the rotational spectrum after 10,000 averages in the time domain where the signal-to-noise ratio is about 300:1 on the PC transitions. The acquisition time is 22 minutes.

There are two important advantages that an ultrabroadband spectrometer has over the BF design. The performance characteristics of sensitivity and resolution (available data channels, set by the gate duration) can be matched to the specific need of the experiment. This versatility makes the ultrabroadband spectrometer ideal for analytical applications. By choosing the degree of signal averaging and resolution, the experimentalist can minimize the analysis time required by the spectrometer to positively identify trace chemical species. We have recently published a detailed comparison of CP-FTMW to the BF spectrometer and have found that 10,000 time domain averages of the CP-FTMW spectrometer is equivalent to a 10 shot average of the BF machine. However, the equivalent spectrum would require over 14 hours to record using the BF design, compared to the 22 minutes required by the CP-FTMW spectrometer.^{15,16}

3.3 Sensitivity

The sensitivity limits of the CP-FTMW spectrometer have been determined using PC as a prototypical molecule. The following procedure was used to prepare successive dilutions of PC. A mixture of 2.275 mL of PC and 4.859 mL of acetone was prepared. A small amount (200 μ L) of this liquid sample was directly injected into an evacuated 5.9 L gas cylinder. The sample was diluted with a buffer gas of 30% He and 70% Ne to achieve 500 ppm PC. To achieve a 250 ppm PC sample, 5 mL of the original liquid sample was diluted with acetone to a total volume of 10 mL in a volumetric flask and directly injected into a 5.9 L gas cylinder which was then diluted to the desired concentration with the same buffer gas. Samples of 100 ppm, 50 ppm, 25 ppm, 10 ppm, 5 ppm, and 1 ppm PC were all similarly prepared by diluting the liquid samples with acetone prior to direct injection of 200 μ L into a 5.9 L gas cylinder and diluting with buffer gas to the desired concentration. This method allowed the PC to be diluted while the concentration of acetone was relatively constant for all samples.

Three rotational transitions were monitored as a function of concentration to determine the linear response and detection limit of the spectrometer. The three lines are marked with asterisks in Figure 3 and correspond to $2_{21}-1_{01}$, $2_{20}-1_{11}$, $4_{04}-3_{13}$, all b-type rotational transitions ($J_{K_aK_c}$ notation²⁴). The resulting plot of intensity vs. concentration in ppm is shown in Figure 4 for all three transitions.

As Figure 4 shows, the intensity response is linear across two orders of magnitude from 5 ppm (μ mol/mol) up to 500 ppm. The intensities shown are after 10,000 averages with a gate duration of 4 μ s. The error bars represent 95% confidence limits on the data points. This graph shows that quantitative determination of gas-phase concentrations is possible using the CP-FTMW technique down to a level of 5 ppm under the current spectrometer configuration. At concentrations above 500 ppm the response becomes non-linear. We attribute this to cluster formation in the supersonic expansion.

The detection limit of PC under the current spectrometer configuration was determined to be 5 ppm after 10,000 averages (22 min.). This sensitivity can be increased in two ways. Increasing the number of averages will increase the sensitivity by \sqrt{N} where N is the number of averages. When the repetition rate of the experiment can

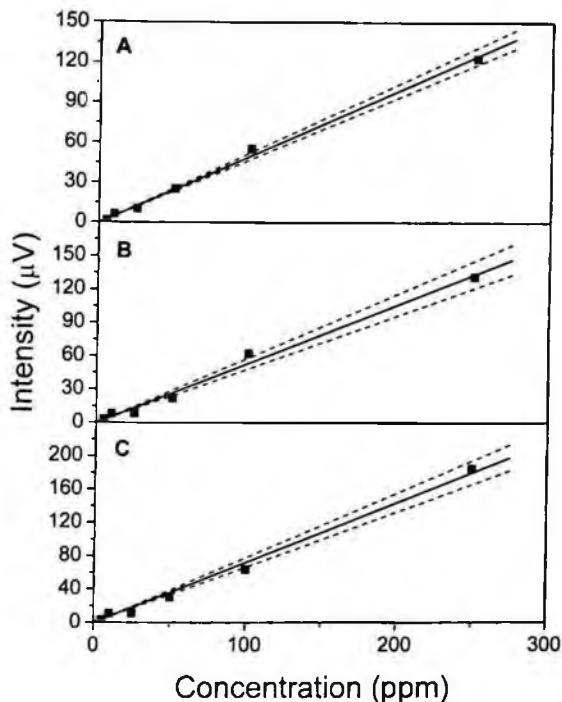


Figure 4: Graphs of signal intensity versus concentration for pinacolone in the CP-FTMW for three transitions: $2_{11}-1_{10}$ at 11503 MHz (A), $2_{20}-1_{11}$ at 11960 MHz (B), and $4_{04}-3_{13}$ at 12685 MHz (C). The solid lines represent best-fit lines with R^2 values of 0.9967 for (A), 0.9889 for (B), and 0.9942 for (C). The dashed lines represent the 95% confidence interval.

be increased as outlined above this becomes an attractive option. A second way to increase the sensitivity of the spectrometer is to pre-concentrate the sample. From the above concentration study, approximately 8 nmol of sample are required to be detected in a single shot. The rotational signal is proportional to the total number of molecules in a shot, so the number of averages (and therefore time) required to achieve a high enough S/N to detect analyte decreases linearly as a function of increasing number of molecules in each shot. Pre-concentration could be achieved by trapping the gas-phase molecules with a cold finger and releasing the adsorbed molecules by shock-heating or laser desorption techniques. In this way the same *amount* of sample detected after 22 minutes of signal averaging could be detected just over the time it takes to concentrate the sample, if released in a single shot.

3.4 Component analysis

In order to function as a CWA detector under real-world conditions, this technique must be able to discriminate between multiple species in a sample. At the current resolution of 420 kHz, there are approximately 22,000 data channels available. The four CWA

simulants used in this study had 3320 unique lines in the library, occupying only 12% of the available data channels. As data handling times decrease spectral resolution could approach 4 kHz (100 μ s gate duration), pushing the number of data channels available close to 3×10^6 . This level of resolving power competes with ultra-high resolution techniques such as FT-ICR.²⁹ It is possible to extract the spectrum of a single component from the spectrum of a complex mixture. As an example, we extracted the spectrum of 2CEES from a mixture of PC, PCA, 2CEES, and DMMP.

To obtain the complex spectrum, the following method was followed. A 1 mL syringe was used to directly inject 200 μ L each of PC, PCA, 2CEES, and DMMP into an evacuated 5.9 L gas cylinder. The cylinder was then filled with 7 atmospheres of first run He/Ne buffer gas. The analysis was performed at the vapor pressure within the gas cylinder at room temperature (no external heating). The resulting microwave signal was digitized and averaged 10,000 times (22 min acquisition time) in the time domain. The resulting frequency spectrum was recovered by fast Fourier transform of the first 4 μ s of the FID. This spectrum is shown as Figure 5.

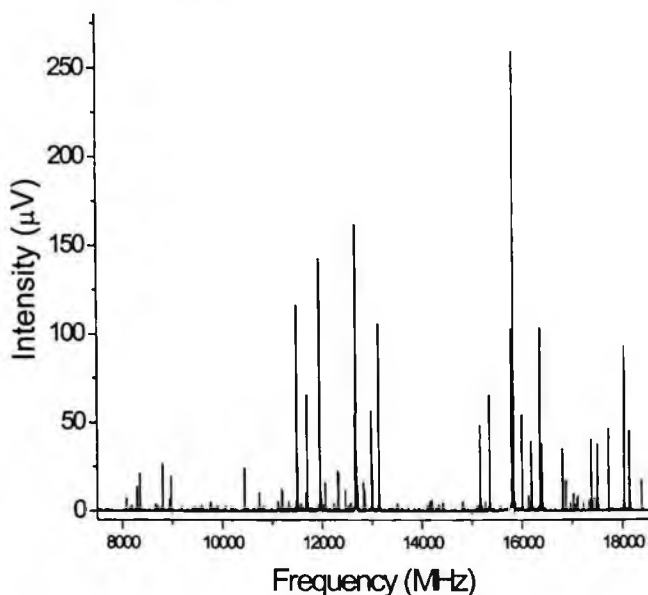


Figure 5: Pure rotational spectrum (10,000 averages) of a mixture of four chemical weapon simulants: pinacolone, pinacolyl alcohol, 2-chloroethyl ethyl sulfide, and dimethyl methylphosphonate.

Figure 6 shows the complex mixture spectrum versus the 2CEES library spectrum on an expanded scale. The more intense transitions in the library are seen in the mixture spectrum. It is evident that the resolution of this technique makes it possible to extract single components analytically from a complex mixture as is necessary for this technique to be applied to the detection of chemical weapons in real (non-laboratory) settings.

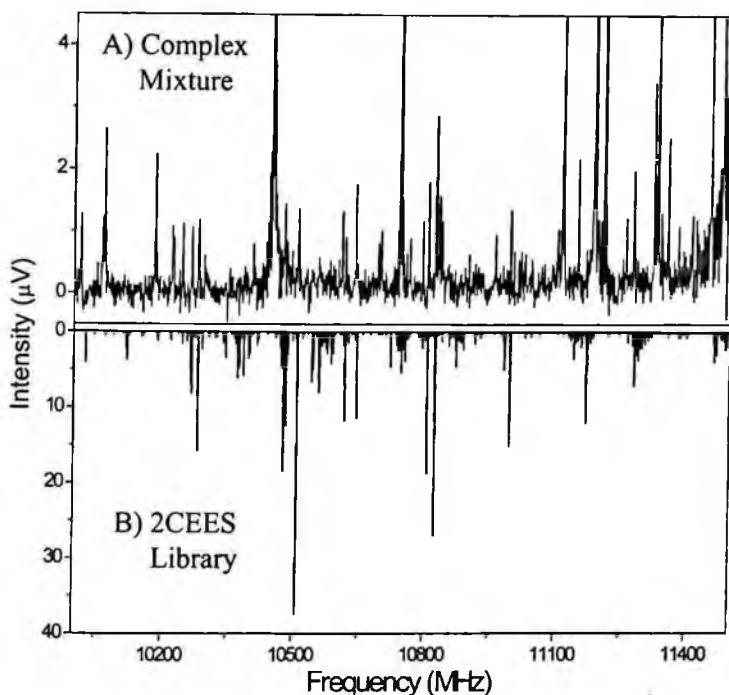


Figure 6: Comparison of the pure rotational spectrum (10,000 averages) of the CWA mixture of the four simulants combined into one sample cylinder (A) with the 2CEES library spectrum (B).

4. Summary

We have developed an ultra-broadband spectrometer that operates in the 7.5-18.5 GHz range. The CP-FTMW spectrometer can record 11 GHz of spectrum in 20 μ s plus processing time for fast Fourier transform with repetition rates of about 10 Hz, making it an ideal candidate for high throughput analytical applications. The high resolution resulting from the large number of data channels combined with the uniqueness of the rotational spectrum of a molecule makes the need for pre-separation techniques such as gas chromatography unnecessary. This discrimination permits the unique identification of isomeric species of the same mass, complementing mass-spectrometry detection methods. The fast analysis time, sensitivity, and selectivity are demonstrated on four chemical warfare agent simulants. The technique is able to positively identify all four species in a mixture without any pre-separation at the ppm level. The linear response of the spectrometer at concentration lower than 500 ppm illustrates the ability to quantitatively identify airborne species including chemical warfare agents even in a complex mixture.

References

1. T.C. Marrs, R. L. Maynard, and F.R. Sidell, *Chemical Warfare Agents: Toxicology and Treatment*. (Wiley, New York, 1996).
2. M.D. Brickhouse, W.R. Creasy, B.R. Williams, K.M. Morrissey, R.J. O'Connor, and H.D. Durst, Multiple-technique analytical characterization of a mixture containing chemical-weapons simulants from a munition, *J. Chromatogr. A*, **883**, 185-198 (2000).
3. N.D. Rezgui, J. Allen, J.G. Baker, J.F. Alder, Quantitative millimetre wave spectrometry. Part I: design and implementation of a tracked millimetre wave confocal Fabry-Perot cavity spectrometer for gas analysis, *Anal. Chim. Acta.*, **311**, 99-108 (1995).
4. N.D. Rezgui, J.G. Baker, J.F. Alder, Quantitative millimetre wave spectroscopy. Part II: Determination of working conditions in an open Fabry-Perot cavity, *Anal. Chim. Acta*, **312**, 115-125 (1995).
5. J.G. Baker, N.D. Rezgui, J.F. Alder, Quantitative millimetre wave spectroscopy. Part III: Theory of spectral detection and quantitative analysis in a millimetre wave confocal Fabry-Perot cavity spectrometer, *Anal. Chim. Acta.*, **319**, 277-290 (1996).
6. J. F. Alder, J. G. Baker, Quantitative millimetre wavelength spectrometry. Part IV. Response curves for oxygen in carbon dioxide and nitrogen at 60 GHz, *Anal. Chim. Acta.*, **367**, 245-253 (1998).
7. N. D. Rezgui, J. G. Baker, J. F. Alder, Quantitative millimetre wavelength spectrometry Part V: Observation of dispersive gas spectra with a millimeter wavelength confocal Fabry-Perot cavity spectrometer, *Anal. Chim. Acta.*, **433**, 269-279 (2001).
8. G. Thirupt, F. Benmakroha, A. Leontakianakos, J. F. Alder, Analytical microwave spectrometer employing a Gunn oscillator locked to the rotational absorption line, *J. Phys E: Sci. Instrum.*, **19**, 823-830 (1986).
9. Z. Zhu, C. Gibson, A. H. Samuel, and I. P. Matthews, Microwave cavity spectrometer for process monitoring of ethylene oxide sterilization, *Rev. Sci. Instrum.* **64**(1), 103-108 (1993).
10. Z. Zhu, I. P. Matthews, and A. H. Samuel, A microwave spectrometer with a frequency control system employing a frequency "scanning window" locked to the rotational absorption peak, *Rev. Sci. Instrum.* **66**(10), 4817-4823 (1995).
11. Z. Zhu, I. P. Matthews, and A. H. Samuel, Quantitative measurement of analyte gases in a microwave spectrometer using a dynamic sampling method, *Rev. Sci. Instrum.*, **67**(7), 2496-2501 (1996).
12. Z. Zhu, I. P. Matthews, and W. Dickinson, Specificity, accuracy, and interpretation of measurements of ethylene oxide gas concentrations during sterilization using a microwave spectrometer, *Rev. Sci. Instrum.* **68**(7), 2883-2890 (1997).
13. J. U. Grabow, W. Stahl, H. Dreizler, A multioctave coaxially oriented beam-resonator arrangement Fourier-transform microwave spectrometer, *Rev. Sci. Instrum*, **67**(12), 4072-4084 (1996).
14. R. D. Suenram, J. U. Grabow, A. Zuban, I. Leonov, A portable, pulsed-molecular-beam, Fourier-transform microwave spectrometer designed for chemical analysis, *Rev. Sci. Instrum.*, **70**(4), 2127-2135 (1999).
15. G. G. Brown, B. C. Dian, K. O. Douglass, S. M. Geyer, and B. H. Pate, A Broadband Fourier Transform Microwave Spectrometer Based on Microwave Chirped Pulse Excitation, *Rev. Sci. Instrum.*, In Preparation.
16. G. G. Brown, B. C. Dian, K. O. Douglass, S. M. Geyer, B. H. Pate, The rotational spectrum of epifluorohydrin measured by chirped-pulse Fourier transform microwave spectroscopy, *J. Mol. Spec.* In Press, available online.
17. J. Ekkers and W. H. Flygare, Pulsed microwave Fourier-transform spectrometer, *Rev. Sci. Instrum.*, **47**, 448-454 (1976).

18. T. J. Balle and W. H. Flygare, Fabry-Perot cavity pulsed Fourier-transform microwave spectrometer with a pulsed nozzle particle source, *Rev. Sci. Instrum.*, **52**, 33-45 (1981).
19. R. H. Dicke and R. H. Romer, Pulsed techniques in microwave spectroscopy, *Rev. Sci. Instrum.*, 915-928 (1955).
20. R. J. Lavrich, A. R. H. Walker, D. F. Plusquellic, I. Kleiner, R. D. Suenram, J. T. Hougen, G. T. Fraser, Conformational analysis of the jet-cooled peptide mimetic ethylacetamidoacetate from torsion-rotation spectra, *J. Chem. Phys.*, **119**(11), 5497-5504 (2003).
21. A. C. Legon, Annual Review of Physical Chemistry, **34**, 275 (1983).
22. K. Suma, Y. Sumiyoshi, and Y. Endo, The Rotational Spectrum and Structure of the HOOO Radical, *Science*, **308**, 1885 (2005).
23. C. H. Townes and A.L. Schawlow, *Microwave Spectroscopy*. (Dover, New York, 1975).
24. W. Gordy and R.L. Cook, *Microwave Molecular Spectra*. (Wiley, New York, 1970).
25. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, Jr., J. A. Montgomery, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma,
26. G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, GAUSSIAN 03, Revision C.02, Gaussian, Inc., Wallingford, CT 2004.
27. Dave Plusquellic, private communications
28. M. J. Tubergen, A. Lesarri, R. D. Suenram, A. C. Samuels, J. O. Jensen, M.W. Ellzy, J. M. Lochner, Rotational spectra, nuclear quadrupole hyperfine tensors, and conformational structures of the mustard gas simulant 2-chloroethyl ethyl sulfide, *J. Mol. Spec.*, **233**, 180-188 (2005).
29. R. D. Suenram, F. J. Lovas, D. F. Plusquellic, A. Lesarri, Y. Kawashima, J. O. Jensen, and A. C. Samuels, Fourier Transform Microwave Spectrum and *ab Initio* Study of Dimethyl Methylphosphonate, *J. Mol. Spec.*, **211**, 110-118 (2002).
30. J. M. Fu, J. M. Purcell, J. P. Quinn, T. M. Schaub, C. L. Hendrickson, R. P. Rodgers, A. G. Marshall, External electron ionization 7 T Fourier transform ion cyclotron resonance mass spectrometer for resolution and identification of volatile organic mixtures, *Rev. Sci. Instrum.* **77**(2), 025102/1-025102/9 (2006).

ENVIRONMENTAL EFFECTS INFLUENCING THE VIBRATIONAL MODES OF DNA: NANOSTRUCTURES COUPLED TO BIOMOLECULES

DINAKAR RAMADURAI

*Bioengineering Department, University of Illinois at Chicago,
851 S. Morgan Street, Chicago, Illinois 60607 USA*

TAKAYUKI YAMANAKA

*Electrical and Computer Engineering Department, University of Illinois at Chicago,
851 S. Morgan Street, Chicago, Illinois 60607 USA*

MILANA VASUDEV

*Bioengineering Department, University of Illinois at Chicago,
851 S. Morgan Street, Chicago, Illinois 60607 USA*

YANG LI

*Electrical and Computer Engineering Department, University of Illinois at Chicago,
851 S. Morgan Street, Chicago, Illinois 60607 USA*

VISWANATH SANKAR

*Electrical and Computer Engineering Department, University of Illinois at Chicago,
851 S. Morgan Street, Chicago, Illinois 60607 USA*

MITRA DUTTA

*Electrical and Computer Engineering, and Physics Departments, University of Illinois at Chicago,
851 S. Morgan Street, Chicago, Illinois 60607 USA*

MICHAEL A. STROSCIO

*Electrical and Computer Engineering, BioEngineering, and Physics Departments,
University of Illinois at Chicago, 851 S. Morgan Street, Chicago, Illinois 60607 USA*

TIJANA RAJH

Chemistry Division, Argonne National Laboratory, Argonne, IL 60439

ZORAN SAPONJIC

Chemistry Division, Argonne National Laboratory, Argonne, IL 60439

SONG XU

Molecular Imaging, Inc., 24605 Helena Drive, Brownstown, MI 48183

The interactions of charges in DNA with the vibrational modes in DNA depend on the spectra of these vibrational modes. Using (a) the Su-Schrieffer-Heeger (SSH) Hamiltonian approach, (b) integrated structures of DNA and manmade nanostructures, and (c) gel electrophoresis techniques,¹ the interaction between charges in DNA and the vibrational modes of DNA are investigated. As is well-known, DNA has a rich spectrum of modes in the THz spectral regime. The use of manmade nanostructures integrated with DNA facilitates the engineering of nanoscale systems useful in studying the role of environmental effects on the vibrational modes of DNA as well as the interaction of these modes with charge carriers in DNA. Among the DNA-based structures considered in this account are: B-DNA and Z-DNA strands related by a conformational change; and DNA molecules bound on one terminal to indirect bandgap semiconductor quantum dots. Gel electrophoresis is used as a tool for the analysis of carrier interactions in novel integrated DNA-manmade-nanostructure complexes, and models based on the SSH Hamiltonian² are employed as a means of analyzing the interactions between the vibrational modes of DNA and charge carriers in DNA.³⁻⁴

Keywords: THz phonons; DNA: polarons; polaron-phonon scattering; quantum wire; quantum dots; polaron trapping; polaron detrapping; optical excitation.

1. Introduction and Background

The research reported in this account deals with environmental effects influencing the vibrational modes of DNA and relates these effects to underlying THz-frequency vibrational modes in DNA as well as their interactions with charge carriers in DNA. To provide a unifying framework for considering environmental effects on the vibrational spectrum of DNA, this account emphasizes the understanding of polaron transport phenomena in DNA in terms of polaron emission and absorption of THz vibrational modes in DNA. In previous works, polaron transport is studied using theoretical calculations of the effects of THz-frequency phonon emission⁵⁻¹⁰ by propagating carriers, and THz-frequency phonon absorption by propagating and trapped carriers. Moreover, previous investigators have considered the role of THz-phonon-mediated charge trapping and detrapping effects near guanine-rich regions of the DNA.^{1,2,5-7} Recently, THz spectra have been determined over a range of frequencies,¹¹ and these observations facilitate the analyses presented in this account. In considering environmental effects on the phonon spectra in DNA-nanostructure complexes, the Su-Schrieffer-Heeger Hamiltonian will serve as a basis for treating the phonon-carrier coupling through an optical deformation potential.¹² In related work, the environmental effect of surrounding the acoustic modes of a quantum dot with another elastic material have been considered previously,¹³ and substantial shifts in the phonon frequencies were predicted. In the case of DNA, the interaction of polarons with the phonon modes in DNA have been modeled theoretically and compared with results of gel electrophoresis analyses of DNA fragments resulting when DNA-quantum-dot complexes are exposed to UV radiation.¹⁴ Moreover, DNA coupled to both quantum dots as well as acceptor and donor molecules have revealed similar DNA cleaving phenomena in electrolytic environments.¹⁶⁻¹⁹ With the growing body of information in the THz spectra of DNA,^{11,20,21} an attempt to understand the role of environmental influences on these spectra becomes feasible. In this account, theoretical results on the phonon-mediated charge transport are reported and gel electrophoresis studies of B-DNA and Z-DNA are reported for the case where the

specific B-DNA and Z-DNA molecules are related by a conformational change that is caused by changing the electrolytic properties of the electrolytic environment of the DNA. In studying these environmental effects on DNA linked to manmade nanostructures, it is essential to consider such effects on these manmade nanostructures as well. Accordingly, this account also presents current-voltage characteristics of TiO₂ quantum dots assembled on Au substrates as measured using current-sensing atomic force microscopy; these results are compared with available models for the current-voltage curves for such structures.²² In particular, current-voltage measurements for TiO₂-nanocrystal layers on Au substrates are analyzed theoretically²² in order to better understand the electrical properties of TiO₂ quantum dots. The study of the role of environmental effects on semiconductor quantum dots bound to biomolecules is further motivated²³⁻³⁰ by a range of studies of semiconductor quantum dots including: the chemical self-assembly of quantum dots into networks using conducting biomolecules as linking structures; the theoretical study of the mechanical, electrical, and optical properties of quantum dots; and in related disciplines, to the use of luminescent quantum dots as biotags of biological structures including cellular and subcellular structures. Similarly, the recent prediction of phonon bottleneck effects in short carbon nanotubes,³¹⁻³² implies that such phonon bottleneck effects may occur in short DNA molecules due to the boundary conditions on the phonons at the ends of the finite-length DNA molecules. The focus of the research presented in this account is on understanding the role of both carrier-phonon interactions on carrier transport interactions in DNA and related environmental effects. As has been discussed in the literature, phenomena such as base pairing,^{33,34} structural fluctuations,³⁵ π -bonding,³⁶ mode of carrier excitation,³⁷ and effective bandgap variations,³⁸⁻³⁹ may play important role in carrier transport in biomolecules.

2. Energetics of Hole Injection from Semiconductor Quantum Dots to DNA

In a number of studies of carrier injection and transport in DNA, optical excitation of TiO₂ quantum dots has been used to inject carriers into DNA.^{1-4,6,7,9,14,16,19} In many of these cases, the subsequent transport of charge in DNA wires has been studied using gel electrophoresis.^{1,6,9,14,16,19} In this account, theoretical results on the phonon-mediated charge transport are reported and gel electrophoresis studied of B-DNA and Z-DNA are reported for the case where the specific B-DNA and Z-DNA molecules are related by a conformational change that is caused by changing the electrolytic properties of the electrolytic environment of the DNA.

In these studies, indirect-bandgap TiO₂ quantum dots are used to facilitate the optoelectronic excitation of charges as a source of charges for injection into DNA. As explained in this account, the optical excitation of the TiO₂-nanocrystal-DNA complexes with UV radiation is shown to result in charge transfer into the DNA. The energetics of these transfer processes are depicted in Fig. 1 where the semiconductor conduction band energy, E_c , the semiconductor valence band energy, E_v , the DNA highest occupied molecular orbital (HOMO), and the DNA lowest unoccupied molecular orbital (LUMO)

are displayed. From the energy alignments in Fig. 1, it is clear that a hole in the valence band of TiO_2 quantum dots has an energy that is equal to an energy in the HOMO band of bases, and it follows that the HOMO energy for guanine is the closest to the vacuum energy level; hence, the HOMO energy for guanine is higher than those of the other three DNA bases, adenine (A), cytosine (C), and thymine (T). As Saito et al. have demonstrated using quantum chemistry calculations,⁴⁰ the local ionization potential of a DNA base depends on its neighboring bases in the sequence of DNA bases, and that of the ionization potentials vary depending of the length of the local sequence of DNA under consideration. These effects are illustrated in Fig. 1 along with the energy alignments discussed previously.

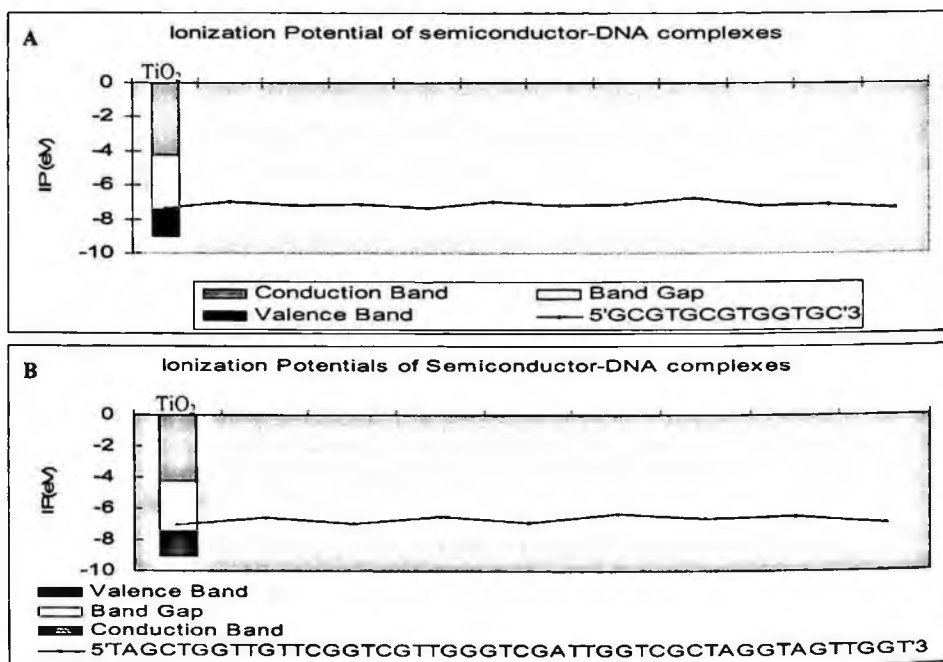


Fig. 1. Conduction and valence bands energies of TiO_2 quantum dots compared with the HOMO energy for two selected DNA sequences: (A) The energies of the ionization potential, IP, for sequence A are represented in terms of HOMO energies calculated for groups of two bases; and (B) the energies of the IP for sequence B are represented in terms of HOMO energies calculated for groups of five bases.⁴⁰

From the results of Fig. 1, it is clear that irradiation of TiO_2 nanocrystals with UV light with a wavelength of 365 nm will result in the production of electron-hole pairs in these TiO_2 quantum dots. Moreover, for the case of TiO_2 -nanocrystal-DNA complexes where nanocrystals of TiO_2 are bound to DNA molecules, the energies of these holes are matched with states in the HOMO band of DNA and have a scattering channel that will lead to hole injection into the HOMO band of the DNA.

3. Phonon Scattering of Carriers Injected into DNA

If these holes emit phonons, they will move to the edge of the HOMO band edge, and if they absorb phonons, they will move away from the HOMO band edge. Each hole will experience a different sequence of phonon absorption and emission events that have probabilities given by the carrier-phonon scattering rate, $S^{(e,a)}(E)$, as a function of carrier energy, E , which is given by,^{2,9}

$$\begin{aligned}
 S^{(e,a)}(E) &= \frac{D_{op}^2}{\rho} \frac{\hbar}{2\hbar\omega_{LO}} \left(n_q + \frac{1}{2} + \varepsilon \frac{1}{2} \right) \cdot 16 \cdot \left(\frac{8}{3\pi} \right)^4 \frac{1}{\sqrt{\frac{\hbar^2}{2m^*} \sqrt{\hbar\omega_{LO}} \sqrt{\frac{E}{\hbar\omega_{LO}} - \varepsilon}}} \quad (1) \\
 &= \frac{D_{op}^2}{\rho} \left(\frac{m^*}{2} \right)^{\frac{1}{2}} \frac{1}{(\hbar\omega_{LO})^{\frac{3}{2}}} \cdot 16 \cdot \left(\frac{8}{3\pi} \right)^4 \left(n_q + \frac{1}{2} + \varepsilon \frac{1}{2} \right) \left(\frac{E}{\hbar\omega_{LO}} - \varepsilon \right)^{-\frac{1}{2}}
 \end{aligned}$$

where

$$\varepsilon = \begin{cases} +1 & \text{for emission} \\ -1 & \text{for absorption} \end{cases}$$

In Eq. (1), the superscript “e” denotes emission of a phonon of angular frequency, ω_{LO} , and the superscript “a” denotes absorption of a phonon of angular frequency ω_{LO} . To

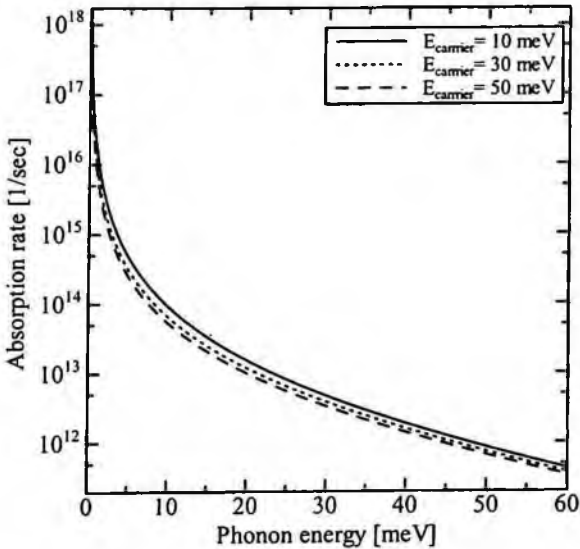


Fig. 2. Phonon absorption rate as a function of phonon energy.

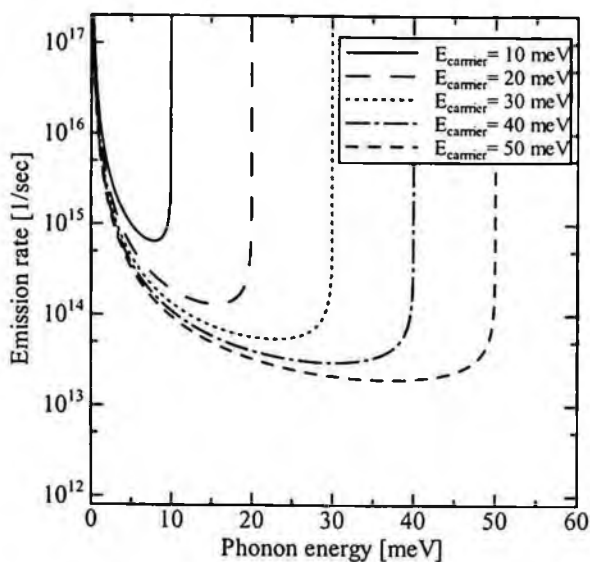


Fig. 3. Phonon emission rate as a function of phonon energy.

determine the phonon scattering rate, values of the optical deformation potential and the density are required. D_{op} is assumed to be 0.6 eV/\AA which is based on the calculated energy change as a function of the distance between bases. The masses of the DNA bases are almost the same, $4.35 \times 10^{-22} \text{ gm}$, thus, ρ is assumed to be uniform along the wire using the distance between bases $\approx 0.34 \text{ nm}$ and its value is, $\rho = 13.05 \times 10^{-16} \text{ kg/m}$. Fig. 2 and 3 are the absorption and emission rates as functions of the phonon energy for several carrier energies. As shown in Fig. 2, the absorption rates are almost the same for different carrier energies. Also, for the emission rates as shown in Fig. 3, the emission rates are of the same order before the threshold for different carrier energies, although each of the plots looks different because of the different threshold in the emission process. From Fig. 2, phonons with lower energies are dominant over those with higher energies. It is also clear from these results that phonon scattering is an important scattering mechanism in DNA. For $E_{\text{phonon}} > 50 \text{ meV}$, the emission rate becomes clearly larger than the absorption rate. Accordingly, phonon scattering processes remain a likely source of carrier energy loss with the resultant trapping at local maxima in the HOMO band. The cleavage of DNA – both linear and T-shaped DNA – has been reported when TiO_2 -nanocrystal-DNA complex are irradiated with radiation above the band gap of TiO_2 ; these experiments have relied heavily on the use of gel electrophoresis to determine the distribution of DNA mass fragments. In this account, these past results will be supplemented with results on the study of the UV irradiation of TiO_2 -nanocrystal-DNA where the DNA has both B-DNA and Z-DNA confirmations.

4. UV Irradiation of TiO₂-nanocrystal-DNA Complexes for both B- and Z-DNA

B-DNA is the most commonly existing form of DNA inside cells of living organisms; it is the classical Watson-Crick structure. Z-DNA is not as common as the other two forms of DNA --- namely A- and B- forms of DNA. Moreover, B-DNA formed out of repeating GC bases (e.g., GCGCGCGC) can form Z-DNA under low salt concentrations at physiological conditions of 370 C; this conformational change results in a change of the helicity of the DNA and is reversible. Thus, probing or studying charge transport phenomena under such transient, unstable conditions and low salt conditions can give us insights on not only their transition *in vivo* but also their switching behavior under non equilibrium conditions.

The DNA-TiO₂-nanocrystal complexes were prepared by techniques discussed by a number of authors.^{1,7,14,16,19} First, TiO₂ quantum dots with a diameter distribution of 3-5 nm (Applied Nanoworks) were suspended in a solution of water at a pH of 3-4, and these TiO₂ nanoparticles were dialyzed against 10 mM NaH₂PO₄ until a pH of 6.5 was obtained. Second, DNA oligonucleotides were synthesized with the 5' terminal carboxyl group (Synthegen Co.) and kept as a 10- μ M solution in 10 - 40 mM phosphate buffer at a pH of 6.5. Third, TiO₂-nanocrystal-DNA complexes were then formed by a condensation reaction through an intermediate N-hydroxy-succinimide ester was used to bind the carboxyl group of the oligonucleotide to the amino group of dopamine by an amide bond. [The DNA terminal carboxyl group is bound to O-N-succinimidyl-NNNN-tetramethyluronium tetrafluoroborate in the presence of N, N-diisopropylethylenamine in dimethyl formamide (DMF).] Next, the succinimidyl group is replaced with dopamine through its terminal amino group in the presence of dioxane. In the next step, the resulting solution was dialyzed thoroughly against water to remove free dopamine unbound to oligonucleotides; TiO₂ particles modified by glycidil isopropyl ether were then bound to dopamine end-labeled oligonucleotides. As has been discussed by a number of authors,^{7,16,19} when dopamine (free or bound to oligonucleotides) is added to TiO₂ colloidal solutions at 8 > pH > 2.5, the immediate development of a red color indicates instantaneous formation of a charge-transfer complex between dopamine and TiO₂.

In the gel electrophoresis experiments on B-DNA and Z-DNA, the sequences used are:

Sequence 1: 5'- COOH TCGCGCGGGCGCGCG 3' (sense)

Sequence 2: 5'- CGCGCGCCCGCGCG 3' (antisense)

Charge transport was studied through photo-oxidation of the natural B-form duplex attached to a TiO₂/Dopamine complex at the 5' terminal of the sense strand. This B-form was converted to Z-form DNA by adding a solution of high salt concentration --- 20 mM Co(NH₃)₆Cl₃ --- which stabilizes the DNA in the Z-form, The experiments were

conducted at physiological pH conditions and temperatures of 37⁰ C. Other salts such as NaCl can be used to bring about this transition; however, multivalent cations are more effective in stabilizing the Z-form than the mono-valent cations. Both the B- and the Z-forms of DNA were excited with UV light for 15 minutes and electrophoresed in a 2.5 % Agarose gel. The bands from the gel are shown in the Fig. 4.

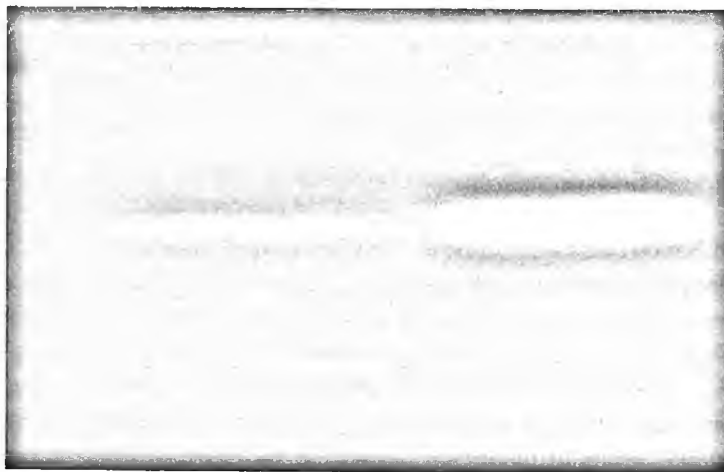


Fig. 4. Lane 1 (on the left) depicts Z-DNA illuminated with UV for 15 minutes, and lane 2 (on the right) B-DNA illuminated with UV for 15 minutes.

It appears from the gel that lane 2 which has the B-DNA - or natural form - has a brighter and wider band than the band for the Z-DNA in lane 1. A possible explanation for this result is that the DNA cleaves in the natural B-form and the cleaved strand being close in molecular weight to the original strand, is fused to form a larger band. For Z-DNA, no obvious cleavage is apparent. This may be attributed to the fact that Z-form of DNA is highly unstable and as a result tends to form coiled or supercoiled structures in order to attain B-Z equilibrium. Moreover, prolonged exposure to high energy UV during photoexcitation of the TiO₂-DNA complex, could alter the thermal stability of the Z-DNA. However, earlier studies have shown that lowering the dielectric constant of the water surrounding the DNA, with agents such as ethanol and methanol, can help stabilize the Z-form. It is evident from the gels, that the mass distributions are not identical for Z-DNA and B-DNA; it is possible that these observations are related to differences in charge transport and charge-induced cleavage of the Z-DNA and the B-DNA and these differences may relate to the stability of the different conformations of DNA. Extensions of these studies will, perhaps, facilitate the understanding of switching phenomenon in self-assembling biomolecules such as B- and Z-forms of DNA and facilitate the design of biomolecular switches based on these transitions. The facts that this transition is also reversible and that it is possible to destabilize the Z-form back into B-form with

intermediate salt concentrations, may facilitate the design of reversible switches based on conformational changes of the bases in DNA.

5. Electronic Properties of the TiO₂ Quantum Dots

To probe the electronic properties of quantum dots, TiO₂ nanocrystals assembled on Au substrates are characterized using current-sensing atomic force microscopy; these results are compared with available models for the current-voltage curves for such structures.²² In particular, current-voltage measurements for TiO₂-nanocrystal layers on Au substrates are analyzed theoretically²² in order to better understand the electrical properties of TiO₂ quantum dots. The current-voltage characteristics of TiO₂ quantum dots assembled on Au substrates are measured using current-sensing atomic force microscopy (AFM) and these results are compared with theory. The current-sensing AFM used in these experiment is a Molecular Imaging AFM. Colloidal TiO₂ nanocrystals were obtained from the Center for Nanoscale Materials with a size distribution of 3-5 nm and were deposited on Au-coated glass substrates at the Center for Nanoscale Materials (CNM) at the Argonne National Laboratory (ANL) at the ANL CNM User Facility. Briefly, the procedure for coating the substrate is as follows:

Step 1. The sample is immersed in a solution containing 0.12 M TiO₂, so that the sample is covered to half the height with solution. Following immersion for two minutes, the sample was heated in an oven for 30 minutes at 150° C.

Step 2. The second stage involved immersion of the same sample pulled slightly up from the bottom of the solution. The sample was heated in the oven again for 30 minutes, following immersion for 5 minutes.

Step 3. The final stage involved repeating step 2 with the sample pulled further up from the bottom of the solution.

This process helps create stepped monolayers of TiO₂ on a sample. A Au tip was used on the current sensing AFM at the University of Illinois at Chicago to measure the I-V characteristics of these structures. The measured conductance gap is about 6 V wide, extending from approximately - 2.5 V to + 3.5 V when the I-V characteristics are measured in the dark. The I-V characteristics of these Au-TiO₂-Au structures are modeled using the formalism of Paulson et al.²⁰ to describe a metal-quantum dot-metal system. In modeling the I-V characteristics of this system, the Fermi level in the TiO₂ is varied over a range of values and the I-V curve for each of these values is compared with the measured I-V.

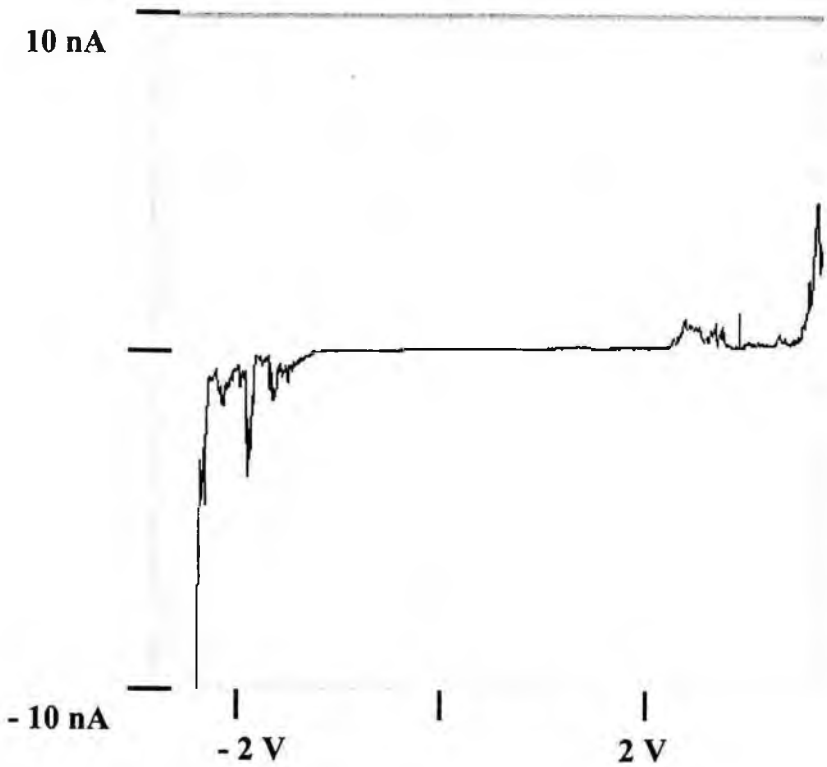


Fig. 5. Current-voltage curve for Au-TiO₂(quantum dot)-Au structure described in the text.

When the Fermi level in the TiO₂ quantum dot is taken to be at -4.2, -4.6, -5.0, -5.4, -5.8, -6.2, -6.6 eV, and -7.4 eV relative to the vacuum. The valence band edge of the TiO₂ quantum dot was taken to be at -7.4 eV and the conduction band edge was taken to be at -4.2 eV. The chemical potential of the Au is taken as -5.1 eV relative to vacuum. When the Fermi level in the quantum dot is varied in this manner the conductance gap varies over the range of 0 to 9 V. In the particular case when the Fermi level is taken to be at -6.6 eV, the conductance gap extends from approximately -3 V to +3 V and is approximately as wide as that measured and displayed in Fig. 5. The calculated conductance curves are depicted in Figs. 6-9.

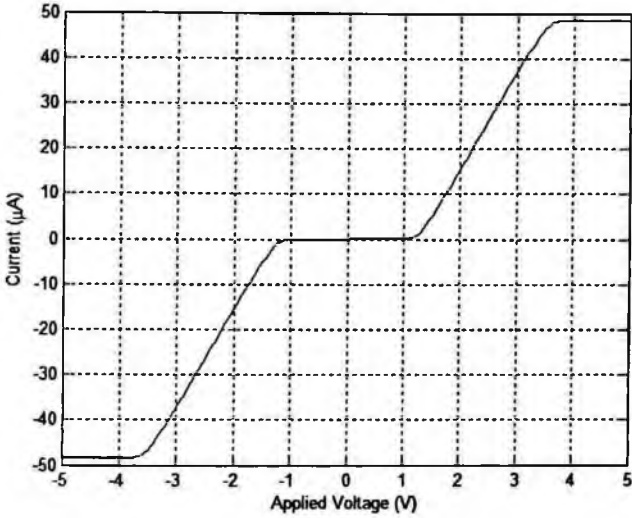


Fig. 6. Conductance curve when the Fermi energy is taken to be - 5.8 eV.

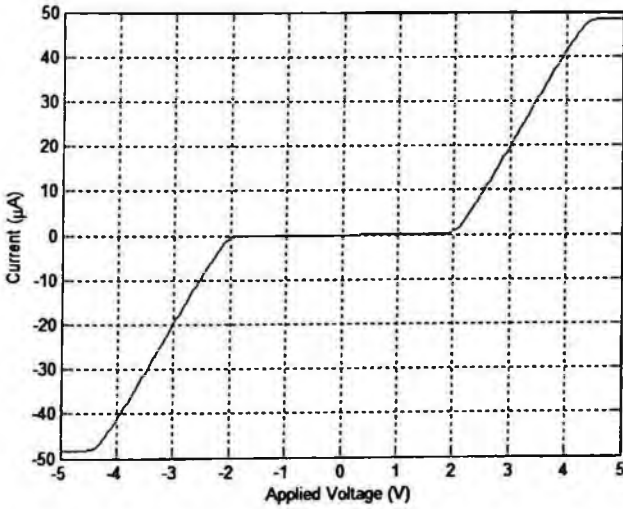


Fig. 7. Conductance curve when the Fermi energy is taken to be - 6.2 eV.

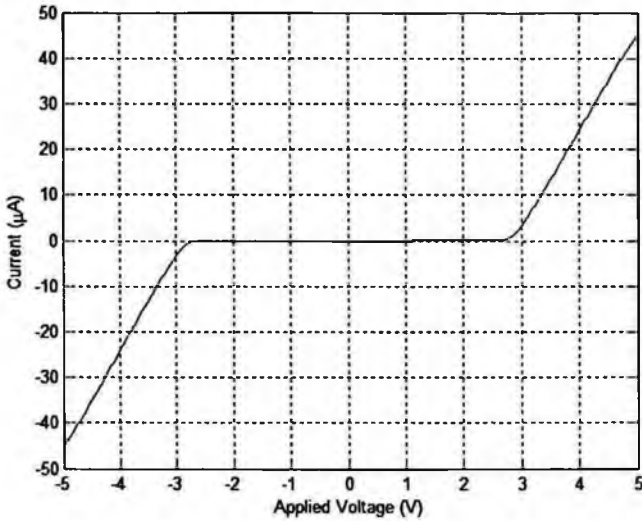


Fig. 8. Conductance curve when the Fermi energy is taken to be - 6.6 eV.

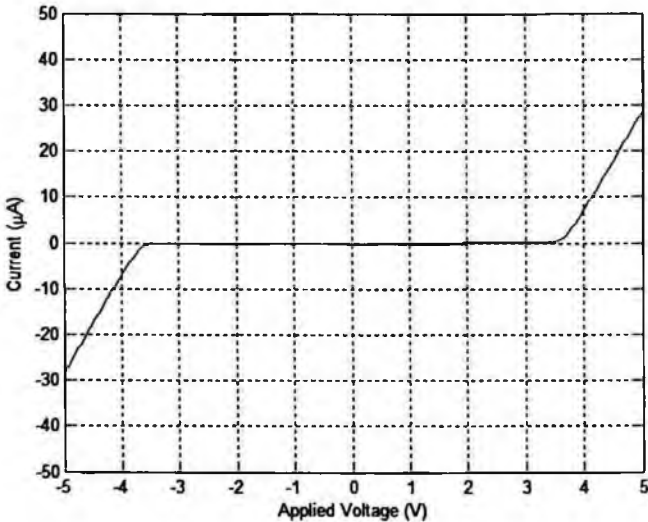


Fig. 9. Conductance curve when the Fermi energy is taken to be - 7.0 eV.

From these results, the calculated and measured values of the I-V curves are in good agreement for a Fermi level in the TiO₂ quantum dot of about - 6.6 eV. This corresponds to a slightly p-type semiconductor since the mid-point between the conduction and valence bands falls at about -5.8 eV.

6. Discussion

This account has focused on the role of environmental – or proximity – effects on the THz vibrational spectra of DNA as well as on the vibrational spectra of manmade nanostructures used in the study of the vibrational, electronic, and optical properties of DNA. Based on the results presented in this account, many factors are responsible for such effects; these include: electrolytic concentrations, materials in the proximity of the nanostructures under consideration, and the sequence of DNA bases.

7. Acknowledgments

For support of the current research under ARO grant DAAD19-03-1-0333, the authors are especially grateful to Dr. Dwight L. Woolard. The authors are extremely grateful Prof. S. Datta for useful insights on modeling I-V characteristics of molecules.

References

1. Dinakar Ramadurai, Yang Li, Takayuki Yamanaka, Dwarakanath Geerapuram, Viswanath Sankar, Milana Vasudev, Dimitri Alexson, Peng Shi, Mitra Dutta, and Michael A. Stroschio, Tijana Rajh, Zoran Saponjic, Nicholas Kotov, Zhiyong Tang, and Song Xu, Colloidal Quantum Dots as Optoelectronic Elements, in *Quantum Sensing and Nanophotonic Devices III*, edited by M. Razeghi, *Proceedings SPIE*, **6127**, 61270-131-143 (2006).
2. T. Yamanaka, M. Dutta, T. Rajh, and M. A. Stroschio, Phonon absorption and emission by holes in the HOMO bands of duplex DNA, *Proceedings of the International Conference on Hot Carriers in Semiconductors*, Springer, in press, 2006.
3. Michael A. Stroschio and Mitra Dutta, Integrated biological-semiconductor devices, *Proceedings of the IEEE*, **93**(10), 1772-1783 (2005).
4. Dimitri Alexson, Hongfeng Chen, Michael Cho, Mitra Dutta, Yang Li, Peng Shi, Amit Raichura, Dinakar Ramadurai, Shaunak Parikh, Michael A. Stroschio, and Milana Vasudev, Semiconductor nanostructures in biological applications, *Journal of Physics: Condensed Matter*, **17**, R637-R656 (2005).
5. E. Conwell, Polarons and Transport in DNA, *Topics in Current Chemistry*, **237**, 73-101 (2004).
6. S. V. Rakhmanova, and E. M. Conwell, Polarons motion in DNA, *J. Phys. Chem.*, **105**, 2056-2061 (2001).
7. Tijana Rajh, Zoran Saponjic, Jianqin Liu, Nada M. Dimitrijevic, Norbert F. Scherer, Manuel Vega-Arroyo, Peter Zapol, Larry A. Curtiss, and Marion Thumauer, Charge transfer across the nanocrystal-DNA interface: Probing DNA recognition, *Nano Letters*, **4**, 1017-1023 (2004).
8. Michael A. Stroschio and Mitra Dutta, *Phonons in Nanostructures* (Cambridge University Press, Cambridge, 2001).
9. M. A. Stroschio and M. Dutta, Biologically-inspired chemically-directed self-assembly of semiconductor quantum-dot-based systems: Phonon-hole scattering in DNA bound to DNA-quantum-dot complexes, *International Journal of High Speed Electronics*, in press, 2006.
10. M. A. Stroschio, Interaction between longitudinal-optical phonon modes of a rectangular quantum wire with charge carriers of a one-dimensional electron gas, *Phys. Rev. B*, **40**, 6428-6432 (1989).

11. D. L. Woolard, T. R. Globus, B. L. Gelmont, M. Bykhovskaia, A. C. Samuels, D. Cookmeyer, J. L. Hester, T. W. Crowe, J. O. Jensen, J. L. Jensen, and W. R. Loerop, Submillimeter-wave phonon modes in DNA macromolecules, *Phys. Rev. E*, **65**, 051903-1-11 (2002).
12. W. P. Su, J. R. Schrieffer, and A. J. Heeger, Soliton excitations in polyacetylene, *Phys. Rev. B*, **22**, 2099-2111 (1980).
13. Michael A. Stroschio, Mitra Dutta, Slavalor Rufo, and Jianyong Yang, Dispersion and damping of acoustic phonons in quantum dots, *IEEE Transactions on Nanotechnology*, **3**, 32-36 (2004).
14. Dinakar Ramadurai, Takayuki Yamanaka, Yang Li, Viswanath Sankar, Mitra Dutta, Michael A. Stroschio, Tijana Rajh, Zoran Saponjic, and Song Xu, Conductive biomolecules and their THz vibrational interactions: Key aspects of bioelectronics, in Terahertz for Military and Security Applications IV, edited by Dwight L. Woolard, R. Jennifer Hwu, Mark J. Rosker, and James O. Jensen, *Proceedings of the Society of Photo-Optical and Instrumentation Engineers*, **6212**, 62120-1-10 (2006); ISBN-0-8194-6268-3.
15. E. W. Schlag, Sheh-Yi Shen, Dah-Yen Yang, H. L. Selzle, and S. H. Lin, Charge conductivity in peptides: Dynamic simulation of a difunctional model supporting experimental data, *Proc. Natl. Acad. Sci. USA*, **97**, 1068-1072 (2002).
16. T. Paunescu, T. Rajh, G. Weiderreich, J. Maser, S. Vogt, N. Stojicevic, N. Portic, B. Lai, J. Oryhan, M. Thurnauer, and G. Woloschak, Biology of TiO₂-oligonucleotide nanocomposites, *Nature Biotechnology*, **2**, 343-346 (2003).
17. G. B. Schuster, and U. Landman, The mechanism of long-distance radical cation transport in duplex DNA: Ion gated hopping of polaron-like distortions," *Topics in Current Chemistry*, **236**, 139-161 (2004).
18. G. B. Schuster, Long-range charge transfer in DNA: Transient structural distortions and control of the distance dependence, *Acc. Chem. Res.*, **33**, 253-260 (2000).
19. T. Rajh, L. X. Chen, K. Lukas, T. Liu, M. C. Thurnauer, and D. M. Tiede, Surface Restructuring of Nanoparticles: An Efficient Route for Ligand-metal Oxide Crosstalk, *J. Phys. Chem. B*, **106**, 10543-10552 (2002).
20. Y. Cao, R. Jin, and C. A. Mirkin, DNA-modified Core Shell Ag/Au Nanoparticles, *J. Am. Chem. Soc.*, **123**, 7961-7962 (2001).
21. M. Bykhovskaia, B. Gelmont, T. Globus, D. L. Woolard, A. C., Samuels, T. H. Duong, and K. Zakrzewska, Prediction of DNA Far-IR Absorption Spectra based on Normal Mode Analysis, *Theor. Chem. Acc.*, **106**, 22-27 (2001).
22. Mangus Paulson, Ferdows Zahid, and Supriyo Datta, Resistance of a molecule, in Nanoscience, Engineering and Technology Handbook, edited by William Goddard, Donald Brenner, Sergey Lyshevski, and Gerald Iafrate, CRC Press, 2004.
23. Michael A. Stroschio, Mitra Dutta, Dinakar Ramadurai, Peng Shi, Yang Li, Milana Vasudev, Dimitri Alexson, Babak Kohanpour, Akil Sethuraman, Vikas Saini, Amit Raichura, and Jianyong Yang, Optical and electrical properties of colloidal quantum dots in electrolytic environments: Using biomolecular links in chemically-directed assembly of quantum dot networks, *J. of Computational Electronics*, **4**, 21-25, 2005.
24. Michael A. Stroschio, Mitra Dutta, Dinakar Ramadurai, Peng Shi, Yang Li, Dimitri Alexson, Babak Kohanpour, Vikas Saini, Amit Raichura, and Jianyong Yang, Electrical and optical properties of colloidal quantum dots and quantum dot networks: Role of surface states and using biomolecular links in network assembly," in *Nanoscale Devices, Materials, and Biological Systems: Fundamentals and Applications*, PV-2004-13, edited by M. Cahay, D. Lockwood, J.P. Leburton, and S. Bandyopadhyay (Electrochemical Society Press, 2005).
25. Michael A. Stroschio and Mitra Dutta, Advances in quantum dot research and technology: The path to applications in biology, *International Journal of High Speed Electronics and Systems*, **12**, 1039-1056 (2004).

26. Salvador Rufo, Michael A. Stroschio and Mitra Dutta, The influence of environmental effects on the acoustic phonon spectra in quantum-dot heterostructures, *International Journal of High Speed Electronics and Systems*, **12**, 1147-1158 (2004).
27. Dimitri Alexson, Yang Li, Dinakar Ramadurai, Peng Shi, Leena George, Lenu George, Muzna Uddin, Preetha Thomas, Salvador Rufo, Mitra Dutta, and Michael A. Stroschio, Binding of semiconductor quantum dots to cellular integrins, *IEEE Transactions on Nanotechnology*, **3**, 86-92 (2004).
28. Peng Shi, Hongfeng Chen, Michael R. Cho, and Michael A. Stroschio, Peptide-directed binding of quantum dots to cellular integrins, *IEEE Transactions on Nanobioscience*, **5**(1), 15-19 (2006).
29. Dinakar Ramadurai, Babak Kohanpour, Dimitri Alexson, Peng Shi, Akil Sethuraman, Yang Li, Vikas Saini, Mitra Dutta, and Michael A. Stroschio, Tunable optical properties of colloidal quantum dots in electrolytic environments, *IEE Proceedings in Nanobiotechnology*, **151**, 189-192 (2004).
30. Chen Chen, Mitra Dutta, and Michael A. Stroschio, Surface-optical phonon assisted transitions in quantum dots, *Journal of Applied Physics*, **96**, 2049-2054 (2004).
31. Amit Raichura, Mitra Dutta, and Michael A. Stroschio, Acoustic Phonons and Phonon Bottleneck in Single Wall Nanotubes, *J. of Computational Electronics*, **4**, 91-95, 2005.
32. Amit Raichura, Mitra Dutta, and Michael A. Stroschio, Continuum model for acoustic phonons in nanotubes: Phonon bottleneck, *Phys. Stat. Sol. (b)*, **241**, 3448-3453, (2004).
33. Kiyohiko Kawai and Tetsuro Majima, Hole transfer in DNA by monitoring the transient absorption of radical cations of organic molecules conjugated to DNA, *Topics in Current Chemistry*, **236**, 117-137 (2004).
34. Kazuhiko Nakatani and Isao Saito, Charge transport in duplex DNA containing modified nucleotide bases, *Topics in Current Chemistry*, **236**, 163-186 (2004).
35. Alessandro Troisi and Giorgio Orlandi, Hole migration in DNA: A theoretical analysis of the role of structural fluctuation, *Journal of Physical Chemistry B*, **106**, 2093-2101 (2002).
36. Yuri A. Berlin, Ogor V. Kurnikov, David Bertran, Mark A. Ratner, and Alexander L. Burin, DNA electron transfer processes: Some theoretical notions, *Topics in Current Chemistry*, **237**, 1-36 (2004).
37. Frederick D. Lewis, Xiaoyang Liu, Jianqiu Liu, Scott E. Miller, Ryan T. Hayes, and Michael R. Wasleteroski, Direct measurement of hole transport dynamics in DNA, *Nature*, **406**, 51-53 (2000).
38. James P. Lewis, Pablo Ordejon, and Otto F. Sankar, Electric-structure-based molecular-dynamics method for large biological systems: Application to the 10 basepair poly(dG)-poly(dC) DNA double helix, *Physical Review B*, **55**, 6880-6887 (1997).
39. Juyeon Yi, Conduction of DNA molecules: A charge-ladder model, *Physical Review B*, **68**, 193103-1-4 (2003).
40. I. Saito, T. Nakamura, K. Nakatani, Y. Yoshioka, K. Yamaguchi, and H. Sugiyama, Mapping of the hot spots for DNA damage by one-electron oxidation: Efficacy of GG doublets and GGG triplets as a trap in long-range hole migration, *J. Amer. Chem. Soc.*, **120**, 12686-12687 (1998).

MULTIDIMENSIONAL IDENTIFICATION OF CHEMICAL WARFARE AGENTS USING SHAPED FEMTOSECOND PULSES

IGOR PASTIRK

*BioPhotonic Solutions, Inc.
Okemos, MI 48864, USA
Pastirk@BioPhotonic Solutions.com*

MARCOS DANTUS

*Department of Chemistry, Michigan State University,
East Lansing, M 48824, USA
dantus@msu.edu*

Detection and identification of chemical warfare simulants based on multidimensional phase shaped femtosecond laser pulses coupled to mass spectrometry (MS) is demonstrated. The presented approach is based on binary phase shaping (BPS) and aims to improve the accuracy and precision required for security applications. It is based on multiphoton intrapulse interference of femtosecond laser pulses. Spectra retrieved by applying n-differently shaped pulses represent n-dimensions of the analysis. We present a multidimensional technique for detection and identification of analogues to chemical agents and mixtures in real-time. Experimental results for dimethyl phosphate, pyridine, and three isomers of nitrotoluene are presented.

Keywords: pulse shaping; chemical detection.

1. Introduction

One of the main concerns for the detection of chemical warfare agents is the reliability and confidence in the measurement. Minimizing or potentially eliminating false positive and negative alarms is of utmost importance. The usual approach for increased reliability is to combine two or more different type of measurements or methods (dimensions), for example, gas chromatography-mass spectrometry GC-MS. We have developed an ultrashort laser-based method capable of fast (one second), accurate (even in a chemically complex environment), robust (stand alone, closed-loop, and portable), and reproducible molecular identification that works even when the molecules have the same chemical formula (isomers, stereoisomers and enantiomers).^{1,2} The method, BPS-MS, utilizes binary pulse shaping technology and relies on Multiphoton Intrapulse Interference (MII) that has been shown to be an invaluable tool for a range of applications.¹⁻⁷ Our approach is to increase the number of measurement dimensions in mass spectrometry using a

number of precisely defined shaped femtosecond pulses to control the fragmentation and ionization processes. The resulting fragmentation patterns are highly sensitive to the quantum mechanical interaction between the analyte molecule and the shaped laser pulse. Therefore, each shaped pulse produces a different fragmentation pattern, or 'fingerprint.' The use of two or more shaped pulses on a given analyte, increases the certainty of the molecular identification. The method is intended to be used for a selected number of otherwise difficult to identify chemical agents.

Finding the best shaped pulses (optimum field) to identify a compound can be a time consuming task given the almost infinite number of possibilities that a pulse-shaper provides. Using 100 pixels and only 100 phase values with 10 different amplitudes results in 10300 possible laser fields (pulse shapes). Such a staggering number of experiments

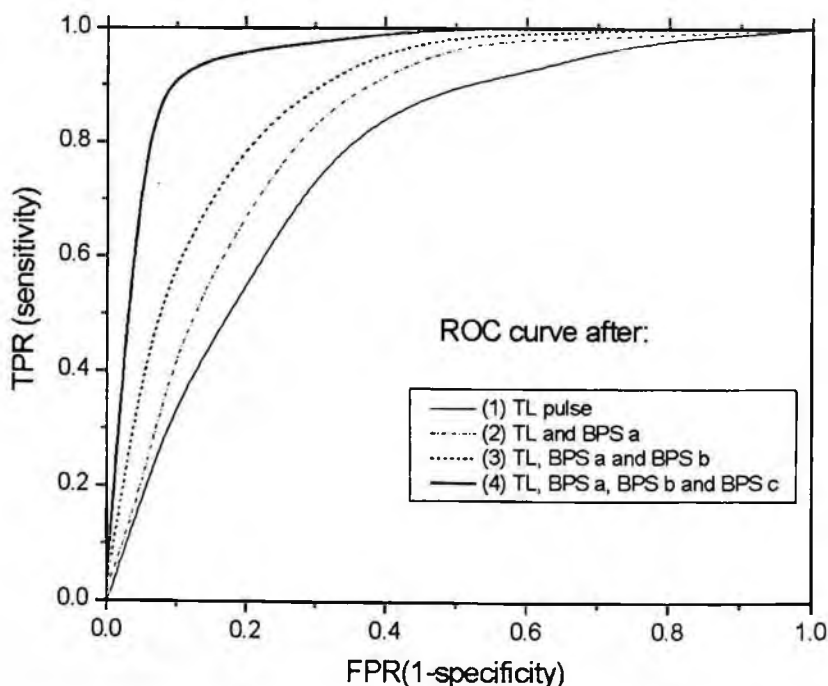


Figure 1. Simulated ROC curve for identification of chemical agent where lines 1-4 represent and increasing number of dimensions (distinct laser fields) used in the identification procedure. TL-transform limited pulse, BPS x - different binary shaped pulses.

that explore this enormous search space can not be performed, and there are several different approaches to investigate and search for the optimum fields. One of the most applied methods is closed-loop (feedback) genetic algorithm-based search or optimization, where a computer program that controls the pulse shaper starts from the random values and tries to mimic the natural selection by keeping the favorable solutions

(fields) and discarding the unfavorable ones.⁸⁻⁹ The results from this approach are rarely applicable outside the facility they were performed, therefore lacking universality. However, this method is easily implemented and requires minimal understanding of the process that is being optimized.

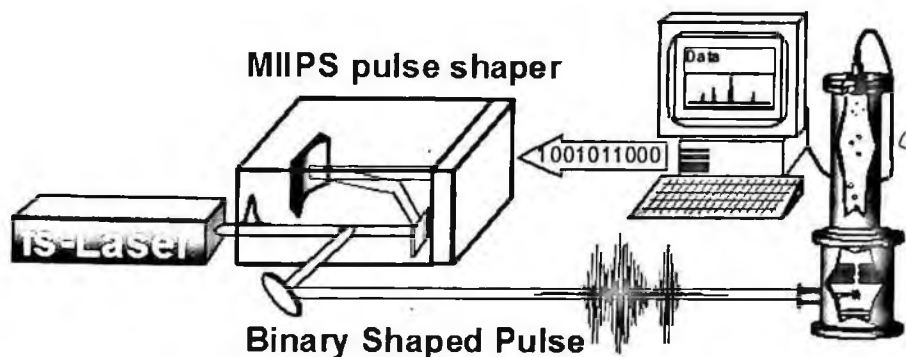


Figure 2. Schematic of the experimental set-up. Pulse shaper can be placed either inside the laser system or after the amplification.

The aim of BPS-MS approach presented here can be summarized by comparing the simulated receiver operator characteristic (ROC) curves shown in Figure 1. Each additional dimension of analysis (a distinct pulse shape, that yields distinct mass spectrum) enlarges the area under the curve and therefore the confidence of the measurement.

2. Experimental

The experiments were performed using 0.8 mJ/pulse amplified titanium-sapphire femtosecond laser (1 kHz repetition rate) that were 35 fs in duration when compressed using MIIPS method,⁵ and centered at 800 nm with full-width at the half maximum of 25 nm. The output of the seed laser (5 nJ, 50 nm full-width at the half maximum) was directed through a folded 2f pulse shaper. The pulse-shaper is based on the dual-mask liquid crystal spatial-light modulator (SLM) where each mask consists of 128 pixels.¹⁰ The SLM is positioned in the Fourier plane of the pulse shaper. Applying different voltages by computer controlled interface to the individual pixels in the SLM will alter the spectral phase of the femtosecond pulses and therefore their shape. After the pulse shaper and amplification in the standard regenerative amplifier, the beam was attenuated to 0.2 mJ/pulse and directed into the vacuum chamber connected to the time-of-flight (TOF) mass spectrometer. Samples (simulants of the chemical agents) were introduced in the vacuum system at the constant pressure in the range of $0.5 \cdot 10^{-5}$ to $1 \cdot 10^{-5}$ Torr. Mass spectra were recorded with 256 laser shots per applied binary phase. Computer controlled MIIPS-enabled pulse shaper was introducing the sequence of well-defined

spectral phase shapes and the mass spectrum was recorded for each of those laser fields. If the loss of the laser intensity is not an issue, it is possible to place the pulse shaper after the amplifier (see the schematic set-up in Figure 2). Results were processed on the computer and formed a database, which was subsequently used to identify presence or absence of the set of chemical agent simulants in the real time.

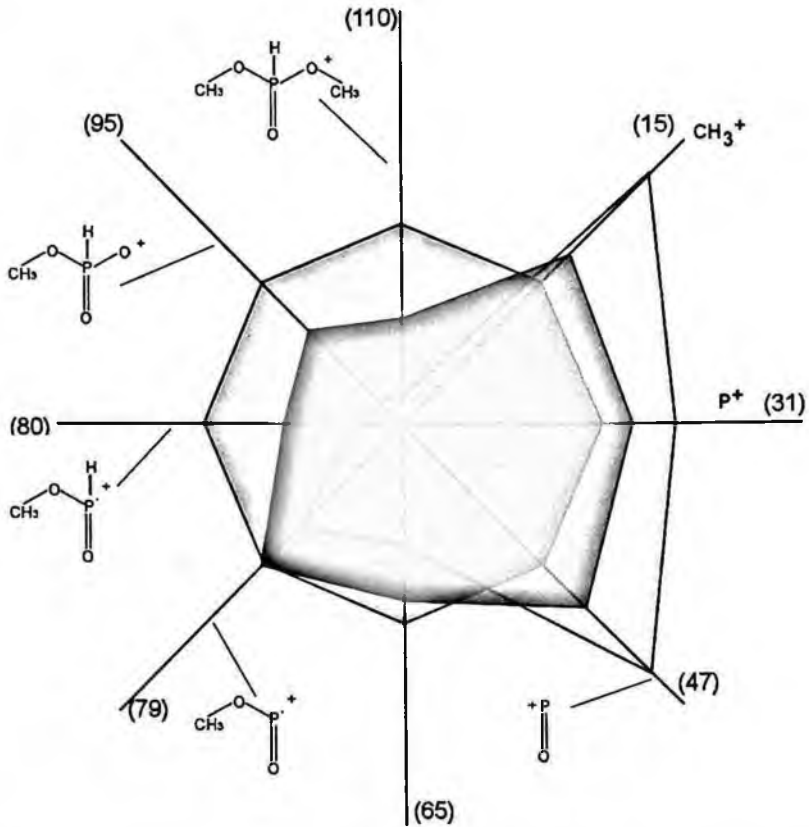


Figure 3. Controlled molecular fragmentation achieved by three different shaped femtosecond pulses yielding the three-dimensional analysis. The sample, dimethylphosphite (a nerve agent simulant), was irradiated with transform-limited (inner octagon) and binary-shaped femtosecond pulses BP 211 and BP 64. Numerals correspond to the sequence number in which the different BP were introduced. Intensities of mass spectra peaks shown here are normalized to intensities recorded for TL pulses. The spider plot clearly indicates significant changes in the fragment ion intensities. This allows enough variation to positively identify the presence of the chemical.

3. Results

In order to limit the search space we have developed a highly efficient method based on binary phase shaping³ which optimizes multiphoton intrapulse interference. Typically, we evaluate all 10-bit binary phase functions, which result in 1024 distinct fields. The results from these experiments are mapped based on criteria that can be used to identify particular chemical warfare agents. A subset of 3-10 shaped pulses are then selected for multidimensional identification purposes. In the field, the selected pulses are applied as a sequence achieving sub-second multidimensional analysis.

Experimental results shown in Figure 3 are collected on dimethyl phosphite (DMP), a nerve agent analogue. Two binary phases showed the largest changes on ion peak (m/z) intensities, most profoundly at m/z at molecular ion (110) and m/z of 95, 79 and 47. Three dimensional analysis shown in the Figure 3 contains MS obtained with TL pulses (octagon in Figure 3) and two binary phase functions to establish presence or absence of DMP. In principle, the larger the number of BP that yield results obtained with more complex and distinctive fragmentation patterns, the better the confidence level that a particular agent is present in a mixture containing interfering compounds. In practice we discovered that a small subset of distinct binary phases (3-10) is sufficient for successful identification. Absolute presence or absence of a given chemical is then determined by the detection algorithm, which takes into account the presence of the various fragmentation patterns. The minimal set of distinct phases needed for identification is illustrated in spider plot in Figure 3 containing the normalized mass spectra of dimethyl phosphite.

Molecular identification using mass spectrometry can be complicated when complex chemical mixtures are involved because of extensive overlap between fragment ions. When isomers are involved fragment-ion overlap is maximum, and typical electron-impact ionization mass spectrometry cannot provide absolute identification. We have tested our multidimensional identification method with a number of isomers and stereoisomers. Detailed results are presented elsewhere.^{2,11} An absolute identification was demonstrated on ortho-, meta- and para- isomers of nitrotoluene. The results (Figure 4) for o-nitrotoluene show the essential requirement for success of the identification procedure; a substantial difference in MS depending on excitation pulse shape. The method is also capable of providing quantitative concentration results.²

It is important to note that the spectrum of the excitation pulses does not change regardless of the binary phase applied, nor does the intensity of the laser beam. The changes to the laser field applied are ultimately very subtle, but provide enough difference in the response to positively identify the molecule of interest. An interesting effect can be observed when data in Figure 3 and 4 are analyzed; the largest suppression is usually achieved on the heaviest fragments. Deeper investigation of this effect is discussed elsewhere.¹²

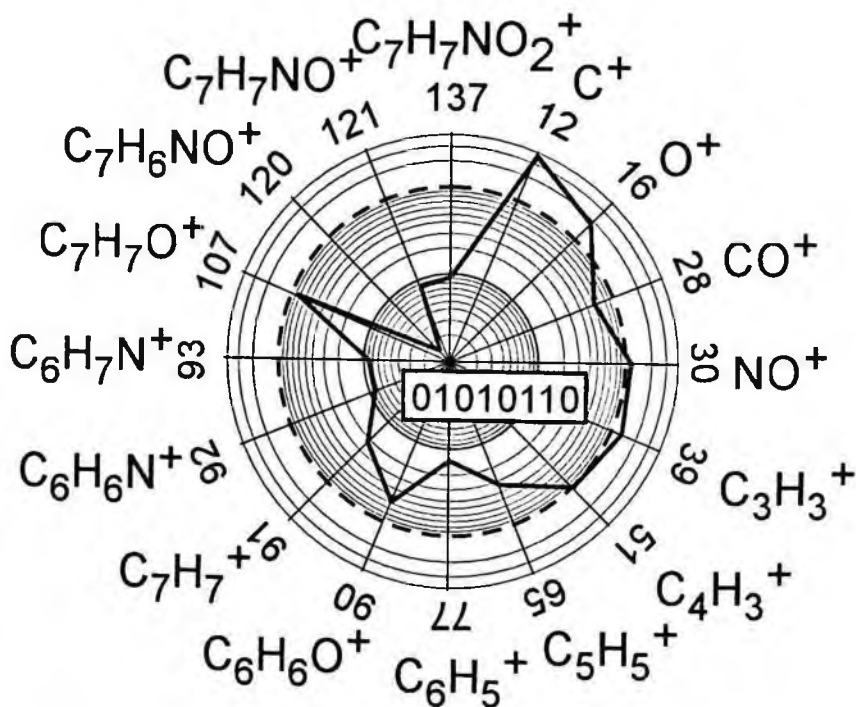


Figure 4. BPS-MS results for ortho-nitro toluene. The changes in the relative yields of a number of fragment ions (with corresponding m/z and chemical formula) as a result of binary phase shaping. Concentric circles that represent the relative intensity are on the logarithmic scale. Results for TL pulses are normalized in such a way that all fragments are assigned a relative yield of unity and shown by the dashed circle. The solid line that encompasses the blue-shaded area represents yields for binary phase function that minimized the ion with $m/z=120$ the most.

4. Conclusions

We have demonstrated chemical agent detection using accurately binary phase-shaped femtosecond pulses in the laboratory environment. Increase in multidimensionality has shown to improve the reliability of method. Future efforts will be focused on ambient sampling and expansion of database. Graphical user interface of the detection software is shown here (Figure 5) for a set of chemical agent simulants currently in the database.

5. Acknowledgments

We are grateful for funding from the Army Research Office through their STTR program for the development of pulse shaping technology that can be used for molecular identification of CB agents.

Chemical	value	Safe	Caution	Danger	Chemical	value	Safe	Caution	Danger
Meta Nitrotoluene	1.004	●	●	●	Meta Nitrotoluene	1.004	●	●	●
Ortho Nitrotoluene	0.9957	●	●	●	Ortho Nitrotoluene	0.9957	●	●	●
Para Nitrotoluene	0.9953	●	●	●	Para Nitrotoluene	0.9953	●	●	●
Benzene	0	○	●	●	Benzene	0	○	●	●
Pyridine	0.149	●	○	●	Pyridine	0.149	●	○	●
4 Chloro Benzonitrile	0	○	●	●	4 Chloro Benzonitrile	0	○	●	●
1 Nitroprazole	0	○	●	●	1 Nitroprazole	0	○	●	●
Dimethyl phosphite	0	○	●	●	Dimethyl phosphite	0	○	●	●
2 Chloro ethyl ethyl sulfide	0	○	●	●	2 Chloro ethyl ethyl sulfide	0	○	●	●
Safe level is below	0.1				Safe level is below	0.1			
Caution level is between	0.1	and	0.5		Caution level is between	0.1	and	0.5	
Danger level is above	0.5				Danger level is above	0.5			

Figure 5. GUI of the identification software based on BPS-MS method. Left panel: positive identification of three isomers of nitro toluene. Right panel: identification on a number of simulants in the complex mixture.

References

- Pastirk, I., Kangas, M., Dantus, M. Multidimensional analytical method based on binary phase shaping of femtosecond pulses *J. Phys. Chem. A*, 109, 2413 (2005)
- Dela Cruz, J.M., Lozovoy, V.V., Dantus, M. Quantitative mass spectrometric identification of isomers applying coherent laser control, *J. Phys. Chem.* 109, 8447 (2005)
- Comstock, M., Lozovoy, V.V., Pastirk, I., Dantus, M. Multiphoton intrapulse interference 6, binary phase shaping *Opt. Express* 12, 1061 (2004)
- Pastirk, I., J. M. Dela Cruz, K. A. Walowicz, V. V. Lozovoy, and M. Dantus, Selective two-photon microscopy with shaped femtosecond pulses. *Optics Express* 11, 1695 (2003)
- Lozovoy, V.V., Pastrik, I, Dantus, M. Multiphoton intrapulse interference. IV. Ultrashort laser pulse spectral phase characterization and compensation, *Opt. Lett.* 29, 775 (2004)
- B. W. Xu, J. M. Gunn, J. M. Dela Cruz, V. V. Lozovoy, M. Dantus, Quantitative investigation of the MIIPS method for simultaneous phase measurement and compensation of femtosecond laser pulses *J. Opt. Soc. Am. B*, 23, 750-759 (2006)
- Dela Cruz, J. M., Pastirk, I., Comstock, M., Lozovoy, V.V., Dantus, M., Use of coherent control methods through scattering biological tissue to achieve functional imaging, *P. Natl. Acad. Sci. USA* 101(49), 16996-17001 (2005)
- Assion, A., Baumert, T., Bergt, M., Brixner, T., Kiefer, B., Seyfried, V., Strehle, M., Gerber, G. Control of Chemical Reactions by Feedback-Optimized Phase-Shaped Femtosecond Laser Pulses *Science*, 282, 919 (1998)
- Levis, R. J., Menkir, G. M., Rabitz, H. Selective Bond Dissociation and Rearrangement with Optimally Tailored, Strong-Field Laser Pulses *Science*, 292, 709 (2001)
- A. M. Weiner, Femtosecond pulse shaping using spatial light modulators, *Rev. Sci. Instrum.* 71(5), 1929-1960 (2000)

11. Pastirk I, Dantus, M, Laser Applications to Chemical, Security and Environmental Analysis Conference, Incline Village, NV, (February 2006)
12. Lozovoy V.V., Gunaratne, T.C., Shane, J.C., and Marcos Dantus, Control of molecular fragmentation using binary phase shaped femtosecond laser pulses, ChemPhysChem (submitted 2006)

SURFACE-PLASMON-RESONANCE BASED OPTICAL SENSING

NORMAN J. MORGENSTERN HORING

*Department of Physics and Engineering Physics, Stevens Institute of Technology,
Hoboken, New Jersey 07030, USA*

H. L. CUI

*Department of Physics and Engineering Physics, Stevens Institute of Technology,
Hoboken, New Jersey 07030, USA*

Over the past twenty years, surface plasmon resonance has been developed as an effective technique for use in real-time biotechnological measurements of the kinetics of label-free biomolecular interactions with high sensitivity.¹⁻¹⁶ On a fundamental level, it is the dielectric-imaging involvement of the adsorbed biomolecular layer (DNA for example) in shifting the surface plasmon resonance (SPR) frequency by means of electrostatic coupling at the interface with the metal film substrate that facilitates SPR-based optical sensing. Of course, there are various factors that can influence surface plasmon resonance, including plasma nonlocality, phonons, multiplicity of layers, all of which should be carefully examined. Moreover, *tunable* SPR phenomenology based on the role of a magnetic field (both classically and quantum mechanically) merits consideration in regard to the field's effects on both the substrate¹⁷ and the adsorbed layer(s).¹⁸ This paper is focused on the establishment of the basic equations governing surface plasmon resonance, incorporating all the features cited above. In it, we present the formulation and closed-form analytical solution for the dynamic, nonlocal screening function of a thick substrate material with a thin external adsorbed layer, which can be extended to multiple layers. The result involves solution of the random phase approximation (RPA) integral equation for the spatially inhomogeneous system of the substrate and adsorbed layer,¹⁹⁻²⁵ given the individual polarizabilities of the thick substrate and the layer. (This is tantamount to the space-time matrix inversion of the inhomogeneous joint dielectric function of the system.) The frequency poles of the resulting screening function determine the shifted surface (and bulk) plasmon resonances and the associated residues at the resonance frequencies provide their relative excitation amplitudes. The latter represent the response strengths of the surface plasmon resonances (oscillator strengths), and will be of interest in optimizing the materials to be employed.

Keywords: Surface Plasmon Resonance; Optical Sensing.

1. Introduction

Surface plasmon resonance has become an important optical technique employed in chemical sensing and biotechnological measurements.¹⁻¹⁶ As a surface phenomenon, surface plasmons sensitively reflect upon the boundary conditions at the interface of a plasma-like medium. In this regard, they react to the introduction of adsorbate material

just outside the substrate, probing the properties of the adsorbate in conjunction with those of the substrate as they modify the joint optical response and reflectivity of the combined system.⁴⁻⁷ Central to such optical variations of the system is the concept of a dynamic, nonlocal and spatially inhomogeneous joint screening function $K(1,2)[1 = \mathbf{r}_1, t_1; 2 = \mathbf{r}_2, t_2, \text{etc.}]$, which is the space-time matrix inverse of the combined dielectric function $\varepsilon(1,2)$ of the composite system:

$$\int d^4 3 \varepsilon(1,3) K(3,2) = \delta^{(4)}(1-2). \quad (1)$$

Its physical significance is that of an electric field propagator in the sense that it propagates an impressed potential field $U(2)$ at space-time point 2 into the actual, effective field $V(1)$ at space-time point 1 through the (linear) relation

$$V(1) = \int d^4 2 K(1,2) U(2). \quad (2)$$

It is closely related to the concept of a longitudinal photon Green's function.

In this paper we describe in some detail the changes of the screening function of a thick substrate brought about by introducing a thin two dimensional (2D) adsorbate layer, illustrating the adsorbate layer as a planar 2D electron gas parallel to the substrate interface. As this system is translationally invariant in the 2D $x-y$ plane as well as translationally invariant in time, we have

$$K(1,2) = K(x_1 - x_2; y_1 - y_2; z_1, z_2; t_1 - t_2), \quad (3)$$

and employing a 2D spatial Fourier transform $[(x_1 - x_2, y_1 - y_2) \rightarrow (q_x, q_y) = \bar{q}]$ in the plane of translational invariance and ω -frequency transform $[(t_1 - t_2) \rightarrow \omega]$, we may rewrite Eq. (1) as

$$\int dz_3 \varepsilon(z_1, z_3; \bar{q}, \omega) K(z_3, z_2; \bar{q}, \omega) = \delta(z_1 - z_2). \quad (4)$$

It is apparent from Eq. (2) that the frequency poles of the screening function K describe resonant oscillations of the system which are longitudinal potential/density collective mode oscillations, i.e.: plasmons. While there is some difficulty in coupling transverse electromagnetic waves to longitudinal plasmons, there are methods by which this can be accomplished. A grating coupler can be used, also resonant Raman scattering and electron energy loss spectroscopy (EELS), but the method preferred experimentally in coupling to surface plasmons seems to be the use of a prism producing an evanescent wave from total internal reflection using TM (transverse magnetic) polarized light, which can excite surface plasmons – resulting in a sharp decrease of reflected light intensity

under appropriate conditions of energy and momentum conservation. In our analysis of the joint screening function carried out below, we arbitrarily take the region outside the substrate to be vacuum, but this can easily be adjusted to account for any other external medium.

2. Formulation of Dielectric Response of a Dynamic Plasma-Like Substrate Coupled to a Thin Adsorbed Layer

Considering a thick substrate medium to be modeled as a local semi-infinite plasma-like medium of dielectric function $\varepsilon = \varepsilon(\omega)$, and taking its boundary with vacuum to be at the plane $z = 0$, we have shown¹⁹ that the direct dielectric function of this semi-infinite system is ($\Theta_+(z)$ is the Heaviside unit step function)

$$\begin{aligned} \varepsilon_{\text{semi}}(z_1, z_2) = & \Theta_+(-z_1) \left[\delta(z_1 - z_2) + \delta(z_2)(\varepsilon - 1)e^{|\bar{q}|z_1/2} \right] \\ & + \Theta_+(z_1) \left[\delta(z_1 - z_2) + \delta(z_2)(1 - \varepsilon)e^{-|\bar{q}|z_1/2} \right]. \end{aligned} \quad (5)$$

Here, \bar{q} is the 2D Fourier transform wavevector conjugate to $\bar{r} = (x, y)$ and reference to it and to ω -frequency will be suppressed when possible. The associated semi-infinite screening function, $K_{\text{semi}}(z_1, z_2)$, which satisfies the inversion relation

$$\int d\bar{z} \varepsilon_{\text{semi}}(z_1, \bar{z}) K_{\text{semi}}(\bar{z}, z_2) = \delta(z_1 - z_2), \quad (6)$$

in the absence of an adsorbate, is given by^{20,21}

$$K_{\text{semi}}(z_1, z_2) = \frac{1}{\eta_+(z_1)} \delta(z_1 - z_2) - \delta(z_2) \Gamma \frac{\eta_-(z_1)}{\eta_+(z_1)} e^{-|z_1||\bar{q}|}, \quad (7)$$

where Γ is the image strength factor

$$\Gamma = \frac{1 - \varepsilon}{1 + \varepsilon} \quad (8)$$

and

$$\eta_-(z) = \Theta_+(z) - \Theta_+(-z) \quad ; \quad \eta_+(z) = \varepsilon \Theta_+(z) + \Theta_+(-z). \quad (9)$$

We characterize the adsorbed layer by a density perturbation response function^{20,21}

$$R_{2D}(z_1, z_2) = \delta(z_1 - z_0)\delta(z_2 - z_0)R^{2D}(\bar{q}, \omega), \quad (10)$$

which is completely localized onto a 2D sheet at $z = z_0$ by the δ -functions, and with $R^{2D}(\bar{p}, \omega)$ describing the dynamic, nonlocal density perturbation response on the 2D sheet. This yields the polarizability of the 2D sheet as

$$4\pi\alpha_{2D}(z_1, z_2) = 4\pi\alpha_0^{2D} e^{-|\bar{q}| |z_1 - z_0|} \delta(z_2 - z_0). \quad (11)$$

Here, $4\pi\alpha_0^{2D}$ is the 2D polarizability on the 2D sheet in \bar{q}, ω -representation. For example, in the case of a cold, local 2D plasma, it is given by $4\pi\alpha_0^{2D} \rightarrow -2\pi\rho_{2D}e^2 |\bar{q}| / m\omega^2$, where ρ_{2D} is the 2D electron sheet density.

The determination of the screening function, $K(z_1, z_2)$, of the combined system involves solution of the inversion relation

$$\int d\bar{z} \mathcal{E}(z_1, \bar{z}) K(\bar{z}, z_2) = \delta(z_1 - z_2) \quad (12)$$

where $\mathcal{E}(z_1, z_2)$ is determined for the combined system of the adsorbate sheet and the thick substrate by summing their polarizabilities so that

$$\mathcal{E}(z_1, z_2) = \mathcal{E}_{\text{semi}}(z_1, z_2) + 4\pi\alpha_{2D}(z_1, z_2). \quad (13)$$

In z -position space matrix notation,

$$\mathcal{E} = \mathcal{E}_{\text{semi}} + 4\pi\alpha_{2D}, \quad (14)$$

and, thus

$$\mathcal{E}K = 1 = \mathcal{E}_{\text{semi}}K + 4\pi\alpha_{2D}K. \quad (15)$$

Applying K_{semi} from the left, and also noting that $K_{\text{semi}}\mathcal{E}_{\text{semi}} = 1$, we have

$$K = K_{\text{semi}} - K_{\text{semi}}(4\pi\alpha_{2D})K, \quad (16)$$

which is the integral equation

$$K(z_1, z_2) = K_{\text{semi}}(z_1, z_2) - \int d\bar{z} \int d\bar{z}' K_{\text{semi}}(z_1, \bar{z}') 4\pi\alpha_{2D}(\bar{z}', \bar{z}) K(\bar{z}, z_2). \quad (17)$$

Recalling the $\delta(\bar{z} - z_0)$ -structure of $4\pi\alpha_{2D}(\bar{z}, \bar{z})$, we have

$$K(z_1, z_2) = K_{\text{semi}}(z_1, z_2) - 4\pi\alpha_0^{2D} \left(\int d\bar{z} K_{\text{semi}}(z_1, \bar{z}) e^{-|\bar{q}||\bar{z}-z_0|} \right) K(z_0, z_2). \quad (18)$$

To solve this, we need $K(z_0, z_2)$ on the right hand side, so we set $z_1 = z_0$ on the left and algebraically obtain

$$K(z_0, z_2) = \left(1 + 4\pi\alpha_0^{2D} \int d\bar{z} K_{\text{semi}}(z_0, \bar{z}) e^{-|\bar{q}||\bar{z}-z_0|} \right)^{-1} K_{\text{semi}}(z_0, z_2). \quad (19)$$

With this, the full solution for the screening function of the combined system is given by

$$K(z_1, z_2) = K_{\text{semi}}(z_1, z_2) - \frac{4\pi\alpha_0^{2D} \left(\int d\bar{z} K_{\text{semi}}(z_1, \bar{z}) e^{-|\bar{q}||\bar{z}-z_0|} \right) K_{\text{semi}}(z_0, z_2)}{1 + 4\pi\alpha_0^{2D} \int d\bar{z} K_{\text{semi}}(z_0, \bar{z}) e^{-|\bar{q}||\bar{z}-z_0|}}. \quad (20)$$

The final integration involved is

$$\int d\bar{z} K_{\text{semi}}(z, \bar{z}) e^{-|\bar{q}||\bar{z}-z_0|} = \frac{1}{\eta_+(z)} e^{-|\bar{q}||z-z_0|} - \Gamma \frac{\eta_-(z)}{\eta_+(z)} e^{-|\bar{q}||z|} e^{-|\bar{q}||z_0|}, \quad (21)$$

and the final result for $K(z_1, z_2)$ is obtained as

$$\begin{aligned} K(z_1, z_2) = & K_{\text{semi}}(z_1, z_2) - 4\pi\alpha_0^{2D} \left\{ \frac{1}{\eta_+(z_1)} e^{-|\bar{q}||z_1-z_0|} - \Gamma \frac{\eta_-(z_1)}{\eta_+(z_1)} e^{-|\bar{q}||z_1|} e^{-|\bar{q}||z_0|} \right\} \\ & \times \left[\frac{1}{\eta_+(z_0)} \delta(z_0 - z_2) - \delta(z_2) \Gamma \frac{\eta_-(z_0)}{\eta_+(z_0)} e^{-|\bar{q}||z_0|} \right] \\ & \times \left[1 + 4\pi\alpha_0^{2D} \left\{ \frac{1}{\eta_+(z_0)} - \Gamma \frac{\eta_-(z_0)}{\eta_+(z_0)} e^{-2|\bar{q}||z_0|} \right\} \right]^{-1}. \end{aligned} \quad (22)$$

3. Dispersion Relation for Coupling of Surface and Adsorbed Layer 2D Plasmons; Coupling of Surface and Quantum Well Plasmons; Conclusions

Plasmons of the combined system may be identified in terms of the frequency poles of the screening function $K(z_1, z_2)$, i.e.: vanishing of its denominator,

We characterize the adsorbed layer by a density perturbation response function^{20,21}

$$R_{2D}(z_1, z_2) = \delta(z_1 - z_0)\delta(z_2 - z_0)R^{2D}(\bar{q}, \omega), \quad (10)$$

which is completely localized onto a 2D sheet at $z = z_0$ by the δ -functions, and with $R^{2D}(\bar{p}, \omega)$ describing the dynamic, nonlocal density perturbation response on the 2D sheet. This yields the polarizability of the 2D sheet as

$$4\pi\alpha_{2D}(z_1, z_2) = 4\pi\alpha_0^{2D} e^{-|\bar{q}|z_1 - z_0} \delta(z_2 - z_0). \quad (11)$$

Here, $4\pi\alpha_0^{2D}$ is the 2D polarizability on the 2D sheet in \bar{q}, ω -representation. For example, in the case of a cold, local 2D plasma, it is given by $4\pi\alpha_0^{2D} \rightarrow -2\pi\rho_{2D}e^2|\bar{q}|/m\omega^2$, where ρ_{2D} is the 2D electron sheet density.

The determination of the screening function, $K(z_1, z_2)$, of the combined system involves solution of the inversion relation

$$\int d\bar{z} \varepsilon(z_1, \bar{z}) K(\bar{z}, z_2) = \delta(z_1 - z_2) \quad (12)$$

where $\varepsilon(z_1, z_2)$ is determined for the combined system of the adsorbate sheet and the thick substrate by summing their polarizabilities so that

$$\varepsilon(z_1, z_2) = \varepsilon_{\text{semi}}(z_1, z_2) + 4\pi\alpha_{2D}(z_1, z_2). \quad (13)$$

In z -position space matrix notation,

$$\varepsilon = \varepsilon_{\text{semi}} + 4\pi\alpha_{2D}, \quad (14)$$

and, thus

$$\varepsilon K = 1 = \varepsilon_{\text{semi}} K + 4\pi\alpha_{2D} K. \quad (15)$$

Applying K_{semi} from the left, and also noting that $K_{\text{semi}}\varepsilon_{\text{semi}} = 1$, we have

$$K = K_{\text{semi}} - K_{\text{semi}}(4\pi\alpha_{2D})K, \quad (16)$$

which is the integral equation

$$K(z_1, z_2) = K_{\text{semi}}(z_1, z_2) - \int d\bar{z} \int d\bar{z}' K_{\text{semi}}(z_1, \bar{z}') 4\pi\alpha_{2D}(\bar{z}', \bar{z}) K(\bar{z}, z_2). \quad (17)$$

Recalling the $\delta(\bar{z} - z_0)$ -structure of $4\pi\alpha_{2D}$ (\bar{z}, \bar{z}), we have

$$K(z_1, z_2) = K_{\text{semi}}(z_1, z_2) - 4\pi\alpha_0^{2D} \left(\int d\bar{z} K_{\text{semi}}(z_1, \bar{z}) e^{-|\bar{q}||\bar{z}-z_0|} \right) K(z_0, z_2). \quad (18)$$

To solve this, we need $K(z_0, z_2)$ on the right hand side, so we set $z_1 = z_0$ on the left and algebraically obtain

$$K(z_0, z_2) = \left(1 + 4\pi\alpha_0^{2D} \int d\bar{z} K_{\text{semi}}(z_0, \bar{z}) e^{-|\bar{q}||\bar{z}-z_0|} \right)^{-1} K_{\text{semi}}(z_0, z_2). \quad (19)$$

With this, the full solution for the screening function of the combined system is given by

$$K(z_1, z_2) = K_{\text{semi}}(z_1, z_2) - \frac{4\pi\alpha_0^{2D} \left(\int d\bar{z} K_{\text{semi}}(z_1, \bar{z}) e^{-|\bar{q}||\bar{z}-z_0|} \right) K_{\text{semi}}(z_0, z_2)}{1 + 4\pi\alpha_0^{2D} \int d\bar{z} K_{\text{semi}}(z_0, \bar{z}) e^{-|\bar{q}||\bar{z}-z_0|}}. \quad (20)$$

The final integration involved is

$$\int d\bar{z} K_{\text{semi}}(z, \bar{z}) e^{-|\bar{q}||\bar{z}-z_0|} = \frac{1}{\eta_+(z)} e^{-|\bar{q}||z-z_0|} - \Gamma \frac{\eta_-(z)}{\eta_+(z)} e^{-|\bar{q}||z|} e^{-|\bar{q}||z_0|}, \quad (21)$$

and the final result for $K(z_1, z_2)$ is obtained as

$$\begin{aligned} K(z_1, z_2) = & K_{\text{semi}}(z_1, z_2) - 4\pi\alpha_0^{2D} \left\{ \frac{1}{\eta_+(z_1)} e^{-|\bar{q}||z_1-z_0|} - \Gamma \frac{\eta_-(z_1)}{\eta_+(z_1)} e^{-|\bar{q}||z_1|} e^{-|\bar{q}||z_0|} \right\} \\ & \times \left[\frac{1}{\eta_+(z_0)} \delta(z_0 - z_2) - \delta(z_2) \Gamma \frac{\eta_-(z_0)}{\eta_+(z_0)} e^{-|\bar{q}||z_0|} \right] \\ & \times \left[1 + 4\pi\alpha_0^{2D} \left\{ \frac{1}{\eta_+(z_0)} - \Gamma \frac{\eta_-(z_0)}{\eta_+(z_0)} e^{-2|\bar{q}||z_0|} \right\} \right]^{-1}. \end{aligned} \quad (22)$$

3. Dispersion Relation for Coupling of Surface and Adsorbed Layer 2D Plasmons; Coupling of Surface and Quantum Well Plasmons; Conclusions

Plasmons of the combined system may be identified in terms of the frequency poles of the screening function $K(z_1, z_2)$, i.e.: vanishing of its denominator,

$$\eta_+(z_0) + 4\pi\alpha_0^{2D} \left\{ 1 - \eta_-(z_0) \Gamma e^{-2|\bar{q}||z_0|} \right\} = 0. \quad (23)$$

Taking the 2D sheet outside the surface, $z_0 < 0$, we have

$$1 + 4\pi\alpha_0^{2D} \left\{ 1 + \Gamma e^{-2|\bar{q}||z_0|} \right\} = 0, \quad (24)$$

and using the polarizabilities of a local bulk plasma ($\varepsilon = \varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2}$, where ω_p is the classical bulk plasma frequency and the background dielectric constant is taken as unity) and of a local 2D electron plasma sheet, ($4\pi\alpha_0^{2D} = -2\pi\rho_{2D}e^2|\bar{q}|/m\omega^2$, where ρ_{2D} is the 2D electron sheet density)

$$1 - \frac{2\pi\rho_{2D}e^2|\bar{q}|}{m\omega^2} \left\{ 1 + \Gamma e^{-2|\bar{q}||z_0|} \right\} = 0, \quad (25)$$

we find (set $\omega_{2D}^2 = 2\pi\rho_{2D}e^2|\bar{q}|/m$ for the 2D plasmon and $\omega_s^2 = \omega_p^2/2$ for the surface plasmon)

$$\omega^2(\omega^2 - \omega_s^2) = \omega_{2D}^2(\omega^2 - \omega_s^2[1 - e^{-2|\bar{q}||z_0|}]). \quad (26)$$

The plasmon roots of Eq.(26) are given by

$$\omega_{\pm}^2 = \frac{[\omega_s^2 + \omega_{2D}^2]}{2} \pm \sqrt{\left(\frac{\omega_s^2 + \omega_{2D}^2}{2}\right)^2 - \omega_{2D}^2\omega_s^2[1 - e^{-2|\bar{q}||z_0|}]}. \quad (27)$$

Clearly, the mode spectrum depends on z_0 . In the limit $z_0 \rightarrow -\infty$, we obtain $\omega_{\pm}^2 = \omega_s^2$ and $\omega^2 = \omega_{2D}^2$, where ω_s describes the decoupled surface plasmon ($z_0 \rightarrow -\infty$), and, separately, ω_{2D} describes the far-removed decoupled 2D plasmon. However, in the limit $z_0 \rightarrow 0$, we have $\omega_{\pm}^2 = \omega_s^2 + \omega_{2D}^2$, where ω_{\pm} is now the hybridized surface plasmon and 2D plasmon. ($\omega^2 = 0$ for $z_0 = 0$ is spurious.) For finite, nonvanishing z_0 , these hybridized plasmon modes are admixed as given by Eq.(27).

Incidentally, it is also of interest to consider the case of a 2D sheet representing a 2D plasma in a quantum well within the semi-infinite bulk, $z_0 > 0$. In this case Eq. (23) becomes

$$\varepsilon + 4\pi\alpha_0^{2D} \left\{ 1 - \Gamma e^{-2|\bar{q}||z_0|} \right\} = 0, \quad (28)$$

whence

$$1 - \frac{\omega_p^2}{\omega^2} - \frac{2\pi\rho_{2D}e^2|\bar{q}|}{m\omega^2} \{1 - \Gamma e^{-2|\bar{q}|z_0}\} = 0, \quad (29)$$

and we find

$$\omega^2 = \omega_p^2 + \frac{2\pi\rho_{2D}e^2|\bar{q}|}{m} \left\{ 1 - \frac{\omega_p^2 e^{-2|\bar{q}|z_0}}{2\omega^2 - \omega_p^2} \right\} = \omega_p^2 + \omega_{2D}^2 \left\{ 1 - \frac{\omega_p^2 e^{-2|\bar{q}|z_0}}{2\omega^2 - \omega_p^2} \right\}. \quad (30)$$

The plasmon roots of Eq.(30) are given by

$$4\omega_{\pm}^2 = (3\omega_p^2 + 2\omega_{2D}^2) \pm \sqrt{(3\omega_p^2 + 2\omega_{2D}^2)^2 - 8\omega_p^2 + \omega_{2D}^2\omega_p^2(1 + e^{-2|\bar{q}|z_0})}. \quad (31)$$

Again, the mode spectrum depends on z_0 . In the limit $z_0 \rightarrow \infty$, we obtain $\omega_+^2 = \omega_p^2 + \omega_{2D}^2$ and $\omega_-^2 = \omega_p^2/2$, where ω_+ describes the coupling of the 3D bulk and 2D plasmons deep in the medium ($z_0 \rightarrow \infty$), and, separately, ω_- describes the decoupled, distant surface plasmon. However, in the limit $z_0 \rightarrow 0$, we have $\omega_+^2 = \omega_p^2$ and $\omega_-^2 = \omega_p^2/2 + \omega_{2D}^2 = \omega_i^2 + \omega_{2D}^2$ where ω_+ is now the decoupled bulk plasmon deep in the medium, and, separately, ω_- describes the coupling of the surface plasmon with the 2D plasmon. For finite, nonvanishing z_0 these hybridized plasmon modes are admixed as given by Eq.(31).

It is clear from the foregoing considerations that the surface plasmon is shifted by interaction with the oscillatory modes of the adsorbed layer, and new coupled modes are introduced. In fact, the adsorbed layer substantially changes all the dielectric response properties of the substrate in accordance with Eq.(22). In consequence of this, its optical properties are modified, in particular in surface plasmon resonance experiments (as well as in all other probes). Analysis of such modifications reflect on the nature of the oscillatory modes of the adsorbate, which can identify it for sensing purposes. It should be noted that the determination of the screening function K (Eq.(22), for example) not only provides the shifted coupled mode spectrum in terms of its frequency poles, but it also provides the relative oscillator strengths of the various modes in terms of the residues at the poles. The analytic technique employed here for the adsorbate layer (in interaction with the substrate) can be extended to multiple layers, wire- and dot-like structures, lattices of such, as well as to the case of a few localized molecular oscillators. It can also take account of spatial nonlocality, phonons, etc., and the frequencies of the shifted surface (and other) plasmon resonances can be tuned by the application of a magnetic field.

Acknowledgements

This work was supported by the U.S. Department of Defense through the DURINT program administered by the Army Research Office, DAAD-19-01-1-05 92.

References

1. R.H. Ritchie, *Phys. Rev.* **106**, 874, (1957).
2. E.A. Stern and R.A. Ferrell, *Phys. Rev.* **111**, 1214 (1958).
3. H. Raether, "Surface Plasma Oscillations and their Applications", in *Physics of Thin Films*, Academic Press, NY, Vol. **9**, 145 (1977).
4. B. Liedberg, C. Nylander and I. Lundstrom, *Sensors and Actuators* **4**, 299 (1983).
5. M. Minunni, *Spectroscopy*, IOS Press, **17**, 613 (2003).
6. I. Pockrand, *Surface Science* **72**, 577 (1978).
7. S. Szunerits, et al., *Electroanalysis* **17**, 2001 (2005).
8. C. Nylander, B. Liedberg and T. Lind, *Sensors and Actuators* **3**, 79 (1982/83).
9. T. Gregory Drummond, Michael G. Hill and Jacqueline K. Barton, *Nature Biotechnology* **21**, 1192 (2003).
10. P. H. Bolivar, et al., *Phil. Trans. R. Soc. London A* **362**, 323 (2003).
11. Vitali Silin and Anne Plant, *Tibtech* **15**, 353 (1997).
12. H. Miyachi, et al., *Analytica Chimica Acta* **407**, 1 (2000).
13. R. Wang, et al., *Biosensors and Bioelectronics* **20**, 967 (2004).
14. R. Wang, et al., *Biosensors and Bioelectronics* **20**, 598 (2004).
15. E. Giakoumaki, et al., *Biosensors and Bioelectronics*, **19**, 337 (2003).
16. A. Kindlund and I. Lundstrom, *Sensors and Actuators* **3**, 63 (1982/83).
17. N.J.M. Horing, et al., *Phys. Lett.* **44A**, 386 (1973); *Solid State Comm.* **12**, 843 (1973); *J. Phys. C*, **6**, 2053 (1973); *Surface Science* **65**, 379 (1977); *Phys. Rev.* **B33**, 3895 (1986); *Physica* **E22**, 843 (2004); *Phys Rev.* **B31**, 4009 (1985); *Phys. Rev.* **B58**, 2001 (1998);
18. N.J.M. Horing, et al., *Ann. of Phys. (NY)* **97**, 216 (1976); *Solid State Comm.* **15**, 1381 (1974); *Solid State Comm.* **105**, 507 (1998); *Superlatt. and Microstr.* **20**, (1997).
19. N.J.M. Horing, T. Jena and H.L. Cui, *Phys. Rev.* **B54**, 2785 (1996).
20. N.J.M. Horing, E. Kamen and H.L. Cui, *Phys. Rev.* **B32**, 2184 (1985).
21. N.J.M. Horing, et al., *Proc. ICPIG 22*, Ed. E. Kunhardt, Vol. **3**, 27 (1995).
22. N.J.M. Horing, G. Fiorenza and H.L. Cui, *Phys. Rev.* **B31**, 6349 (1985).
23. N.J.M. Horing and J.D. Mancini, *Phys. Rev.* **B34**, 8954 (1986).
24. N.J.M. Horing, et al., *J. Phys C* **16**, 2215 (2004); also *Microelectronic Engineering* **51-2**, 219 (2000).
25. V. Fessatidis and N.J.M. Horing, *Physica E* **487** (2005); also *Proc. IEEE NANO 2005*.

LONG-WAVE INFRARED AND TERAHERTZ-FREQUENCY LASING BASED ON SEMICONDUCTOR NANOCRYSTALS

VALERY I. RUPASOV and SERGEI G. KRIVOSHLYKOV

*ALTAIR Center, LLC
1 Charnwell Circle, Shrewsbury, MA 01545, USA
altairctr@aol.com*

ALTAIR Center develops long-wave infrared (LWIR) and terahertz-frequency (THz) lasers operating at room temperature employing intraband luminescence in colloidal semiconductor nanocrystals, in which the optical transition frequencies can be easily tuned to the desired values by an appropriate choice of the semiconductor material and radius of the nanocrystals.

Keywords: Semiconductor nanocrystals, Terahertz lasers, Long-wave infrared lasers.

1. Lasing on Intraband Optical Transitions in Semiconductor Nanocrystals

The Long-Wave Infrared (LWIR) and terahertz (THz) frequency windows are of great scientific, commercial, and military interest for a variety of applications, ranging from fundamental research in solid-state physics to molecular detection in chemistry and biology, medical imaging and biological sensing, including military radar and spectroscopic sensing for homeland security.

A concept of THz optical emission source based on self-assembled quantum dots in GaAs/AlGaAs semiconductor structures has been proposed recently by Solomon *et al*¹. A GaAs/AlGaAs microdisc also plays a role in the high quality of an optical cavity, in which the quantum dot emitters are coupled to so-called whispering gallery modes. However, implementation of this concept requires a very complicated and expensive technology for fabrication, which is available only in few laboratories and could hardly be commercialized in the near future.

We propose the concept of the Long-Wave Infrared (LWIR) and THz-frequency lasing employing the intraband optical transitions in colloidal semiconductor nanocrystals, which can be fabricated using simple and cost-efficient methods of colloidal chemistry. The inversion population of the intraband transitions in the semiconductor nanocrystals can be created as a result of transfer of only one electron from the valence band to the conduction band. That completely prevents the Auger losses, which are presently the main obstacle for lasing on interband transitions in nanocrystals. The semiconductor nanocrystals placed in a high quality optical cavity should enable development of very efficient lasers having low threshold and power consumption.

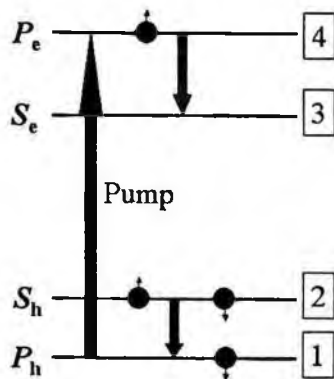


Fig. 1. Schematic of lasing on intraband transitions.

In the simplest model of a semiconductor nanocrystal (particle-in-a-box), the quantized energy levels are determined by the equation²:

$$E_{n,l}^{(\mu)} = \frac{\hbar^2 [\phi_{n,l}]^2}{2m_\mu R^2}. \quad (1)$$

Here, \hbar is the Planck constant, $\mu = e, h$ stands for the conduction and valence bands, respectively, m_μ are the effective masses of electron ($\mu = e$) and holes ($\mu = h$), and $\phi_{n,l}$ is the n -th zero of the spherical Bessel functions of the order l , $g_l(\phi_{n,l}) = 0$, and R is the nanocrystal radius. Two lowest energy levels in the conduction band and two highest energy levels in the valence band are shown in Figure 1, where letters S and P stand for the levels with the angular momenta $l = 0$ and $l = 1$, respectively.

Figure 1 illustrates how a pump laser radiation resonance to the transition $1 \rightarrow 4$ ($P_h \rightarrow P_e$) transfers an electron from state P_h in the valence band to state P_e in the conduction band. The electron transfer creates the inversion population for both intraband transition $4 \rightarrow 3$ in the conduction band and intraband transition $2 \rightarrow 1$ in the valence band. Although the number of electrons on P_h level is bigger than that on S_h level, we use the term "inversion population" because absorption of photons on the transition $1 \rightarrow 2$ is absent since all states on the level S_h are occupied.

Because most semiconductors have a sufficiently large energy gap, temperature population of conduction levels and depletion of valence levels are negligibly small. That makes the proposed intraband lasing temperature insensitive.

2. Radiative Decay Rate

The rate of radiative spontaneous decay, Γ_0 , of a dipole transition of frequency ω and dipole momentum d in a point-like emitter, such an atom or a molecule, embedded into a dielectric host of permittivity ϵ_{host} , is described by well-known expression³ (speed of light $c = 1$):

$$\Gamma_0 = \frac{4}{3} \omega^3 d^2 \sqrt{\epsilon_{\text{host}}} \quad (2)$$

However, in the case of colloidal semiconductor nanocrystals experiments show an essential depression of the decay compared to the value determined by Eq. (2). To explain this depression, Wehrenberg, Wang, and Guyot-Sionnest⁴ noted that the above expression, derived for point-like emitters, does not take into account the local field effect, or screening of the radiation field inside a nanocrystal of finite size. If a spherical nanocrystal of permittivity ϵ_{NC} is embedded into a dielectric host of permittivity ϵ_{host} and $\epsilon_{\text{NC}} > \epsilon_{\text{host}}$, then the internal electric field (inside the nanocrystal) E_{int} is weaker in comparison with the external field E_{ext} in the host, $E_{\text{int}} = S E_{\text{ext}}$, where the screening factor S is found to be:

$$S = \frac{3}{2 + \epsilon_{\text{NC}}/\epsilon_{\text{host}}} \quad (3)$$

Therefore, the expression (2) should be replaced by:

$$\Gamma = S^2 \Gamma_0 = \frac{4}{3} \omega^3 d^2 \left(\frac{3}{2 + \epsilon_{\text{NC}}/\epsilon_{\text{host}}} \right)^2 \sqrt{\epsilon_{\text{host}}} \quad (4)$$

The estimates of radiative lifetime performed using Eq. (4) for lead selenide (PbSe) and cadmium selenide (CdSe) nanocrystals show a good agreement with experimental data⁴. Although, expressions (3) and (4) are derived under assumption that the field variation is small on the nanocrystal size scale, while the radiation field exhibits a strong angular dependence, more accurate computations⁵ give the same result. It is clear also that Eq. (4) should be essentially modified in the case of so called core-shell nanocrystals with the dielectric permittivity $\epsilon(r) = \epsilon_{\text{core}}$ ($0 < r < R_1$), ϵ_{shell} ($R_1 < r < R_2$), where R_1 is the core radius, and R_2 is the external radius of the shell.

Solution of the Maxwell equations with appropriate boundary conditions on the core-shell and shell-host interfaces shows⁵ that the radiative decay rate of the dipole transition in the core is still given by the expression $\Gamma = W^2 \Gamma_0$, where the screening factor W is found to be:

$$W = \frac{9}{\left(2 + \frac{\epsilon_{\text{core}}}{\epsilon_{\text{shell}}} \right) \left(2 + \frac{\epsilon_{\text{shell}}}{\epsilon_{\text{host}}} \right) + 2 \left(1 - \frac{\epsilon_{\text{core}}}{\epsilon_{\text{shell}}} \right) \left(1 - \frac{\epsilon_{\text{shell}}}{\epsilon_{\text{host}}} \right) \left(\frac{R_1}{R_2} \right)^3} \quad (5)$$

3. Nonradiative Relaxation

Fast nonradiative relaxation of intraband transitions in semiconductor nanocrystals with the relaxation times shorter than 1 picosecond is obviously the major obstacle of lasing on these optical transitions. Although, at present time mechanism of the nonradiative relaxation is not clear, it has been recently demonstrated experimentally⁶ that fast nonradiative relaxation of intraband transitions is mainly determined by their strong

coupling to optical vibrations in organic ligands on the nanocrystal surface. An appropriate choice of ligands has resulted in depression of the nonradiative relaxation by 2 orders of magnitude. In core-shell nanocrystals, the shell increases the distance between an intraband optical transition localized in the core and optical vibrations in ligands on the external surface of the shell. That should obviously result in the further essential depression of the nonradiative relaxation.

4. Enhancement of Radiative Decay

For efficient lasing on intraband transitions one need not only to depress nonradiative relaxation but also to enhance the radiative decay rate of the transitions.

4.1. Bare nanocrystals

If the nanocrystal permittivity is greater than the host permittivity, $\epsilon_{\text{NC}} > \epsilon_{\text{host}}$, the screening factor S describes depression of the decay due to screening of the radiation field inside the nanocrystal. Frohlich was the first to note^{7,8} that for a metal nanocrystal with the complex frequency-dependent dielectric function

$$\epsilon_{\text{NC}}(\omega) = \text{Re } \epsilon_{\text{NC}}(\omega) + i \text{Im } \epsilon_{\text{NC}}(\omega) \quad (6)$$

the system metal nanocrystal plus dielectric host exhibits a resonance at the frequency $\omega = \Omega$ when real part of the denominator in expression (3) vanishes:

$$\text{Re } \epsilon_{\text{NC}}(\Omega) = -2\epsilon_{\text{host}} \quad (7)$$

At the resonance frequency, the screening factor is found to be:

$$S(\Omega) = \frac{3i \text{Re } \epsilon_{\text{NC}}(\Omega)}{2 \text{Im } \epsilon_{\text{NC}}(\Omega)} \quad (8)$$

If $\text{Re } \epsilon_{\text{NC}} \gg \text{Im } \epsilon_{\text{NC}}$, the above expression results in a huge resonance enhancement of the radiation field inside metal nanocrystals.

In our case of semiconductor nanocrystals, the real part of the dielectric function becomes negative in the vicinity of the transverse-longitudinal splitting, i. e. in the frequency range between the transverse, Ω_{T} , and longitudinal, Ω_{L} , frequencies of an optical phonon, $\Omega_{\text{T}} < \omega < \Omega_{\text{L}}$. In this frequency range, the dielectric function is well modeled by the expression:

$$\epsilon_{\text{NC}}(\omega) = \epsilon_{\infty} \frac{\omega^2 - \Omega_{\text{L}}^2 + i\omega\gamma}{\omega^2 - \Omega_{\text{T}}^2 + i\omega\gamma} \quad (9)$$

where ϵ_{∞} is the high-frequency permittivity (at $\omega \gg \Omega_{\text{L}}$), and γ is the relaxation parameter. Separating the real and imaginary parts of the dielectric function, we find from Eq. (7) the following equation for the Frohlich resonance frequency:

$$\epsilon_{\infty} \frac{\Omega^2 - \Omega_{\text{L}}^2}{\Omega^2 - \Omega_{\text{T}}^2} = -2\epsilon_{\text{host}} \quad (10)$$

If frequency of an optical transition in the nanocrystal Ω_{tr} equals to the resonance frequency Ω , $\Omega_{tr} = \Omega$, then the radiative decay rate of the transition is enhanced by the gain factor $G(\Omega)$, $\Gamma(\Omega) = G(\Omega)\Gamma_0$, where

$$G(\Omega) = |S(\Omega)|^2 = \left| \frac{3i \operatorname{Re} \varepsilon_{NC}(\Omega)}{2 \operatorname{Im} \varepsilon_{NC}(\Omega)} \right|^2 \quad (11)$$

Frequencies of optical vibrations in different semiconductor materials range from approximately 10 microns to submillimeters. For example, in silicon carbide (SiC) $\Omega_T = 23.8$ THz and $\Omega_L = 29.1$ THz, while in lead selenide (PbSe), $\Omega_T = 1.96$ THz and $\Omega_L = 6.14$ THz⁹. That obviously enables one to design high-efficiency optical media operating from mid-IR to terahertz-frequency ranges.

As an example, we estimate the resonance enhancement of an intraband optical transition in silicon carbide (SiC) nanocrystals. The dielectric function of SiC is well modeled by the expressions (7) with¹⁰ $\varepsilon_\infty = 6.52$, $\Omega_T = 793.9 \text{ cm}^{-1}$ (and the wavelength $\Lambda_T = 12.6 \text{ }\mu\text{m}$), $\Omega_L = 970.1 \text{ cm}^{-1}$ ($\Lambda_L = 10.3 \text{ }\mu\text{m}$), and $\gamma = 4.763 \text{ cm}^{-1}$. Note that the relaxation parameter γ is much less than the optical phonon frequencies, $\gamma/\Omega_T = 0.006$ and $\gamma/\Omega_L = 0.005$. The solution of the resonance condition (6) results in $\Omega \approx 902 \text{ cm}^{-1}$ and the corresponding resonance wavelength $\Lambda \approx 11 \text{ }\mu\text{m}$. Here and hereafter, in all our numerical estimates we accept the permittivity of a host matrix $\varepsilon_{\text{host}} = 2.25$, because this value is typical for many solvents, glasses, and polymers. Then, the gain factor $G(\Omega)$ is estimated to be approximately 3.6×10^2 .

4.2. Core-shell nanocrystals

For optical transitions localized in the core of core-shell nanocrystals, the screening factor is given by Eq. (5). If in this expression the real part of the complex dielectric function of the shell

$$\varepsilon_{\text{shell}}(\omega) = \operatorname{Re} \varepsilon_{\text{shell}}(\omega) + i \operatorname{Im} \varepsilon_{\text{shell}}(\omega) \quad (12)$$

is negative, $\operatorname{Re} \varepsilon_{\text{shell}}(\omega) < 0$, then the real part of denominator in Eq. (5) vanishes at two resonance frequencies Ω_1 and Ω_2 .

To avoid simple but more tedious computations, we restrict our further analysis to the case of a relatively thick shell of the thickness $\Delta = R_2 - R_1 \geq R_1$. In such core-shell nanocrystals, the geometry factor $\sigma = (R_1/R_2)^3$ is much less than 1, $\sigma \ll 1$. Therefore, computing zeroes of the real part of denominator in Eq. (5), we can omit the term proportional to σ . Then, the resonance frequencies are determined by the equations:

$$\operatorname{Re} \varepsilon_{\text{shell}}(\Omega_1) = -2\varepsilon_{\text{host}}, \quad (13)$$

$$\operatorname{Re} \varepsilon_{\text{shell}}(\Omega) = -\frac{1}{2} \varepsilon_{\text{core}}. \quad (14)$$

Note that at $\varepsilon_{\text{core}} = 4 \varepsilon_{\text{host}}$ we deal with the degenerate case when $\Omega_1 = \Omega_2$.

In core-shell nanocrystals we still have one free parameter - the geometry factor σ . An appropriate choice of the geometry factor enables increasing one of the magnitudes: $W(\Omega_1)$ or $W(\Omega_2)$. If $\varepsilon_{\text{core}} < 4 \varepsilon_{\text{host}}$, then at $\sigma = \sigma_1$, where

$$\sigma_1 = \frac{4\varepsilon_{\text{host}} - \varepsilon_{\text{core}}}{2(\varepsilon_{\text{core}} + 5\varepsilon_{\text{host}})} \quad (15)$$

we find:

$$W(\Omega_1; \sigma_1) = \left[\frac{3 \operatorname{Re} \varepsilon_{\text{shell}}(\Omega_1)}{2 \operatorname{Im} \varepsilon_{\text{shell}}(\Omega_1)} \right]^2 \quad (16)$$

That results in the gain factor:

$$G(\Omega_1; \sigma_1) = [W(\Omega_1; \sigma_1)]^2 = \left[\frac{3 \operatorname{Re} \varepsilon_{\text{shell}}(\Omega_1)}{2 \operatorname{Im} \varepsilon_{\text{shell}}(\Omega_1)} \right]^4 \quad (17)$$

The expression in square brackets in the above formula is nothing but the screening factor at the Frohlich resonance frequency, and hence, the gain factor in core-shell nanocrystals is equal to the square of the Frohlich gain factor, $G_{\text{core-shell}} = [G_{\text{Frohlich}}]^2$.

For nanocrystals with the CdSe core of permittivity $\varepsilon_{\text{core}} = 6.2$ in a host matrix of permittivity $\varepsilon_{\text{host}} = 2.25$, the required geometry factor $\sigma_1 = 0.08$, and we find $\Delta \approx 1.3 R_1$.

If $\varepsilon_{\text{core}} > 4 \varepsilon_{\text{host}}$, then at $\sigma = \sigma_2$, where

$$\sigma_2 = \frac{\varepsilon_{\text{core}} - 4\varepsilon_{\text{host}}}{2(2\varepsilon_{\text{core}} + \varepsilon_{\text{host}})}, \quad (18)$$

we derive analogous expressions for the magnitude:

$$W(\Omega_2; \sigma_2) = -\frac{3 \operatorname{Re} \varepsilon_{\text{shell}}(\Omega_2)}{2 \operatorname{Im} \varepsilon_{\text{shell}}(\Omega_2)} \frac{3\varepsilon_{\text{host}}}{\operatorname{Im} \varepsilon_{\text{shell}}(\Omega_2)} \quad (19)$$

and the gain factor:

$$G(\Omega_2; \sigma_2) = \left[\frac{3 \operatorname{Re} \varepsilon_{\text{shell}}(\Omega_2)}{2 \operatorname{Im} \varepsilon_{\text{shell}}(\Omega_2)} \right]^2 \left[\frac{3\varepsilon_{\text{host}}}{\operatorname{Im} \varepsilon_{\text{shell}}(\Omega_2)} \right]^2. \quad (20)$$

For nanocrystals with PbSe core of permittivity $\varepsilon_{\text{core}} = 22.9$ in a host matrix of permittivity $\varepsilon_{\text{host}} = 2.25$, the required geometry factor $\sigma_2 = 0.14$, and we find $\Delta \approx R_1$.

Thus, core-shell nanocrystals exhibit a huge enhancement of an optical transition localized in the core provided that the geometry factor is chosen to maximize the gain factor either at the frequency Ω_1 or at the frequency Ω_2 . An appropriate choice of materials, including host material, and sizes of core-shell nanocrystals enables the development of extremely efficient optical media operating in an extremely wide spectral range - from mid-IR to terahertz.

To demonstrate high efficiency of core-shell nanostructures we estimate the gain factor in the core-shell nanocrystals with the shell made of SiC, focusing on the

resonance frequency Ω_1 that is controlled by host permittivity only. As in the case of bare SiC nanocrystals in a host of permittivity $\epsilon_{\text{host}} = 2.25$, the resonance frequency $\Omega_1 = 902 \text{ cm}^{-1}$ ($\Lambda_1 = 11 \text{ }\mu\text{m}$). The core of the structure can be made out of any semiconductor material providing an intraband optical transition with the transition frequency $\Omega_{\text{tr}} = \Omega_1$. However, permittivities of the core and host control the required geometry factor σ_1 . For CdSe core of radius R_1 the shell thickness $\Delta = 1.3 R_1$.

In the case of bare nanocrystals the Frohlich gain factor is given by Eq. (11), and was estimated to be $G_{\text{Frolich}} = 3.6 \times 10^2$. Therefore, for core-shell nanocrystals with the SiC shell, we immediately find:

$$G_{\text{core-shell}}(\Omega_1; \sigma_1) = [G_{\text{Frolich}}]^2 = 1.3 \times 10^5 \quad (21)$$

Thus, semiconductor core-shell nanocrystals can exhibit a huge enhancement of optical transitions lying in the spectral range from 10 microns to submillimeters.

5. Conclusion

Core-shell nanocrystals exhibit a huge enhancement (up to 10^5) of an optical transition localized in the core provided that the geometry factor is chosen to maximize the gain factor either at the resonance frequency Ω_1 or at the frequency Ω_2 . In addition, an appropriate choice of ligands could result in the depression of the nonradiative relaxation of intraband transitions by 2 orders of magnitude. Thus, a proper selection of materials, including the host material, sizes of core-shell nanocrystals, and ligands enables the development of efficient and cost-effective laser media operating in an extremely wide spectral range - from mid-IR to terahertz.

6. Acknowledgments

This work is supported by the Army Research Office STTR Phase II contract W911NF-05-C-0124 (Dr. Dwight Woolard).

References

1. G. S. Solomon, Z. Xie, and M. Agrawal, *Int. J. of High Speed Electronics and Systems* **13**, 1099 (2003).
2. Al. L. Efros and M. Rosen, *Annu. Rev. Mat. Sci.* 2000, 30: 475-521.
3. L. D. Landau and E. M. Lifshits, *Electrodynamics of Continuous Media* (Elsevier, Oxford, 1984).
4. B. L. Wehrenberg, C. Wang, and P. Guyot-Sionnest, *J. Phys. Chem. B* **106**, 10634 (2002).
5. V. I. Rupasov, to be published.
6. P. Guyot-Sionnest, B. Wehrenberg, and D. Yu, *J. Chem. Phys.* **123**, 074709 (2005).
7. H. Frohlich, *Theory of Dielectrics* (Oxford Univ. Press, London, 1949).
8. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, NY, 1983).
9. O. Madelung, *Semiconductors - Basic Data* (Springer, Berlin, 1996).
10. W. G. Spitzer and D. A. Kleinman, *Phys. Rev.* **121**, 1324 (1961).

ULTRASENSITIVE WIDEBAND INTEGRATED SPECTROMETER FOR CHEMICAL AND BIOLOGICAL AGENT DETECTION

IGOR V. VERNIK

*HYPRES, Inc.,
175 Clearbrook Rd., Elmsford, NY 10523, USA
vernik@hypres.com*

A novel concept of a compact mm/submm integrated spectrometers for environmental monitoring for hazardous materials of chemical and biological origin as well as for remote monitoring of the Earth atmosphere is discussed. The agents will be exactly identified by their unique spectral signatures. The assembled on a multi-chip module, cryocooler-mounted Superconducting Integrated SPectrometer (SISP) exploits the superior performance of superconducting Josephson junction technology and unique on-chip integration of analog components, analog-to-digital converter, and digital components. Analog components include a superconductor-insulator-superconductor (SIS) mixer with integrated quasioptical antenna, mm-wave local oscillator, and SQUID amplifier for the down-converted (IF) signals. Upon amplification, the IF signal is digitized using a bandpass delta-sigma modulator, followed by real time processing with rapid single flux quantum (RSFQ) circuitry. Experimental results showing both operation of spectrometer components and the way to their successful integration are presented.

Keywords: Biological and chemical agents, remote detection, sensing and imaging, millimeter and submillimeter wave, THz spectrometer, superconductor, RSFQ.

1. Introduction

Lightweight and compact ultra-sensitive submm integrated spectrometer is very attractive for environmental monitoring for hazardous materials of chemical and biological origin, radio-astronomical research and remote monitoring of the Earth atmosphere. The Superconducting Integrated SPectrometer (SISP) offers a unique on-chip integration of different planar components such as a SIS (superconductor-insulator-superconductor) mixer with quasioptical antenna, a superconducting local oscillator (LO), an intermediate frequency (IF) SQUID amplifier and the circuits for digitizing of down converted signals and their real time processing.

Over the past years, the proliferation of chemical and biological agents as instruments of warfare and terrorism has become a major national security issue. The threat of chemical/biological terrorist attack after the September 11th 2001 and the 1995 nerve gas attack in Tokyo underscore the need for reliable instruments capable of detecting dangerous or potentially lethal chemical and biological agents. The new ambitious radio-astronomy multi-dish projects would gain considerably by using a single-chip SISP due

to their low price and better serviceability as compared to conventional approaches. Finally, a remote study of atmospheric pollution is possible by using air or satellite borne SISP in order to detect the spectrum lines of elements.

A lot of activities recently arise in both the technology and scientific arenas associated with the THz frequencies—i.e., usually defined as the portion of the submillimeter-wavelength electromagnetic spectrum between approximately 1 mm (300 GHz) and 100 μ m (3 THz). The THz regime inherently offers important technical advantages such as component compactness, wider bandwidths and improved spatial resolutions when compared to RF electronics and holds the promise for new and novel sensing applications (e.g., chemical and biological agent detection, inspection of sealed packages, concealed weapons detection, medical diagnostics, etc.). However, the same THz regime presents significant challenges such as extreme atmospheric attenuation, weak interaction signatures and standing wave interference to name few for practical implementation within traditional scenarios. There are also scientific and engineering problems that have to date either prohibited or severely limited the implementation of conventional electronics within this regime where wavelength is on the order of component size. In addition, there are the strong scientific payoffs of the THz regime, which is the most richly populated portion of the spectrum in terms of spectral signature information. As an example, interstellar dust (which contains an array of light to heavy molecules) was predicted long ago to contain literally tens of thousands of individual spectral lines. However, only a few thousands have been resolved and many have not yet even been identified¹.

In this paper, we propose to use the same proven spectroscopic approach further advanced by unsurpassed speed of superconductor digital electronics for the real-time DSP for identification of chemical and biological agents. These same spectral signatures, which are so interesting for interstellar and intragalactic science, are also present in planetary atmospheres and therefore have relevance for atmospheric monitoring applications. The unique physical properties associated with how THz radiation interacts with matter (i.e., molecular resonances) have fueled a long term focus on laboratory-based molecular spectral analysis and on astronomical and atmospheric remote sensing.

The SIS mixer itself is undoubtedly the device of choice for a low noise front-end detector at THz frequencies. Since the noise temperature of an SIS mixer is ultimately limited only by the fundamental quantum value hf/k^2 , SIS heterodyne spectrometers have been successfully used in radio astronomy for observation of spectra with the lowest possible noise temperature in the mm and sub-mm wave range. Many applications lack a compact and easily tunable submm LO. At frequencies above 30 GHz there is a steep increase in the cost and complexity of solid state radiation sources. Gunn diodes are widely used as LOs. They have highest fundamental frequency of operation between commercially available mm wave sources up to 110 GHz. Extension to higher frequencies requires low efficiency harmonic generators. Moreover, Gunn diode sources are voltage tunable over a range typically less than $\pm 10\%$ of the resonant frequency of the oscillator. This tuning range is inadequate for many applications. We propose to use unidirectional fluxon (flux quantum $\Phi_0 = h/2e$) propagation in a long Josephson tunnel junction (LJJ) as a way to achieve practical and compact on-chip LO. This is Flux-Flow Oscillators (FFOs) that satisfy the following important requirements: they deliver enough power to pump a mixer and have high frequency tunability. For standard superconducting technology when Nb is used to fabricate the FFO, such an oscillator can

be tuned from 0 to 700 GHz. For other superconducting materials (such as NbN) the upper frequency can be increased to over 1 THz. As the result of this tunability, the same integrated spectrometer can be used for chemical/biological agent detection in extremely wide (from 0 to 1THz) frequency range.

The characteristics of two *Bacillus globigii* (BG) signatures (see review by Woolard et al³) prove real viability of the proposed SISP approach. BG has signatures at ~260 GHz and ~420 GHz with signature width depending from BG form (dense, diluted and aerosol) from 6 GHz to 30 GHz³. These characteristics are very good match for proposed SISP approach. Harmless to humans, BG is ubiquitous and found easily in samplings of wind-borne dust. BG is safely used in biological studies as a stand-in for pathogenic bacteria. BG is used as a biological tracer for anthrax because its particle size and dispersal characteristics are similar to those of anthrax. A household bleach-and-water solution easily kills BG.

Digital low-temperature superconductor (LTS) technology, using rapid single flux quantum (RSFQ) logic, is the fastest integrated circuit (IC) technology today. Individual logic gates operating at 750 GHz have been demonstrated in this technology⁴. Large-scale integrated circuits with thousands of gates are being routinely manufactured at the HYPRES superconductor electronics (SCE) facility. These complex circuits exhibit synchronous operation with clock rates of 20 GHz⁵ and are expected to increase up to 40 GHz with on-going fabrication enhancements. Such performance is well beyond any other IC technology. SCE is also an extremely low-energy digital technology. Switching energy of only 10^{-18} J/gate results in negligible power dissipation (~100 nW) even at clock rates of 100 GHz. Moreover, the fundamental quantum mechanical nature of RSFQ logic ensures high-fidelity data conversion between analog and digital formats, resulting in very high linearity analog-to-digital converters.

The specific benefits of superconducting integrated spectrometer technology for chemical and biological agent detection are:

- **Precise Agent Identification** – The chemical and biological agents will be exactly identified by their unique fingerprint - their spectral signatures.
- **Enhanced Sensitivity** – The noise temperature of the wideband cryogenic spectrometer is dramatically smaller (from a factor of 4 to as much as a factor of 25 depending on the bandwidth) than a conventional spectrometer, with the advantage of being greater for wider bandwidths. This makes possible detection of lower concentration of chemical and biological agents.
- **Both Passive and Active Spectrometer** – The spectrometer may employ both the passive and active architecture. In former case it would provide spectra excited by ambient conditions (sun light, normal temperature...) and not require powerful excitation source. In later case external excitation (e.g. by GaAs Schottky diode) may be provided.
- **Ultra-wideband Agent Detection** – Use of long Josephson junctions as continually tunable LOs will allow dc-to-sub THz spectrometry.
- **Compactness** – The whole spectrometer will be implemented on a few $10 \times 10 \text{ mm}^2$ and $5 \times 5 \text{ mm}^2$ chips mounted on the multi-chip module.
- **Low Cost and High Reliability** – All components of the spectrometer are fabricated using standard low-cost, thin film, all Nb fabrication process. The final product is packaged in an ultra-reliable cryocooler with a Mean-Time-Between-Failures (MTBF) exceeding 99 years. Superconductor circuits are extremely radiation-hard, which make them attractive for space applications.

The concept of an integrated spectrometer is not new. It was first suggested by Koshelets and Shitov⁶ (see also references therein) and further developed by us⁷. The efforts in its development targeted the successful on-chip integration of such *analog* components as LO and SIS mixer with quasioptical antenna⁶. In this paper we propose a real breakthrough with the integration of superconducting thin film *analog* components (LO with phase-lock loop, mixer, and IF amplifier) with superconducting *digital* RSFQ circuitry on a multi-chip module (MCM). This spectrometer MCM will be packaged on cryocooler.

2. Integrated Digital Spectrometer

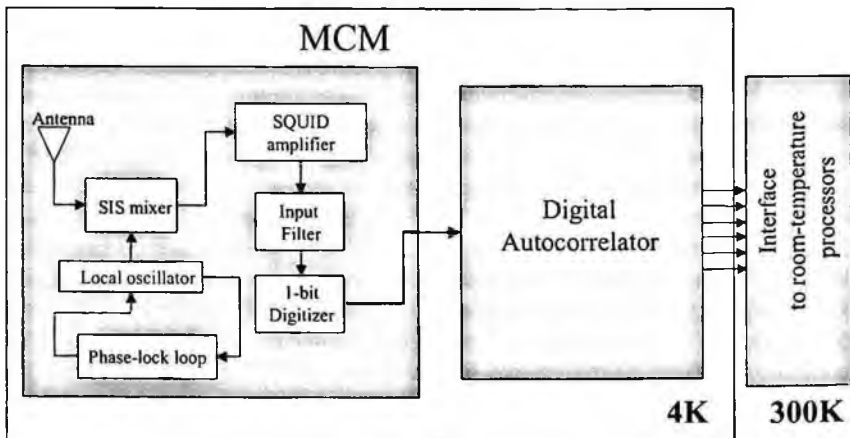


Fig. 1. The block diagram of integrated spectrometer mounted on MCM module. See text for more details.

Fig. 1 shows the block diagram of the integrated spectrometer under development. The superconducting part of the receiver consists of two chips mounted on a multi-chip module (MCM).

This chip die sizes are 5×5 and 10×10 mm². The first chip (5×5 mm²) of the spectrometer, the "front-end" chip, has the local oscillator together with the SIS mixer and impedance matching structure for better power delivery and to tune out the SIS junction capacitance. The frequency resolution of SISP determined by both the instant linewidth of the LO and its long-time stability along with the noise temperature is one of the major parameters in spectral measurements. The phase-lock loop consisted both on-chip circuitry and room-temperature electronics is employed.⁸

The first part of digital circuitry consists of a simple, fast ADC that takes the IF analog output and generates a one-bit digital pulse code at a clock rate of the order of 20 GHz or greater. The typical noise temperature of the 1-bit digitizer is of the order of few hundred K. This is significantly higher than the noise temperature of about 40 K at 475 GHz for a reference SIS mixer pumped by an external local oscillator⁶. Therefore, an IF amplifier with a power gain G_A of about 10 dB and a noise temperature T_N of about 100 K necessary to achieve the best possible overall noise temperature is designed on the same front-end chip. An IF amplifier is attached to an IF port of SIS mixer, since it helps to avoid losses of the long cable. A semiconducting IF amplifier 'integrated' with SIS mixer has a drawback of significant heat load (typically 20mW per stage). This makes an RF amplifier based on a SQUID that has an ultra-low power consumption, small size and

“natural” compatibility with any SIS based structure a natural choice for integration with SIS mixer.

The second chip (10 X 10 mm²) has a digital-signal processing circuit - the digital autocorrelator. We have selected the correlator circuit as the core component of our digital signal processor. One classic way to obtain the frequency spectrum of a time-domain signal $f(t)$ is to take the correlation between $f(t)$ and a time-delayed version of the same function $f(t-\tau)$. The autocorrelation function is then given by

$$R(\tau) = \frac{1}{T} \int f(t)f(t-\tau)dt, \text{ where the integral is over a long time period } T. \text{ That will}$$

selectively increase coherent frequency components that are periodic with time τ ; other components will exhibit a random walk. The power spectral density function $S(f)$ (or simply the spectrum of the signal) is then the Fourier transform of $R(\tau)$. If the signal is first digitized, then both the autocorrelation and the Fourier transform can be obtained in the discrete digital domain. The resolutions in the time and frequency domains are similar.

For a rapidly changing signal, this correlation function must be computed in real-time, by using a large number of correlators and accumulators in parallel. These parallel channels with distinct delays are generally known as “lags”. However, for radiometric imaging, the spectrum of the signal is relatively stable and changes only very slowly. In this case, one can compute $R(\tau)$ serially for various discrete values of τ , and then Fourier transform to obtain the spectrum. This is particularly true for RSFQ electronics, where the serial computations are so fast. This approach has the advantage of minimizing the circuit area, total power consumption, and the number of input/output (I/O) lines.

As an example, a 128-channel autocorrelator formed by 16-lag autocorrelator and a 112-bit programmable delay shift register (PRSR) was designed, fabricated and successfully tested. The PRSR acts as programmable (in increments of 16) delay line; data flow from a 16-lag autocorrelator to the PRSR and then come back.

The experimental results pertinent to front-end chip and digital autocorrelator follow in next paper sections.

3. Front-End Chip

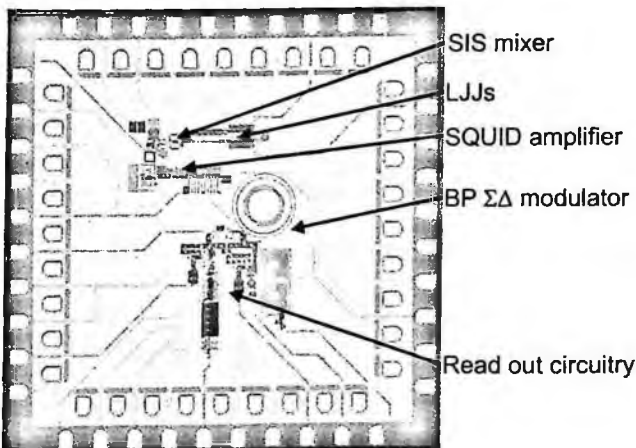


Fig. 2. The photo of front-end chip. See text for details.

For easy testability we are developing a self-contained spectrometer, e.g. no actual outside signal is investigated and this outside signal is modeled by an on-chip oscillator, very similar to the LO only with weaker coupling to mixer.

Unidirectional fluxon (flux quantum $\Phi_0 = h/2e$) propagation in a long Josephson tunnel junction (LJJ) is used as a way to achieve a practical and compact on-chip LO. A Josephson junction is defined as "long" when its physical length $L \gg \lambda_j$ and the width $W \ll \lambda_j$, where Josephson penetration depth λ_j is the characteristic dimension of an unperturbed fluxon. A fluxon in a long Josephson junction (LJJ) carries a magnetic flux equal to one flux quantum Φ_0 . A dc magnetic field H is applied in the plane of the junction. With increasing the external magnetic field, the screening current at the junction edge becomes unstable and forms a closed loop, which enters the interior of the junction. Fluxons are continuously generated at one junction end, and accelerated by the bias current toward the other end. Here the reflections generate a regular standing wave pattern resulting in emitted radiation. The oscillation frequency is described by the formula: $f = V_{dc} / \Phi_0 = u d \mu_0 H / \Phi_0$, where d is the magnetic thickness of the junction and μ_0 is the permeability free space. The average velocity u of fluxons is proportional to the bias current.

The oscillation frequency can be tuned over the wide range by changing the applied field, by varying the fluxon density in the junction, and by changing the bias current.

All circuits in this paper are fabricated using the standard HYPRES 3 μm 1 kA/cm² process⁹. Fig. 2 shows the layout of the chip that incorporates all front-end structures. The chip consisted of the following major parts: the LO in the form of long Josephson junction, SIS mixer, weakly coupled LJJ to model external signal, various transformers for better impedance matching of SIS mixer to LJJs, SQUID amplifier, and digital circuitry with bandpass delta-sigma modulator to test this structure as a whole. The LJJs/mixer circuitry is designed for better matching for the frequency of 350 GHz. The IF SQUID amplifier has been designed for the 1 GHz bandwidth. LO LJJ and weakly coupled LJJ have to be adjusted within 1 GHz difference around 350 GHz.

All components of the front-end chip have been tested. For the LO/SIS mixer structure, the current-voltage characteristic (I-V curve) of the LO, when a 7 mA current is applied through the control line, exhibited a pronounced flux-flow step at 0.75 mV. At the same time, the I-V curve of the SIS mixer irradiated by the LO, while its Josephson current is suppressed by current through control line and when LO is biased at flux-flow step, exhibits the photo-assisted tunneling step at hf_{RF}/e below the gap.

The main purpose for the SQUID amplifier on the front-end chip is to match low input impedance of bandpass modulator ($Z_{in} \ll 1 \Omega$) and high output impedance of SIS mixer ($Z_{out} \approx 50 \Omega$) with simultaneous filtering of all out of IF band signals. The realized amplifier consists of a front end and a power amplifier. The front-end is a simple two-junction voltage-mode SQUID to transform the input signal current into the stream of SFQ pulses. The pulses are amplified in power by passing over the Josephson transmission line (JTL) with exponentially increasing critical and bias currents of the Josephson junctions. The final stage of JTL and resistive-inductive network provide both the power gain and low-pass filtering in order to suppress the SQUID's Josephson oscillations and to refine the amplified signal.

The output signal is the current through a load inductor, which is inductively coupled to the modulator. The single Josephson junction connected in series with the load acts as the signal limiter.

The resonance frequency of mixer-amplifier tank circuit is measured to be 1.5 GHz. The SQUID amplifier has the following experimentally measured parameters: critical current and normal state resistance per junction of front end SQUID are $I_C=40 \mu\text{A}$ and $R_N=4 \Omega$, SQUID inductance is 40 pH, mutual input inductance is 400 pH and input inductance is 5 nH providing the amplifier noise temperature $T_N=90 \text{ K}$.

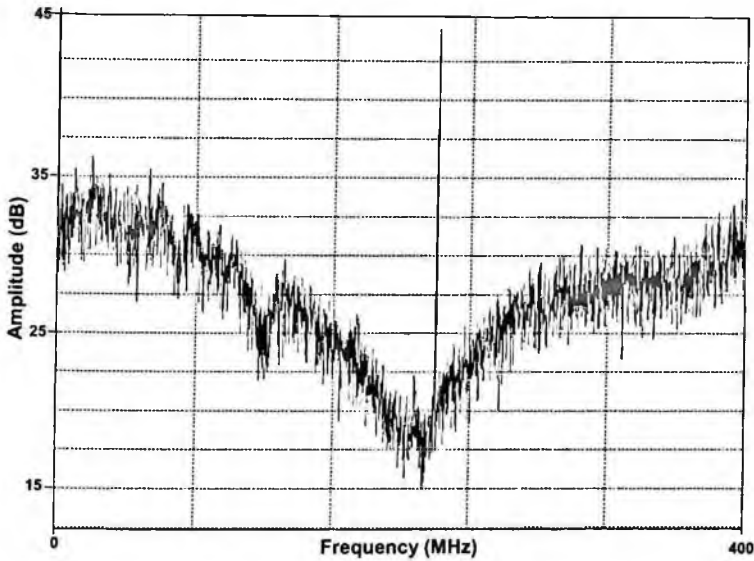


Fig. 3. Spectrum from bandpass delta-sigma modulator (1-bit digitizer) with clock signal at 800 MHz and input signal at 1020 MHz.

A first-order bandpass delta-sigma modulator with a center frequency of 1 GHz that used a lumped-element LC resonator is designed and fabricated on the front-end chip. The modulator uses implicit feedback – the switching of the clocked comparator automatically feeds $-\Phi_0$ back to the resonator – which is a fundamental advantage of superconductor delta-sigma modulators. In order to be able to verify performance of stand-alone front-end chip, the modulator is connected to read out digital circuitry, e.g., a high-speed output driver. The data from the modulator's output are converted to non-return-to-zero voltage-potential form, amplified and transferred to room temperature for additional amplification and processing.

Fig. 3 shows the measured spectrum of bandpass modulator when -20 dBm (-3 dB full scale) 1020 MHz input signal is sampled with 800 MHz clock. The undersampled spectrum clearly shows tone corresponding to 1020 MHz and expected noise shaping with dip of 15 dB in vicinity of 1 GHz.

4. Digital Signal Processing

The correlator circuit as the core component of our digital signal processor has been selected. The 128-channel autocorrelator that gives 8 MHz frequency resolution, e.g. appropriate for space-borne application has been designed and its correct operation has been verified. Our approach in design of 128-channel autocorrelator is to compose it from a 16-channel autocorrelator and 112-stage PRSR. The shift register should have number

of switches. By closing these switches data can be directed through or bypassing different digital delay lines, obtaining the data delay in increments of 16.

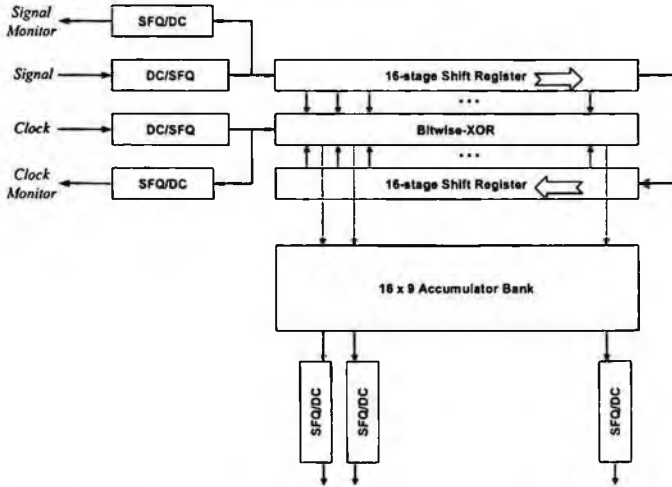


Fig. 4. The block diagram of 16-channel digital autocorrelator.

Fig. 4 shows the block diagram of 16-channel digital autocorrelator. The digital 16-channel autocorrelator forms a linear array with two main parts: digital delay lines with multipliers and an array of binary counters. The binary counters are composed of T flip-flop gates, outputting only the most significant bits. The individual components of the design - the circular shift register, XORs and T flip-flop gates are well known and have been reported to have very wide operating margins in many studies. Straightforward integration of these gates into an operational subsystem is possible only for the T flip-flop counters, due to their asynchronous mode of operation and unusually wide parameter margins. Integration of the XOR gates into the circular shift register is a much less trivial task, mostly due to numerous timing requirements.

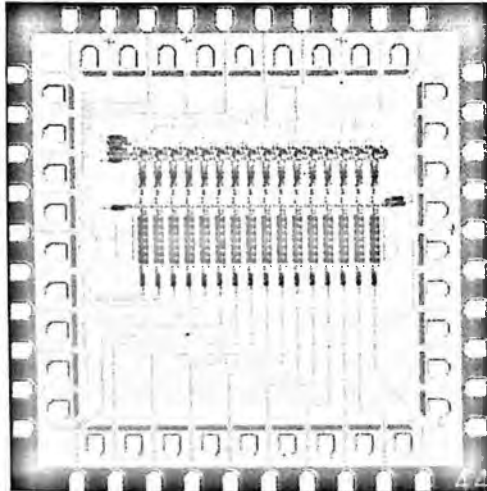


Fig. 5. The photo of 16-channel digital autocorrelator chip.

The main part of the autocorrelator is the digital delay line, which is based on a circular shift register with XOR multipliers built into every stage. The digital test at low speed is performed to verify the correct operation of a 32-stage circular shift register. The total number of D flip-flops in the shift register is 34:2 D flip-flops per each regular stage plus 2 in the “0-th” stage, where data makes a U-turn.

The 16-channel autocorrelator has been fabricated (the layout is shown in Fig. 5) and successfully tested using OCTOPUX test system capable of performing exhaustive low speed (500 Hz) tests of large digital and analog circuits.

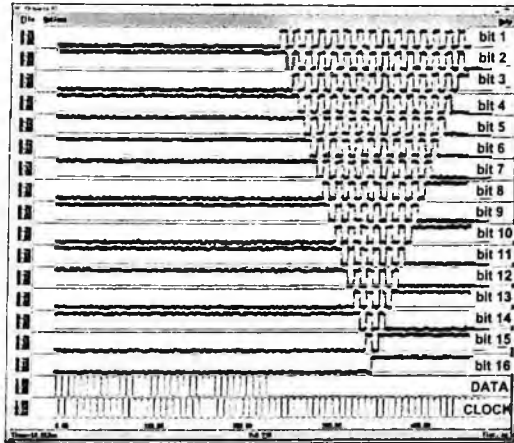


Fig. 6. Low-frequency testing of the 16 channel autocorelator. For the data and clock inputs each rectangular pulse corresponds to an SFQ pulse.

Fig. 6 shows the correct operation of the 16 channel autocorrelator without accumulator bank for the test sequence when a train of $2 \times 16 + 2 = 34$ ‘1’s (signal “DATA” in Fig. 6) is loaded into the circular shift register. High as digital ‘1’ and low as digital ‘0’ are interpreted. The correct operation of all 16 XOR gates (outputs O1-O16) for all possible combinations of inputs (“00”, “10”, “01”, “11”), as well as correct operation of the circular shift register under full load is demonstrated.

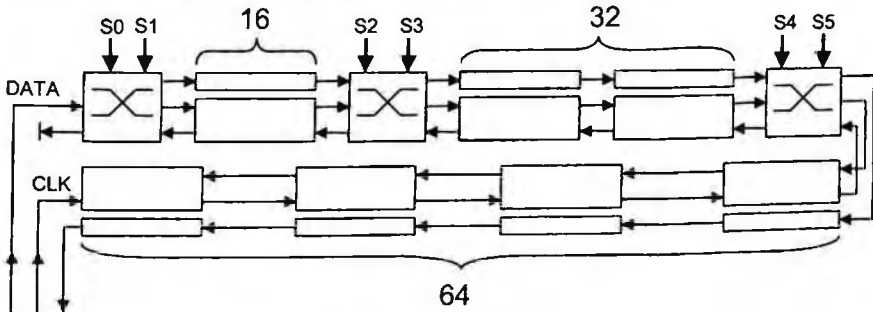


Fig. 7. The block diagram of programmable delay 112-stage shift register.

Fig. 7 shows the block diagram 112-bit PRSR. This shift register includes 112 stages divided into three chunks of 16, 32 and 64 stages together with 6 switches S0-S5. It employs more robust counter-flow design¹⁰, e.g. the stream of data and clock pulses flow in opposite directions. By applying DC current to different combination of the switches

we are able to choose any delay from 0 to 112 in increments of 16. The 112-stage PRSR has been fabricated and a photo of the chip is presented in Fig. 8.

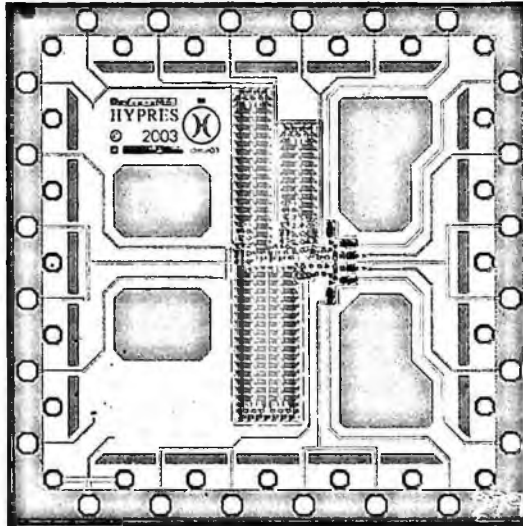


Fig. 8. The photo of 112-bit programmable shift register chip.

The operation of the shift register has been successfully verified by using OCTOPUX. As an example, fig. 9 presents test results for the 112-bit PRSR with correct operation of the shift register with switches S0, S3 and S5 closed while switches S1, S2 and S4 are opened, showing the delay of data at the output for 16 clock cycles. Fig. 9 shows the following three groups of data traces: data and clock inputs – inputs that have been generated by OCTOPUX and sent to the chip, data and clock monitors – traces experimentally measured immediately after appropriate entrances to the chip, and clock and data outputs- traces measured after clock and data propagated through all 112 stages of shift register. The data output exhibits the correct 16-clock cycles delay.

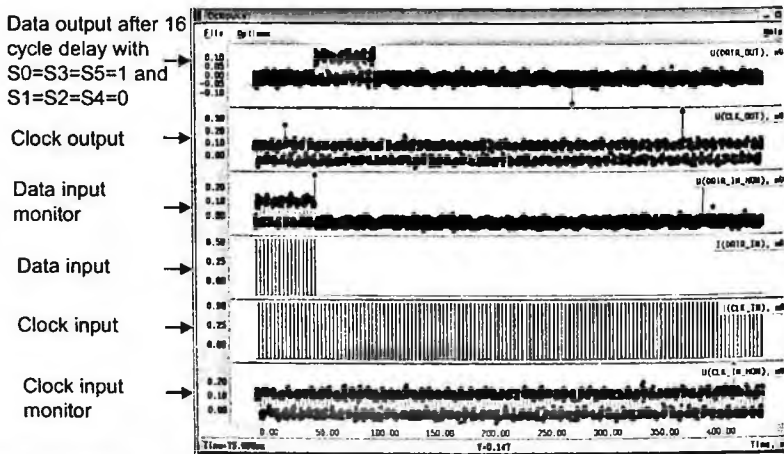


Fig. 9. Test results for the 112-bit PRSR with 16 clock cycles delay.

The correct performance of 16-lag autocorrelator and 112-lag PRSR has been verified. Their high-speed operation has been already realized before: autocorrelator at 11 GHz¹¹ and more complex shift register at 19 GHz¹⁰. After integration, our DSP circuit is expected to operate at slightly lower speed, still sufficient for intended applications – to acquire spectra of slow changing signals.

5. Packaging

A key element of the spectrometer package is the cryocooler – a cryogenic refrigerator needed to maintain the superconducting circuits at these low operating temperatures. Historically, superconducting systems required cooling with liquid helium, rendering its use expensive and inconvenient in terms of cryogen logistics. Recently, reliable closed cycle refrigerators (cryocoolers) have become commercially available. These are turnkey electrical machines with no liquid cryogens, analogous to home freezers and air-conditioning units, requiring only standard electrical power to deliver continuous operation for years. HYPRES has sold commercial SCE systems mounted on a commercial-off-the-shelf (COTS) cryocooler, which have exhibited extended performance superior to their liquid-He cooled counterparts. The technology for producing these cryocoolers has evolved over several decades, and recent developments now enable reliable products of unprecedented efficiency and reliability. Products utilizing superconductor materials for analog signal processing are commercially deployed and several hundred units have demonstrated over 3.5 Million hours of operation with mean-time-between-failures (MTBF) exceeding 99 years. Fig. 10 shows a concept sketch of a two-channel (capable of screening for the presence of chemical and biological agents at two different locations) spectrometer in a cryocooler fit for a standard 19" rack. The number of channels can be extended to as many as ten.

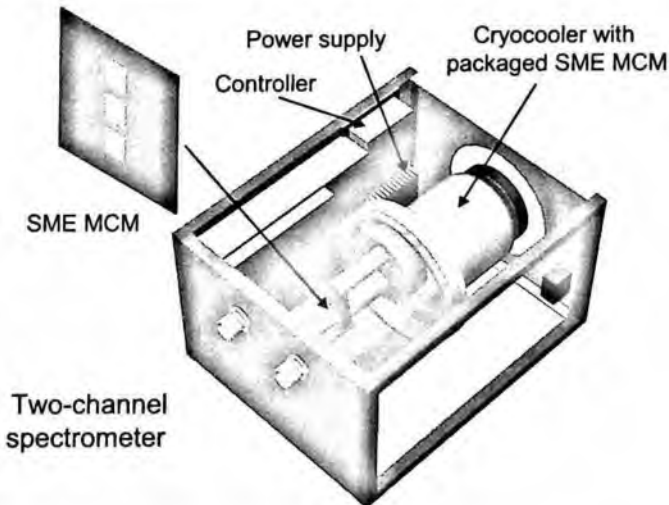


Fig. 10. The concept sketch of a two-channel spectrometer packaged in the cryocooler.

6. Conclusions

A compact ultrasensitive wideband superconducting mm/submm integrated digital spectrometer for chemical and biological agent detection is being developed. In addition, this spectrometer may be used in radio-astronomical research and remote monitoring of the Earth atmosphere. Correct performance of all integrated spectrometer analog and digital components, such as local oscillator/SIS mixer structure, IF SQUID amplifier, bandpass delta-sigma modulator, autocorrelator and programmable delay shift register have been successfully verified.

Assembled on a multi-chip module and packaged on COTS cryocooler the complete spectrometer offers integration of thin film analog components such as a quasioptical antenna with mixer, superconducting local oscillator and an IF SQUID amplifier together with superconducting digital circuitry.

7. Acknowledgments

This work was supported in part by Missile Defense Agency contract no. DASG60-03-C0007.

References

1. P. H. Siegel, THz technology, *IEEE Trans. Microw. Theory Tech.* **50**(3), 910–928, (2002).
2. J. R. Tucker and M. J. Feldman, Quantum Detection at Millimeter Wavelength, *Rev. of Mod. Phys.* **4**, 1055-1113, (1985).
3. D. L. Woolard, E. R. Brown, M. Pepper, and M. Kemp, Terahertz Frequency Sensing and Imaging: A Time of Reckoning Future Applications? *Proc. of IEEE* **93**(10), 1722-1743, (2005)
4. W. Chen, A. V. Rylyakov, Vijay Patel, J. E. Lukens, and K. K. Likharev, Superconductor digital frequency divider operating up to 750 GHz, *Appl. Phys. Lett.* **73**, 2817-2819, (1998).
5. O. M. Mukhanov, V. K. Semenov, I. V. Vernik, A. M. Kadin, D. Gupta, D. K. Brock, I. Rochwarger, T. V. Filippov, and Y. A. Polyakov, High resolution ADC operating up to 19.6 GHz clock frequency, *Supercond. Sci. Technol.* **14**, 1065-1070, (2001).
6. V. P. Koshelets and S. V. Shitov, Integrated Superconducting Receivers, *Supercond. Sci. Technol.* **13**, 53-69, (2000).
7. I. V. Vernik, D. E. Kirichenko, S. Sarwana and D. K. Brock, Integrated Millimeter/Submillimeter Superconducting Digital Spectrometer, *IEEE Trans. Appl. Supercond.* **15**, 419-422, (2005).
8. D. E. Kirichenko and I. V. Vernik, High Quality On-Chip Long Annular Josephson Junction Clock Source for Digital Superconducting Electronics, *IEEE Trans. Appl. Supercond.* **15**, 296-299, (2005).
9. The standard HYPRES Nb process flow and design rules are available via the HYPRES, Inc. website at <http://www.hypres.com>
10. O. A. Mukhanov, "RSFQ 1024-bit shift register for acquisition memory," *IEEE Trans. Appl. Supercond.* **3**, 3102-3113, (1993).
11. A. V. Rylyakov, D. F. Shneider and Yu. A. Polyakov, A Fully Integrated 16-channel RSFQ Autocorrelator Operating at 11 GHz, *IEEE Trans. Appl. Supercond.* **9**, 3623-3626, (1999).

COMPACT OPTICAL CHARACTERIZATION PLATFORM FOR DETECTION OF BIO-MOLECULES IN FLUIDIC AND AEROSOL SAMPLES

PETER KIESEL, OLIVER SCHMIDT, MICHAEL BASSLER, NOBLE JOHNSON

*Palo Alto Research Center, 3333 Coyote Hill Rd.,
Palo Alto, CA 94304, USA
Peter.Kiesel@parc.com*

An optical characterization unit based on fluorescence spectroscopy-on-a-chip is described. It comprises a compact fluidic platform that is integrated onto a chip-size spectrometer. The analyte is continuously excited within a novel waveguide. Fluorescence spectra are recorded as the analyte traverse the detection area.

In order to achieve a strong interaction between excitation light and analyte we use an anti-resonant waveguide, in which the light is guided within the target-containing medium, thereby enabling a continuous excitation of a large volume. The excitation light is guided in the lower-refractive-index fluid when the light is coupled into the waveguide at an appropriate angle.

Compact spectrometers can be integrated along the fluidic channel. The spectrometers are composed of a detector array which is coated with a linear variable band-pass filter. The filter converts the spectral fluorescence information into a spatially dependent signal that is analyzed by the detector array. These chip-size spectrometers are especially applicable for characterization of moving analytes.

Keywords: spectroscopy-on-a-chip; chip-size spectrometer; optofluidics.

1. Introduction

Many applications require fast, sensitive, and specific detection methods to identify nanoparticles, micro-organisms, bio agents, and toxins in water, blood, food, aerosols and other specimens. This generally involves a laboratory procedure that is expensive, time consuming, and requires skilled personnel. Yet many applications actually require a compact, fast and automated detection system for point-of-care detection. It is generally a necessity to detect the analyte of interest “on-the-fly”, that is, as it is being transported in order to allow for continuous and real-time detection. Optical detection methods offer high sensitivity, but current approaches capture only a “snapshot” of the moving particle which yields limited information. Furthermore, most detection schemes have limited performance due to weak interaction between the excitation light and analyte, as well as the high cost and large size of the typical optical instruments (e.g., commercial spectrometers).

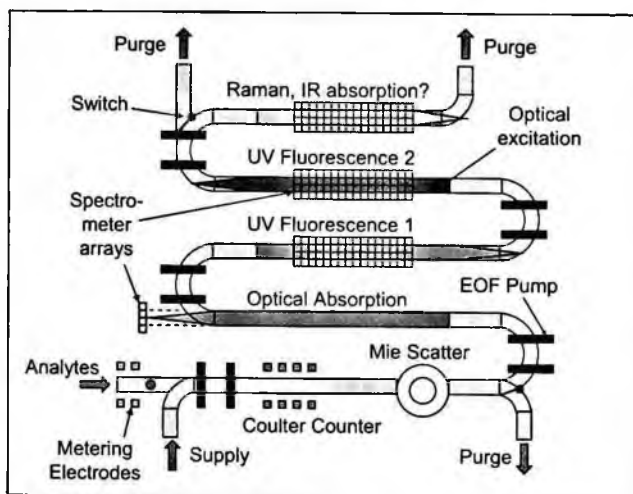


Fig. 1. Compact identification platform based on orthogonal optical methods. The analyte is characterized step by step during the movement through the characterization unit

The ultimate goal of our research is a compact identification platform based on orthogonal optical methods. As sketched in Fig. 1, various chip-size wavelength detectors designed for different wavelength ranges and enabling different optical characterization methods will be aligned along a fluidic channel that is acting as an anti-resonant waveguide. This will allow real-time analysis of analytes (e.g., bio molecules) during their passage through the fluidic system. One advantage of this system is the possibility of interactive detection schemes. With fast data acquisition, the result of a prior characterization can trigger or influence a certain subsequent measurement. For example, the light scattering and fluorescence experiment may indicate the presence of a certain analyte. This could trigger a more specific investigation for this species by using, for example, multi-wavelength fluorescence, Raman or IR spectroscopy. This might be particularly useful for systems with an integrated switch or sorter, which is triggered by previous results and allows for purging or accumulation of the analyte in a certain investigation channel (as indicated in Fig. 1). This feature is especially interesting for refined investigations like Raman and Far-IR spectroscopy, which are known to provide molecular fingerprints of the target but usually provide only very weak signals. For these methods an enhanced photonic interaction and larger integration times are highly desirable features.

This paper describes an optical subsystem that is able to record fluorescence spectra of an unknown analyte in water "on the flow" and can be assembled into an extremely compact system. Our approach integrates a method for enhanced light-target interaction with a chip-size wavelength detector for spectral characterization of a moving analyte. Unlike most other approaches for detecting and identifying an analyte, ours does not require concentrating or immobilizing suspect particles for interrogation. Rather, it takes advantage of the general necessity to detect such particles in real time as they are moving.

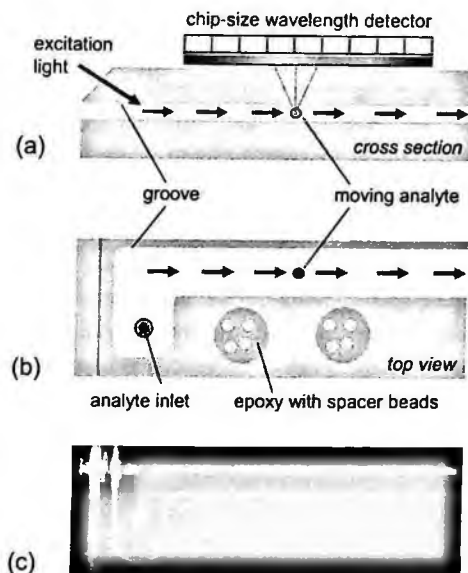


Fig. 2. a) Schematic cross section of our characterization platform comprising a chip-size wavelength detector and a fluidic channel that incorporates an anti-resonant waveguide. b) Top view of the fluidic chamber illustrating the key features of the unit and the fabrication process c) Photograph of the fluidic chamber: the channel is emitting fluorescence light from an Alexa dye which is anti-resonantly excited along the channel.

Our concept permits warning (triggering) of suspicious particles and class identification simultaneously. This information can be used to initiate further processing in a second stage, such as identification, removal, disinfection, capturing or sorting of agents.

2. Fabrication of the Characterization Platform

The compact optical characterization platform is schematically illustrated in Fig. 2a. The main building blocks of the fluidic platform are an anti-resonant waveguide to enhance light-target interaction and a chip-size spectrometer. These components are discussed in the sections below. The fabrication of the fluidic platform can be performed with various technologies including injection molding, hot embossing or laser micromachining. It requires only conventional, readily available technology which should enable a cost effective manufacturing process.

Here we briefly describe the fabrication of the prototype shown in Fig. 2a. The chamber is composed of two $25 \times 75 \text{ mm}^2$ plastic (acrylic) slides each with a thickness of 1.5mm. The distance between the slides is defined by glass micro-spheres with a diameter of $100 \mu\text{m}$. A blend of micro-spheres and UV curable epoxy has been used to assemble the chamber and enable a precise spacing of the slides. The channel structure on the slides was defined by creating $50 \mu\text{m}$ grooves with a laser cutter (see Fig. 2b). In a second glue-step UV epoxy with a low viscosity was sucked in from the side of the chamber by capillary forces. The epoxy penetration stopped at the provided grooves and

thereby defined and sealed the fluidic channel. The channel of the example shown in Fig. 2 is 5mm wide and 70mm long. In order to provide an inlet and outlet for the analyte pinholes through the cover slides were thrilled with the laser cutter. One facet of the cover slides was polished for an angle of 45°. This enabled efficient light coupling into the fluidic channel that represents an anti-resonant waveguide, as discussed in the next section. In a final step the fluidic chamber has been mounted on top of the chip-size spectrometer.

For many of the measurements presented in the next sections the fluidic chamber was imaged onto the chip-size spectrometer rather than mounted in order to enable more flexibility in changing the measuring parameters. Fig 2c shows a photograph of the fluidic chamber. The fluidic channel emits fluorescence light from a dye-containing fluid. The dye was excited within a large volume by anti-resonant waveguide excitation, as explained in Sect. 4.

3. Chip-size spectrometer

Compact spectrometers which can be integrated along the fluidic channel are crucial components for the spectroscopy-on-a-chip concept explained above. Most presently used spectrometers are bulky and expensive systems in which the incident wavelength spectrum is split into its components with gratings or prisms in order to determine the intensity of the various wavelength fractions. Many approaches for miniaturization exist, based on various technologies, e.g., MEMS, but they are still very expensive, not compact enough and most importantly not suited for integration into small systems as required in our case. We have identified concepts that enable the realization of a compact and low-cost spectrometer that is especially well suited for the spectral characterization of moving particles. The size of the spectrometer can be only slightly larger than that of the detector array itself, and fabrication requires only conventional, readily available processing technology which should enable a cost effective manufacturing process.

The concepts are based on light sensitive elements, that are usually used to provide spatial information (e.g., detector array, CCD chip, CMOS detector array or avalanche photo-diode array), in combination with mirrors or cavities with inhomogeneous (spatially-dependent) transmission/reflection properties. The laterally varying transmission/reflection properties of the coating on top of the light-sensing element define a correlation between position and wavelength. Therefore, the spatially dependent signal of the detector contains information about the incident spectrum. For instance, the output signal of a coated detector array will provide the wavelength spectrum of the incident light. One can select from a large variety of light sensing elements and coatings covering the whole spectral range from the deep UV to the Far-IR. Systems demanding very high wavelength resolution (e.g., absorption spectroscopy or Raman spectroscopy) as well as system for broadband application (e.g., fluorescence spectroscopy) can be realized by choosing the appropriate coating.

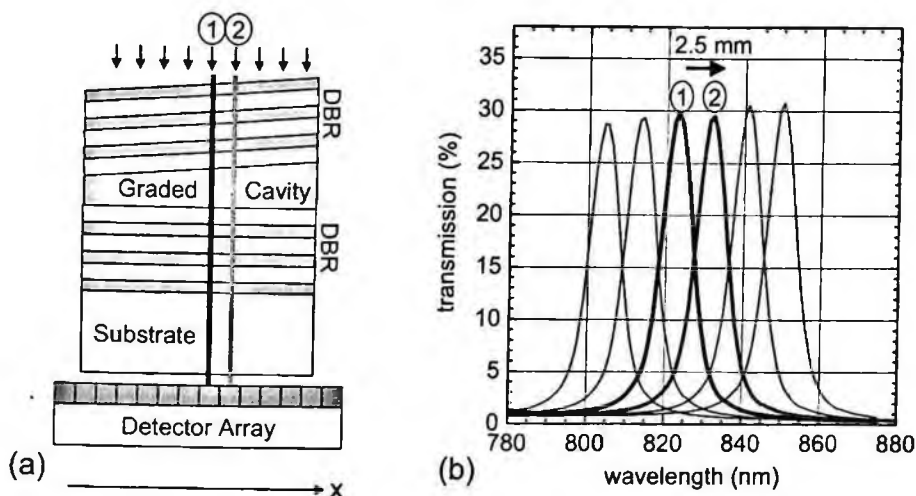


Fig. 3. (a) Filter with laterally varying transmission properties transforms spectral into spatial information. A photodetector records the intensity distribution behind the filter. (b) Transmission properties of a multiple dielectric Fabry-Perot-Cavity at 6 different positions of the filter. A position change of 2.5 mm induces a 10 nm transmission shift, resulting in a 4 nm/mm gradient.

To illustrate, a detector array was covered with a laterally-graded Fabry-Perot structure (Fig 3a). In this case the coating was deposited onto a thin substrate and afterwards packaged together with the detector array; it can also be directly deposited on the light sensing elements during the fabrication process of the detector array. Each detector pixel is sensitive only to a narrow spectral range. As indicated in Fig. 3b adjacent pixels of the array are sensitive to a slightly shifted spectral range. A parallel read-out of the array provides the spectrum of the incident light. In order to prevent errors due to inhomogeneous illumination certain rows of the detector array remain uncovered and thus serve to determine the spatial distribution of the light intensity (reference spectrum).

Figure 4 illustrates the spectrometer concept. For this demonstration we have covered a CCD camera with a linear variable filter which spans the spectral range from 400 to 700nm. The filter has a spectral gradient of about 32nm/mm. This means that the transmission maximum shifts by 32nm between two locations which are 1mm apart. The FWHM of the transmission peak of the filter is about 10nm. This linear variable filter was attached on top of a CCD chip. We illuminated the coated CCD chip with a tunable monochromatic light source (halogen lamp spectrally filtered by a monochromator). After calibration the usual read-out of the CCD chip can be displayed directly as a wavelength spectrum. The spectral resolution (FWHM) in this example was about 4nm. Figure 4b shows a logarithmic plot of the 540nm curve in order to illustrate the excellent stray light suppression. Even for this configuration, which was not optimized, the stray light is suppressed by almost two orders of magnitude.

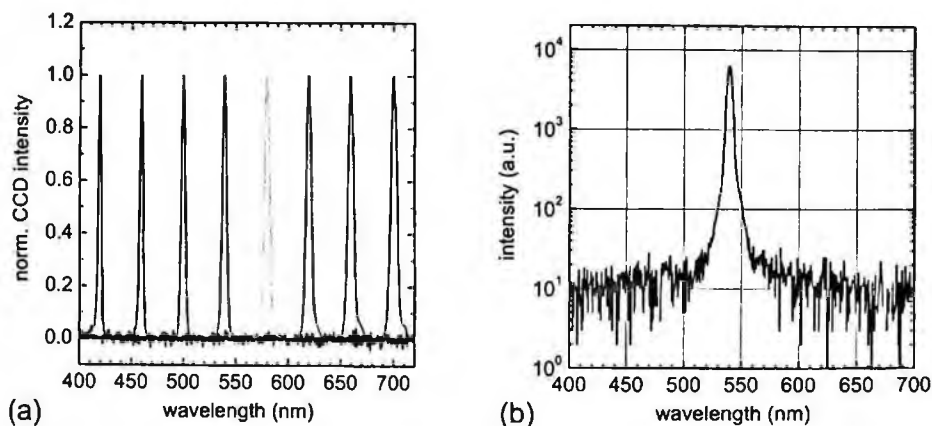


Fig. 4. (a) Various wavelength spectra taken with a CCD chip coated with a laterally inhomogeneous Fabry-Perot coating. As a light source we have used a halogen lamp spectrally filtered by a monochromator. (b) Logarithmic plot of the 540nm spectra in order to illustrate the stray light suppression.

The most convincing advantage of the spectrometers described here is their compactness and their multi-functional use in systems. Different rows of a 2D array can be provided with different coatings addressing different wavelength ranges. This way the data obtained from a single CCD chip can provide a detailed analysis of the incident light over a broad spectral range. In addition, uncoated rows within the CCD chip can be used to provide a spatially resolved real-time reference signal.¹ The compactness of the detectors in combination with the fast data acquisition allows for an interactive detection scheme. Subsequent or adjacent detector arrays can exchange information or trigger events.

4. Large-area fluorescence excitation

The fluorescence detection concept described in Sect. 1 requires a continuous excitation of the analyte as it moves across the detection area. In order to enable fluorescence excitation over an extended area, conventionally evanescent field excitation is used. In this case the light is guided within a traditional optical waveguide within a layer of high refractive index adjacent to the fluidic analyte. As a consequence the interaction between light and analyte is typically very weak because the evanescent field of a guided mode carries only a small fraction of the intensity. In order to improve this interaction we are using an anti-resonant waveguide, in which the core region has a lower refractive index than the cladding layers. With this concept the light can be guided within the target-containing medium, thereby enabling an extended interaction length. Anti-resonant waveguides are especially compatible with a fluidic biosensor because the fluidic channel itself can be used as the core of the anti-resonant waveguide. Of course, this approach can also be used in various other configurations, e.g., aerosols in glass capillary or liquid films between glass slides. The anti-resonant waveguide concept is particularly well

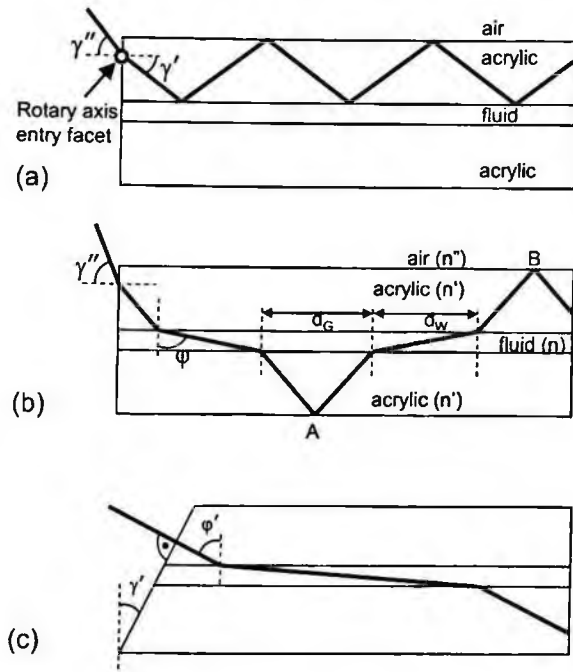


Fig. 5. Schematic illustration of light rays propagating through the waveguide structure at different coupling angles (a) below and (b) above the critical angle. Figure (a) also indicates the position of the rotary axis of measurement setup. (b) From the path length of light in water and glass the confinement factor can be calculated. (c) Coupling mechanism for excitation of anti-resonant waveguide modes for tilted entry facet.

suited for a multi-signal analysis approach, since the performance is relatively unaffected by changes in both wavelength and film thickness.

Figure 5 shows the structure of an anti-resonant waveguide which consists of a liquid film which is sandwiched between two plastic slides that have a higher refractive index than the liquid film. Understanding light propagation in waveguides is commonly achieved by examining the optical modes that are obtained by solving the Helmholtz equation. From this analysis the intensity distribution over the cross section of the waveguide can be obtained, and the confinement factor of the light within a given layer of the structure can be determined. Calculating the optical modes yields a number of conventional and anti-resonant modes, the latter revealing very high light confinement in the liquid compared to the conventional modes (e.g., up to 90%). The anti-resonant modes can be excited by coupling light under a certain angle into the structure. For simplicity, this is explained in the following within the ray picture of light.

Figure 5a and b shows light entering the structure through the entry facet and its propagation at different coupling angles. For angles below a critical angle γ_c , light is guided in the upper glass slide by total internal reflection (TIR) as in a conventional waveguide (see Fig. 5a). No light is coupled into anti-resonant waveguide modes and no light is guided within the water. The quantity γ_c is determined by Snell's law as follows:

$$\sin \gamma' = \frac{n''}{n'} \sin \gamma_c'' \quad (1)$$

At the critical angle light at the glass/liquid interface is no longer blocked by TIR and anti-resonant guided optical modes in the liquid are excited. Note that the critical angle is independent of the layer thickness.

The light is nearly parallel to the channel if the coupling angle is close to the critical angle. For this case we obtain the largest interaction because the ray propagates a long distance through the water layer before coupling into the opposite glass slide. The length inside the water layer becomes shorter with increasing coupling angle. The ratio between the horizontal distance that the beam travels through the water and distance it travels through the entire stack from A to B (see Fig. 5b) can be interpreted as the confinement factor:

$$\Gamma = \frac{d_w}{d_w + d_G} \quad (2)$$

The confinement factors obtained from this simple ray picture are in good agreement with the experimentally determined values. A more detailed discussion of anti-resonant waveguides was published by Schmidt *et al.*^{2,3}

Light coupling through a regular end facet as shown in Figs. 5a and b is not very practical. For the case of a water film sandwiched between two acrylic slides, the critical coupling angle is about 41°. Coupling through the end facet at such a large angle would result in considerable reflection loss. The problem is overcome by tilting the entrance facet as shown in Fig. 3c. Maximum coupling efficiency is achieved with an incident beam perpendicular to the entrance facet. Thus, the facet is tilted by about 26°.

As discussed in Sect. 2, we have designed our fluidic chamber as an anti-resonant waveguide with a tilted end facet. Figure 2c shows a picture of the fluidic chamber filled with Alexa Fluor 594 fluorescent dyes (maximum emission at 620nm) dissolved in water. The 532nm emission from a frequency-doubled Nd:Yg laser was coupled into the anti-resonant waveguide structure through the polished entry facet. The coupling angle was just above the critical coupling angle in order to achieve a large overlap (confinement factor) between excitation light and the dyes in the fluidic channel. As shown in Fig. 2c, the dyes are homogeneously excited over the entire 7 cm length of the fluidic channel, which demonstrates the large volume fluorescence excitation afforded by using an anti-resonant waveguide.

5. Fluorescence spectroscopy-on-a-chip

We have measured the fluorescence spectrum emitted from the anti-resonantly excited Alexa dye in the fluidic channel with a chip-size spectrometer as discussed above and with a conventional spectrometer (see Fig.6). The spectra are normalized to their 633nm intensity in order to ease comparison. Both show a sharp peak at 532nm caused by scattered excitation light and a broad peak around 633nm from the Alexa fluorescence. In

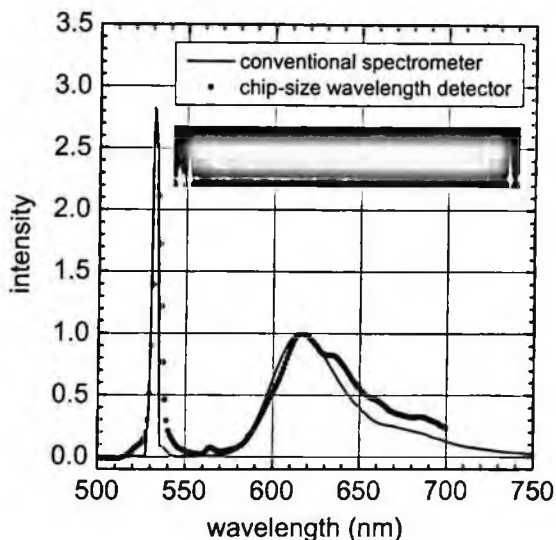


Fig. 6. Comparison of the fluorescence spectra from Alexa dye obtained measured with conventional spectrometer and a chip-size wavelength detector

these measurements no filters have been used in order to block the scattered excitation light. This result confirms the inherently great stray light suppression of the anti-resonant waveguide geometry. The excitation light is guided within the fluidic channel and significant stray light only originates from fabrication imperfections and impurities within the fluidic channel. This can be seen from the distributed green spots within the red emission in the picture from the fluidic chamber (see inset in Fig. 6). In addition significant scattered light from the green excitation is observed at the entrance facet, the epoxy-stop grooves (see also Fig. 2b) and around the inlet and outlet.

Overall, the spectra recorded with the chip-size spectrometer and the conventional spectrometer are in good accordance. The minor differences will be discussed in the following. In the spectrum recorded by the chip-size spectrometer, the laser line appears slightly broader and weaker. This is due to the lower spectral resolution of the chip-size spectrometer which is defined by the transmission function of linear variable filter. The spectra measured with the chip-size spectrometer also show a shoulder on the long wavelength side of each peak. Since the incident light is not collimated, the angle dependence of the filter transmission affects the spectrum asymmetrically. Simulations for this setup resulted in a red shift of about 10nm which is in good agreement with the measurement. If correction is necessary for particular applications, appropriate optics (e.g., micro optics) between fluorescence chamber and filter can be introduced to collimate the light. However, for the application we have in mind ("class identification of analyte based on multicolor fluorescence spectroscopy") there is not even a need for correction since all analytes are affected equally and a spectral resolution of about 15nm seems to be sufficient for class identification.⁴

As a last feature we will discuss the slight oscillations added to the spectrum measured with the chip-size spectrometer. These are caused by Fabry-Perot oscillations within the protection layer on top of the CMOS chip in the camera. An appropriate anti-reflection coating on top of the protection layer will efficiently remove this artifact. Moreover, for a real device the protection layer will be replaced by depositing the graded filter directly on top of the CMOS chip.

6. Summary

We have described a spectroscopy-on-a-chip platform that is comprised of a chip-size spectrometer and fluidic chamber that incorporates an anti-resonant waveguide. The waveguide enables guiding light within the fluid containing the analyte. Therefore, the particles can be continuously excited as they traverse the chip size spectrometer. The system is capable of recording the fluorescence spectrum of an analyte while it is moving along the detector chip. The measured fluorescence spectra are in good agreement with reference spectra taken with a conventional spectrometer. The concept will enable a compact characterization platform for point-of-care detection and analysis of pathogens.

7. Acknowledgments

The authors are pleased to acknowledge helpful discussions with Meng Lean, and Armin Völkel. The research was funded by ONR under grant N00014-05-C-0430 monitored by Jeremy Walker, Janet Jensen, and Paul Armistead.

References

1. O. Schmidt, P. Kiesel, M. Bassler, N. M. Johnson, "Chip-size wavelength detectors", *submitted for publication to International Journal of High Speed Electronics and Systems (same Vol.)*
2. O. Schmidt, M. Bassler, P. Kiesel, N. M. Johnson, G. H. Döhler, *Appl. Phys. Lett.* **88**, 151109 (2006).
3. O. Schmidt, M. Bassler, P. Kiesel, O. Wolst, G. H. Doehler, *SPIE Proc.* **6094**, 80 (2006).
4. M. Bassler, O. Schmidt, P. Kiesel, N. M. Johnson, "Class Identification of bio-molecules based on multi-color native fluorescence spectroscopy", *submitted for publication to International Journal of High Speed Electronics and Systems (same Vol.)*

RESONANT TERAHERTZ SPECTROSCOPY OF BACTERIAL THIOREDIXIN IN WATER: SIMULATION AND EXPERIMENT

ALEXEI BYKHOVSKI

*Dept. of Electrical and Computer Engineering, University of Virginia
351 McCormick Rd, P.O. Box 400743, Charlottesville, VA 22904, USA
ab4k@virginia.edu*

TATIANA GLOBUS

*Dept. of Electrical and Computer Engineering, University of Virginia
351 McCormick Rd, P.O. Box 400743, Charlottesville, VA 22904, USA
tg9a@virginia.edu*

TATYANA KHROMOVA

*Dept. of Electrical and Computer Engineering, University of Virginia
351 McCormick Rd, P.O. Box 400743, Charlottesville, VA 22904, USA
tbk4b@virginia.edu*

BORIS GELMONT

*Dept. of Electrical and Computer Engineering, University of Virginia
351 McCormick Rd, P.O. Box 400743, Charlottesville, VA 22904, USA
gb7k@virginia.edu*

DWIGHT WOOLARD

*U.S. Army Research Laboratory, Army Research Office
Research Triangle Park, NC 27709, USA
dwight.woolard@us.army.mil*

The experimental and computational study of bacterial thioredoxin, an E. coli protein, at THz frequencies is presented. The absorption spectrum of the entire protein in water was studied numerically in the terahertz range (0.1 – 2 THz). In our work, the initial X-ray molecular structure of thioredoxin was optimized using the molecular dynamical (MD) simulations at room temperature and atmospheric pressure. The effect of a liquid content of a bacterial cell was taken into account explicitly via the simulation of water molecules using the TIP3P water model. Using atomic trajectories from the room-temperature MD simulations, thioredoxin's THz vibrational spectrum and the absorption coefficient were calculated in a quasi harmonic approximation.

For our terahertz transmission measurements, we used solutions of thioredoxin in distilled water obtained from Sigma. The experimental and simulated signatures are correlated and dominant peaks are close in frequencies. The results of this study demonstrate that terahertz spectroscopy is a promising tool in generating spectral data for cellular components of bio agents such as bacterial cells and spores.

Keywords: THz, absorption, thioredoxin, *E. coli*.

1. Introduction

A systematic approach to the THz fingerprinting of biological (bio) warfare agents and their simulants requires a detailed experimental and theoretical study of cellular components. Therefore, contributions from significant components into whole cell THz spectra have to be analyzed. Recently, we reported our experimental and theoretical THz spectra studies of genetic components of bio warfare agents and their simulants^{1, 2}. In particular, THz signatures of bacterial chromosomal DNAs from *E. coli*, *B. subtilis* and *B. anthracis* were studied¹, as well as transfer RNA specific for tyrosine and valine from *E. coli*². The developed modeling and simulation approach has enabled us to accurately predict the important THz absorption features of the nucleic acids in 10-25 cm⁻¹ range from the studied bio warfare agents and their simulants.

Proteins contribute ~ 50% dry weight of *E. coli* compared to ~25% contribution from nucleic acids (DNA and RNA)³. In this work, we report the results of our studies of the THz signature for a bacterial thioredoxin, a widespread protein from inside the cell of a potential bio warfare agent *E. coli*. The results of our modeling and simulations are compared to our experimental spectra of thioredoxin from *E. coli*.

2. Model Description

Thioredoxin from *E. coli* has been studied extensively using biochemical, spectroscopic and X-ray diffraction techniques. The protein consists of a single polypeptide chain of 108 amino acid residues of known sequence. The protein has been cloned and expressed. Thioredoxin of *E. coli* is a compact molecule with 90% of its residues in helices, beta-strands or reverse turns. This protein transports electrons via an oxidation-reduction active disulfide⁴. The oxidized form thioredoxin-(S₂) is reduced to thioredoxin-(SH)₂. In particular, this protein was found to participate in the reduction of ribonucleotides to deoxyribonucleotides. In Fig. 1, the optimized structure is shown with a carbon backbone for clarity only. The molecule consists of two conformational domains, connected by two helices. The beta-sheet forms the core of the molecule packed on either side by clusters of hydrophobic residues. Helices form the external surface. We used a crystal structure of the oxidized form of thioredoxin from *Escherichia coli* that has been refined by the stereochemically restrained least-squares procedure at 1.68 Å resolution⁴.

We developed two models for the THz spectra simulation of thioredoxin. Amber's Generalized Born approximation⁵ was selected for our initial draft model. This approximation takes into account the effects of solvation in water implicitly, using a

dielectric model. In this model, an analytical interpolation formula for the dielectric function of a macromolecule in a solution is used. The covalent bond energy, covalent angle energy, proper and improper torsions, non-bonded interactions including electrostatic, and Van der Waals interactions were taken into account using AMBER 03 protein force field. Molecular dynamical simulations and quasi harmonic analysis were performed to extract normal vibrations and polarization vectors from the calculated atomic trajectories.

It has been noted that water molecules are involved in establishing intramolecular hydrogen-bonding interactions between protein atoms and thus serves as an integral part of the folded protein structure. The effect of the water molecules on simulated THz spectra of the thioredoxin was investigated, since water can substantially affect the dynamics of the thioredoxin. To do this, an explicit simulation of water molecules was also conducted using a more advanced model which is described below.

The molecular structure was optimized using the molecular dynamical (MD) simulations at room temperature and atmospheric pressure. Again, the covalent bond energy, covalent angle energy, proper and improper torsions, non-bonded interactions including electrostatic, and Van der Waals interactions were taken into account using the AMBER 03 protein force field. In contrast with our initial model, the effect of the liquid content of the bacterial cell was taken into account explicitly via the simulation of water molecules using the TIP3P water model^{5, 6}. This water model represents a rigid water



Fig. 1. The structure of thioredoxin. Ribbon represents the carbon backbone. ZY plane (left), ZX plane (middle), XY plane (right). The asymmetry of the molecule is clearly seen.

monomer with three interaction sites⁶. Coulomb and Lennard-Jones potentials are taken into account in the TIP3P model. A TIP3P water monomer geometry is represented by O-H distance of 0.09572 nm and HOH angle 104.52 deg. Pre-equilibrated box of water provided in Amber was used to build initial atomic coordinates for the simulated system. In order to use the constant pressure simulations, the periodic boundary conditions were imposed.

In Fig. 2, the initial X-ray structure (left) is compared with the structure optimized using the generalized Born model (middle) and the explicit water model (right). The

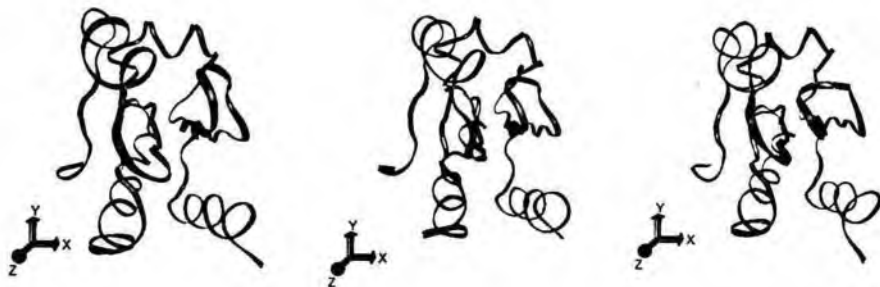


Fig. 2. Initial X-ray carbon backbone structure of the thioredoxin (left). The end structure for generalized Born-based MD simulations (middle). The end structure for explicit water MD simulations (right). All structures are shown at the same orientation.

carbon backbone of the thioredoxin is shown for clarity only. All structures in Fig. 2 are in the same orientation. The conformational domains and the backbone structure remain essentially unchanged during optimization. Since conformations of bio molecules are linked to their bio function, the stability of the conformation during the simulations confirms that the optimization procedures were correct and the biologically valid conformation of the thioredoxin was well preserved.

For both models, we used atomic trajectories from our room-temperature MD simulations to calculate thioredoxin's THz spectra in a quasi harmonic approximation. The absorption coefficient was calculated for different orientations of the molecule with respect to the electric field polarization.

3. Results of Simulations

The results for the calculated absorption coefficient averaged over all orientations are plotted in Fig. 3 both for the generalized Born implicit water model and the explicit water model. There is a substantial difference between the two models, especially at 15 cm^{-1} , pointing to the significance of the effect of water molecules on THz dynamics. Significant decreases in the absorption peak intensities for the explicit water model can probably be attributed to the damping of vibrations in the liquid environment. Still, multiple absorption peaks at approximately 12.8 cm^{-1} , 14.4 cm^{-1} , 15.3 cm^{-1} , and 22.5 cm^{-1} are predicted in the $10\text{-}25\text{ cm}^{-1}$ frequency band by this more realistic model. We will discuss the results obtained using this model in more detail, including the effect of orientation. Fig. 4 shows our calculated absorption coefficient spectra for x- and y-polarized light, and Fig. 5 shows the results for (x+y) and z-polarized light. According to our simulations, there is a substantial orientation effect in the absorption of thioredoxin. As it is shown in Fig. 5, we can expect very distinguishable spectra for z-polarized light compared to x+y polarization. The peak at 22.5 cm^{-1} mostly comes from z-component (Fig. 5). The x-polarized light does not show any contribution into 11 cm^{-1} peak (Fig. 4).

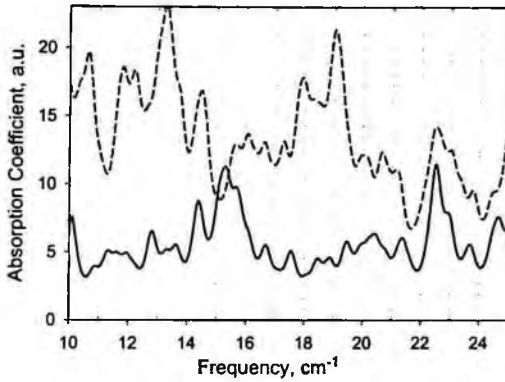


Fig. 3. Thioredoxin of *E. coli*'s averaged spectrum calculated using Amber, quasi harmonic analysis. Dashed line- generalized Born approximation, solid line- MD with explicit water.

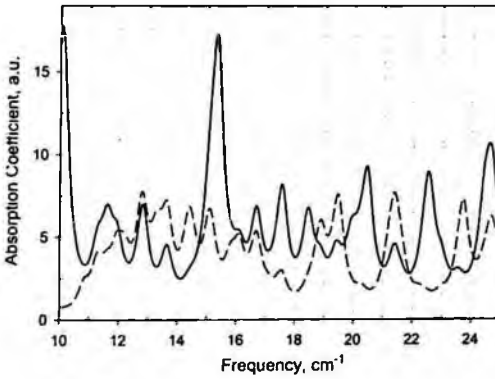


Fig. 4. The absorption coefficient for X- polarized light (solid line) and Y-polarized light (dashed line). Thioredoxin, Amber. MD with explicit water.

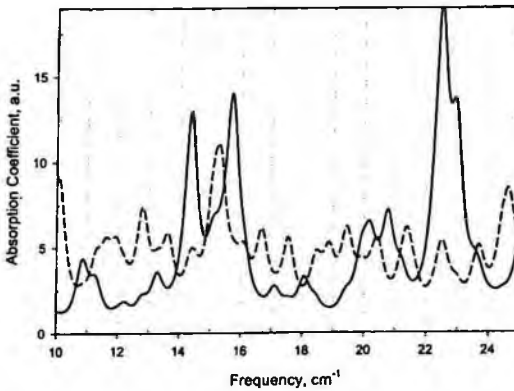


Fig. 5. The absorption coefficient for (X+Y)- polarized light (dashed line) and Z- polarized light (solid line). Thioredoxin, Amber. MD with explicit water.

On the other hand, z-polarized light has a distinct contribution into a peak at 11 cm^{-1} (Fig. 5). These results are important, since the orientation seems to affect the experimental transmission spectra as well. The molecule is not symmetric (see Figs. 1 – 2). The direction of z-axis (in blue) is somewhat more closely associated with directions of carbon atoms' backbones. Therefore, an anisotropy in distribution of biological molecules can be detected using the THz spectroscopy.

4. Terahertz Transmission Measurements

For our THz transmission measurements, we used solutions of thioredoxin in distilled water obtained from Sigma. The sample preparation technique and experimental procedure has been described in Ref. ⁷. Samples of diluted water solutions with the concentrations between 0.1 mg/ml and 3 mg/ml of proteins were measured. Several samples were made from each solution. Samples are in the form of a cell assembled from two polycarbonate films and a spacer ($12\text{ }\mu\text{m}$ thick) placed between them, with $\sim 10\text{ }\mu\text{l}$ solution inside the spacer. Each sample was measured several times and the resulting spectra were calculated against the background of similar cells made from pure de-ionized water between two PC films.

Figure 6 shows resulting spectra for 3 different samples of the thioredoxin in water with the same concentration, and the averaged result. We have a rather good reproducibility in the position of almost all resonance modes. In current measurements, we tried to avoid orientation effects. Since there was no preferable direction for the solution to flow during the preparation procedure, we expected to get the results averaged over all orientations of the molecules in the plane of the supporting film. However, some orientation effects are present. One of the possibilities to obtain reliable data is to average results from many samples and from different concentrations.

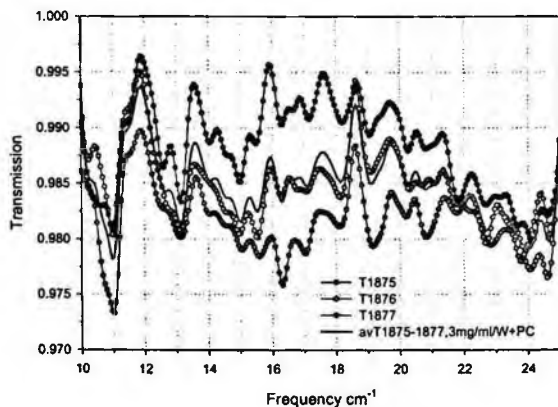


Fig. 6. Spectra of 3 samples of thioredoxin water solution with the concentration 3 mg/ml (between 2 PC films against the background of water between 2 PC).

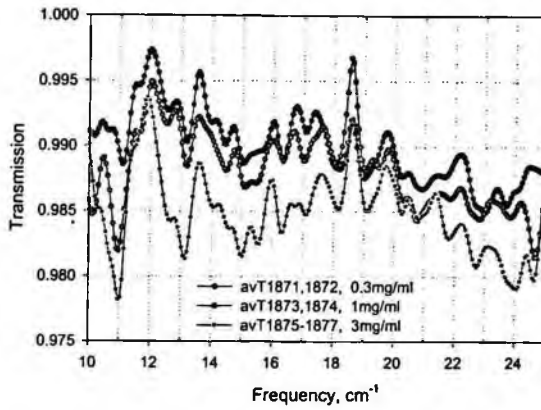


Fig. 7. Comparison of spectra of samples of thioredoxin water solutions with 3 different concentrations (between 2 PC films against the background of water between 2 PC).

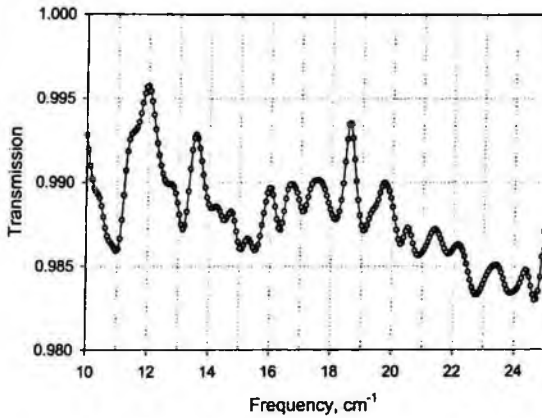


Fig. 8. Averaged spectrum of thioredoxin water solution.

Figure 7 demonstrates that there is the reproducibility of resonance features in the spectra of samples from 3 different concentrations. At last, the averaged results for thioredoxin at all concentrations are demonstrated in Fig. 8. The transmission spectra presented in Fig. 8 was recalculated for absorption which was compared with the modeling and simulation results (see the next Section).

5. Comparison Between the Modeling and Experimental Results

In Fig. 9, the absorption coefficient (blue dotted line) is plotted for the experiment (dotted line), the MD simulation with explicit water for non-polarized or averaged over all polarizations (solid line) and z-polarized light (dashed line). The damping coefficient of 0.5 cm^{-1} was used to account for dissipation processes in our simulations, which

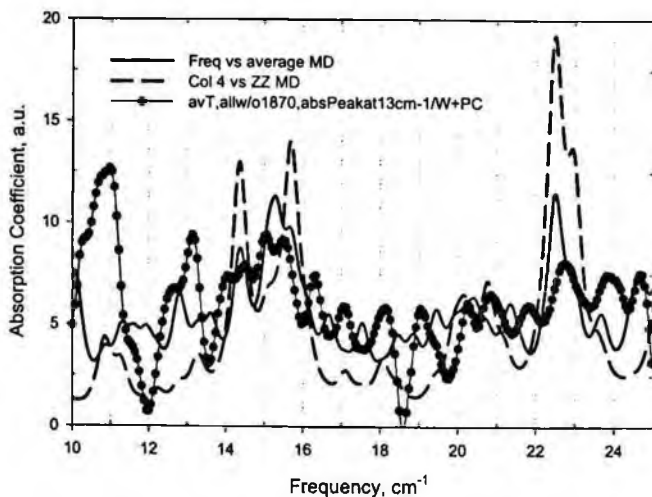


Fig. 9. Experimental absorption coefficient (blue dotted line) v averaged (green) MD simulation with explicit water and results for z-polarized light (dark red). Z-axis direction is shown in Figs. 1-2.

appears to be reasonable for our spectra. The Z-axis direction is shown in Figs. 1-2. Our averaged simulation spectrum describes quite well the experimental structure between 13 and 16 cm^{-1} . The peak at $\sim 22.5 \text{ cm}^{-1}$ is also well predicted, however, the experimental feature at 11 cm^{-1} is only seen in the predicted spectrum for the z-polarized light. Therefore, the orientation effects have probably played some role in our experiments. The significant absorption activity, in particular at $\sim 12.8 \text{ cm}^{-1}$, $\sim 14.4 \text{ cm}^{-1}$, $\sim 15.3 \text{ cm}^{-1}$, 22.5 cm^{-1} , 23.7 cm^{-1} and 24.7 cm^{-1} , predicted by our explicit water MD simulations agrees reasonably well with our experimental data.

We conclude that the developed THz modeling and simulation approach can be successfully used to predict and study THz absorption/transmission spectra of real proteins in water, in particular, proteins of biological warfare agents. The effects of orientation on THz spectra appear to be noticeable in water in the studied thin film protein – water samples. The results of this study demonstrate that terahertz spectroscopy is a promising tool in generating THz spectral data for protein components of bio agents. This is especially important since protein components might be significant for the whole cell/spore THz fingerprints.

6. Acknowledgments

This work was supported by the US DOD under Contract No W911NF-05-1-0033 and DAAD19-00-1-0402, and in part by U.S. Army NGIC Contract # DASC01-01-C0009.

References

1. Alexei Bykhovski, Xiaowei Li, Tatiana Globus, Tatyana Khromova, Boris Gelmont, Dwight Woolard, Alan C. Samuels, and James O. Jensen, THz absorption signature detection of genetic material of *E. coli* and *B. subtilis*, Chemical and biological standoff detection III, *Proc. SPIE* 5995, 198-207 (2005).
2. Alexei Bykhovski, Tatiana Globus, Tatyana Khromova, Boris Gelmont, and Dwight Woolard. An Analysis of the THz Frequency Signatures in the Cellular Components of Biological Agents. SPIE Defence & Security Symposium, terahertz for military and security applications IV, *Proc. SPIE* 6212, 132-141 (2006).
3. A. L. Lehninger, *Biochemistry*, 2nd edition, (1981).
4. Katti, S. K., LeMaster, D. M., Eklund, H., Crystal structure of thioredoxin from *Escherichia coli* at 1.68 Å resolution, *J Mol. Biol.* 212, 167 (1990).
5. D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, J. Wang, W.S. Ross, C.L. Simmerling, T.A. Darden, K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crowley, V. Tsui, H. Gohlke, R.J. Radmer, Y. Duan, J. Pitera, I. Massova, G.L. Seibel, U.C. Singh, P.K. Weiner and P.A. Kollman (2004), AMBER 8, University of California, San Francisco. <http://amber.scripps.edu/>.
6. W.L. Jorgensen, J. Chandrasekhar, J. Madura & M.L. Klein. Comparison of Simple Potential Functions for Simulating Liquid Water, *J. Chem. Phys.*, 79, 926-935 (1983).
7. Tatiana Globus, Tatyana Khromova, Boris Gelmont, Dwight Woolard, and Lukas K. Tamm, Terahertz characterization of diluted DNA solutions, *Proceedings of SPIE Photonics West 2006 (Biomedical Vibrational Spectroscopy III: Advances in Research and Industry)*, 6093, (2006).

STUDY OF LENNARD-JONES CLUSTERS: EFFECTS OF ANHARMONICITIES FAR FROM SADDLE POINTS

L. Y. CHEN

*Department of Physics, University of Texas at San Antonio, San Antonio, Texas 78249-0697
liao.chen@utsa.edu*

N. J. M. HORING

Department of Physics and Engineering Physics, Stevens Institute of Technology, Hoboken, New Jersey 07030

We study the transition pathways of a Lennard-Jones cluster of seven particles in three dimensions. Low lying saddle points of the LJ cluster, which can be reached directly from a minimum without passing through another minimum, are identified without any presumption of their characteristics, nor of the product states they lead to. The probabilities are computed for paths going from a given minimum to the surrounding saddle points. These probabilities are directly related to prefactors in the rate formula. This determination of the rate prefactors includes all anharmonicities, near or far from saddle points, which are pertinent in the very sophisticated energy landscape of LJ clusters and in many other complex systems.

Keywords: Lennard-Jones cluster; anharmonicities; transition state.

1. Introduction

In the past two and a half decades, a vast amount of theoretical work has been devoted to the study of Lennard-Jones (LJ) clusters.¹⁻⁵ Some investigations have been chiefly concerned with the difficult problem of optimizing geometries, locating important minima as well as saddle points. Other investigations have been aimed at determining the thermodynamic properties and understanding several related phenomena, namely, melting, isomerizations, solid-solid phase changes, and thermodynamic stability. The study of LJ clusters is important because the field of weakly bound complexes remains an important source of frontier explorations, both experimental and theoretical. The ultimate goal of all such investigations is a greater understanding of the details of the forces at play; these, in turn, are fundamental for a whole array of scientific and technological applications in material science, engineering, and molecular biology. From a theoretical point of view, weakly bound complexes provide a convenient, realistic model to study solvation effects and to study in detail all the phenomenology observed experimentally, as matter makes the transition from the gas phase to the condensed phase.

In this paper, we employ a recently derived formula to study a LJ cluster of seven particles in three dimensions (LJ7).⁵ Starting from a minimum on the potential energy surface (PES), we sample all possible transition paths out of this given minimum through the saddle points in the phase-space sphere surrounding the minimum. For each of those transition paths that goes from the given minimum through a saddle point to another minimum, we compute its statistical weight factor including the Boltzmann exponential factor and the pre-exponential factor. In this, we determine the transition rate prefactors along with locating saddle points. It should be pointed out that, for this well studied system of LJ7, our research finds new low-lying saddle points and adds the determination of rate prefactors including anharmonicities both far from and close to the saddle points.

The rest of the paper is organized as follows: In Section 2, we briefly present the theoretical formulation and the algorithm for sampling transition paths and evaluating the corresponding prefactors. In Section 3, we discuss the numerical results for LJ7. In Section 4, we provide a summary.

2. Formulation and algorithm

We consider the rate of transitions out of a reactant state (a minimum on the PES) at temperatures for which thermal energy $k_B T = 1/\beta$ is much less than the activation energy barriers (energy differences between the saddle points and the minimum). The Arrhenius formula is applicable and it can be expressed as

$$r = A \sum_i p_i e^{-\beta(E_{TS_i} - E_R)}, \quad (1)$$

where A is an “overall” prefactor and p_i is the probability distribution of the pathway that goes through i -th saddle point (apart from the Boltzmann exponential factor):

$$\sum_i p_i = 1. \quad (2)$$

E_R is the energy at the minimum. E_{TS_i} is the energy at the i -th saddle point. When multiple saddle points with similar energy barriers surround a minimum of interest, it is through the prefactors that some reaction pathways are favored over others.

Within the harmonic approximation of the transition state theory (TST), the prefactors are determined as the ratios between the product of all the eigenmode frequencies at the minimum and that of all the real eigenmode frequencies at the saddle points. This approximation is adequate for systems with slowly varying PES's, to the extent that no anharmonicities are involved. Going beyond the harmonic approximation of TST, in Ref.[5], the rate prefactor probabilities are determined as

$$p_i = \frac{1}{C} \sum_{\omega_{0,i}} \exp[-I_{eff}[r_{0,i}(t)]] \cdot \quad (3)$$

Here, $r(t)$ represents the set of position vectors for all particles of the system at time t . For LJ7, it is a collection of 21 coordinates at time t . The summation involved in Eq.(3) is over all paths, $r_{0,i}(t)$, that go from all initial states r_0 in the local equilibrium ensemble near the reactant state to the i -th saddle point r_{TSi} . The normalization factor

$$C = \sum_i \sum_{r_{0,i}} \exp[-I_{eff.}[r_{0,i}(t)]] \quad (4)$$

involves summation over all paths to all saddle points.

Eq.(3) is valid wherever the overdamped Langevin equation is applicable. Its utility rests on the facts that it can be used with all algorithms for finding saddle points starting from a local minimum on the PES, and that the required computational effort for determining the prefactors is minimal. One has only to evaluate the following integral,

$$I_{eff.}[r_{0,i}(t)] = \frac{\beta}{4\gamma} \sum_{l=1}^{N+1} \left[-\gamma \frac{r_{0,i}(t_l) - r_{0,i}(t_{l-1})}{dt} + \nabla V(r_{0,i}(t_{l-1})) \right]^2 dt, \quad (5)$$

along each sampled path, $r_{0,i}(t)$, which goes from r_0 at $t = t_0$ near the local minimum to the saddle point r_{TSi} at $t = t_f$. The prefactors for transitions from a given minimum on the PES to other minima on the PES through the saddle points surrounding that minimum are determined through Eq.(3) when paths are sampled from each of the initial states, r_0 , of a local equilibrium ensemble near that minimum.

Many studies have been focused on developing or refining methods for finding saddle points from a minimum without knowledge of the final states.⁶ Here, pointing out that Eq.(3) works with any one-point boundary algorithm for reaching saddle points from a minimum, we pursue the "gentlest ascent" approach.⁷ Starting from an initial state $r(t_0) = r_0$ near a minimum on the PES, the system state is advanced in time according to the following algorithm,

$$r(t) = r(t_{l-1}) + \frac{1}{\gamma} [\nabla V(r(t_{l-1})) + F(t_{l-1})] dt. \quad (6)$$

Its continuous form is the following differential equation,

$$-\gamma \frac{dr(t)}{dt} + \nabla V(r(t)) + F(r(t)) = 0. \quad (7)$$

Here, we have the usual force field $-\nabla V$ and an "extra" force introduced as

$$F(t) = -\alpha \left[1 - \frac{\nabla V(r(t)) \cdot \nabla V(r(t))}{|\nabla V(r(t))|^2} \right] \nabla V \left(r(t) + \frac{1}{\gamma} \nabla V(r(t)) dt \right). \quad (8)$$

The adjustable parameter α measures the strength of the “extra” force field. A path is generated as follows: Starting from r_0 , Eq.(6) is run until a saddle point is reached or a cut-off value of potential energy is exceeded. Each iteration in Eq.(6) involves two steps: First is to climb up along the gradient of the potential, ∇V , from $r_l = r(t_{l-1})$ to $r_l^* = r(t_{l-1}) + \nabla V(r(t_{l-1}))dt/\gamma$. Second is to evaluate the force $-\nabla V(r_l^*)$ at r_l^* and project out its component that is parallel to the force $-\nabla V(r(t_{l-1}))$ at $r(t_{l-1})$. Adjusting the strength of this projected force with α , the system state is advanced from r_l^* to $r_l = r_l^* + Fdt/\gamma$. The path so sampled, $r(t_l)$ ($l = 1, 2, \dots, N$), can be readily used in the integration of Eq.(5) in the evaluation of prefactors using Eq.(3).

Obviously, this algorithm may or may not lead to a saddle point in a given trial. If we define the sampling efficiency as the ratio between the number of paths that actually go from the reactant state to one of the saddle points and the total number of paths sampled, the efficiency of this algorithm can be over 50% when α is adjusted to its optimal value in our studies of the LJ cluster as a 21-dimensional system. At the end of the path, the Langevin equation,

$$\gamma \frac{dr(t)}{dt} + \nabla V(r(t)) = \xi(t), \quad (9)$$

can be run for the system to relax down to a nearby local minimum. Naturally, some of the paths so generated will settle down to new stable/metastable (product) states and others return to the vicinity of the initial state. Note that $\xi(t) = (\xi_1(t), \xi_2(t), \dots, \xi_{21}(t))$ in Eq.(9) is the Gaussian white noise random force in 21 dimensions, with mean value being zero and correlations being,

$$\langle \xi_i(t) \xi_j(t') \rangle = \frac{2k_B T}{\gamma} \delta_{ij} \delta(t - t') \quad \text{for } i, j = 1, 2, \dots, 21. \quad (10)$$

3. Numerical results

Our system is a cluster of seven particles in three dimensions that interact with one another by the LJ pair potential,

$$V_{LJ}(r) = w_0 \left(\frac{1}{r^{12}} - \frac{2}{r^6} \right) \quad (11)$$

where r is the distance between two particles and we take $w_0 = 1$. This pair potential has its minimum $V_{min} = -1$ at $r = 1$ and its well depth is 1. Rescaling the distance r to $r/2^{1/6}$, Eq.(11) immediately takes the standard form in the literature.² Throughout this paper, we assume the damping constant $\gamma = 1$ and the inverse temperature $\beta = 10$.

LJ7 as defined here has one stable state (global minimum) M1 and three meta-stable states (local minima) M2, M3, and M4. They are tabulated in Table I. This system has been well studied in the literature.² The present study adds the determination of

anharmonic prefactor probabilities for transitions through low-lying saddle points. These prefactor probabilities are computed beyond the local anharmonicity correction to the harmonic TST approximation.

Table 1: Four minima of LJ7 cluster. The error bar for the energy is ± 0.001 .

	Minimum 1	Minimum 2	Minimum 3	Minimum 4
Energy	-16.505	-15.935	-15.593	-15.533
Point group	D5h	C3v	C3v	C2

Fig. 1 exhibits histograms of two typical paths sampled with the algorithm given in Eq.(6). One path goes from M1 through TS1 to M2. The other goes from M1 through TS4 to M2. States M1 and M2 are shown in Fig.2.

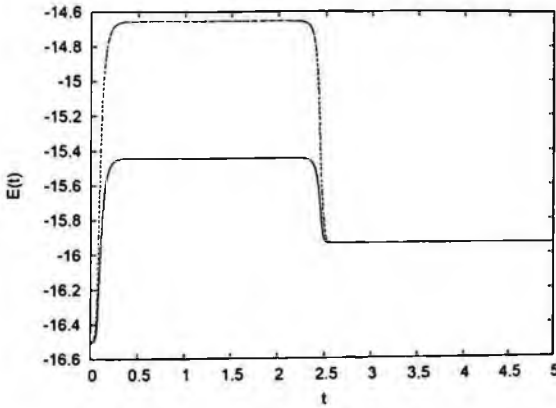


Fig.1: Transition paths from M1 through TS1 (solid line) and through TS4 (dashed line) to M2.

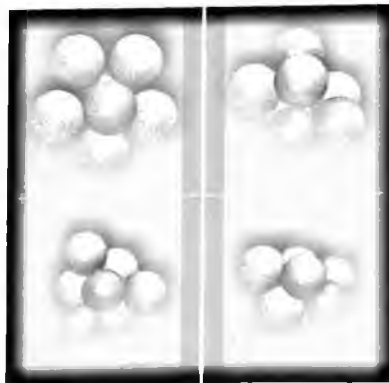


Fig.2: Stable state M1 (top left) and meta-stable states M2 (top right), M3 (bottom left), and M4 (bottom right).

Fig. 3 illustrates low lying saddle points TS1 and TS4 that can be reached directly from M1 without going through other metastable states. Both saddle points involve concerted displacements of multiple particles out of the equilibrium positions they had in their stable state M1. These concerted displacements involve stretching fewer “bonds”, thus corresponding to a lower energy barrier. This is a general feature of LJ clusters in three dimensions as well as in two dimensions.⁵

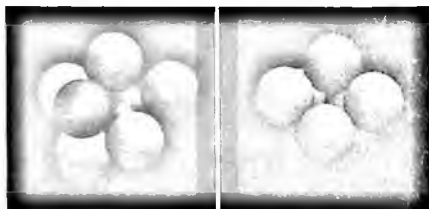


Fig.3: Low energy saddle points TS1 (left) and TS4 (right) that can be reached directly from stable state M1.

The saddle point energies and the prefactor probabilities are tabulated in Tables II through V. They are numbered in order from the lowest energy barrier to the highest. As shown in Table II, the lowest energy saddle point, TS1, is dominantly favored. No other saddle points are either energetically favored (having low energy barriers) or entropically favored (having large prefactor probabilities). Transitions out of M1 basically follow one route: M1-TS1-M2 as TS1 only connects M1 and M2. Table III shows the energy barriers and prefactor probabilities for transitions out of M2. TS1 that leads to M1 is favored both energetically and entropically. The next most probable route is through TS2, which leads to M3. Table IV shows transitions out of M3. TS1, which leads to M1 and M2, has the lowest energy barrier but a very small prefactor probability. TS2, TS3, and TS4 have similar energy barriers and similar prefactor probabilities. Table V shows transitions out of M4. TS3, TS4, and TS5 have slightly higher barriers than TS2 but much greater prefactor probabilities and therefore they are more important than TS2 if the temperature is not too low, $k_B T > 0.2$. Comparing TS3 and TS4, their barriers are nearly identical. Their relative importance is completely determined by the prefactor probabilities. The same thing can be said about TS6 and TS7 as well as TS14 and TS15. All these clearly lead to the conclusions that the prefactor probabilities are important and that anharmonicities far from the saddle points need to be taken into account in the rate calculations.

Table 2: Saddle point energies and prefactors for transitions from Minimum 1 (D5h) to the surrounding saddle points. The error bar for the energy barrier is ± 0.001 and that for the prefactor is ± 0.01 .

	TS1	TS2	TS3	TS4	TS5	TS6	TS7	TS8	TS9
E_{TSi}	-15.445	-15.029	-15.022	-14.656	-14.447	-14.266	-14.192	-14.129	-14.004
p_i	0.990	0.00035	0.016	0.0006	0.00035	0.00075	0.00005	0.0001	0.0006

Table 3: Saddle point energies and prefactors for transitions from Minimum 2 (C3v) to the surrounding saddle points. The error bars are the same as in Table II.

	TS1	TS2	TS3	TS4	TS5	TS6
E_{TSi}	-15.445	-15.320	-15.283	-15.098	-15.029	-15.022
p_i	0.915	0.051	0.0076	0.0069	0.00035	0.00045
	TS7	TS8	TS9	TS10	TS11	
E_{TSi}	-14.876	-14.749	-14.723	-14.555	-14.348	
p_i	0.013	0.0053	0.00005	0.00035	0.0002	

Table 4: Saddle point energies and prefactors for transitions from Minimum 3 (C3v) to the surrounding saddle points. The error bars are the same as in Table II.

	TS1	TS2	TS3	TS4	TS5	TS6
E_{TSi}	-15.029	-14.749	-14.568	-14.555	-14.447	-14.393
p_i	0.0042	0.308	0.258	0.153	0.012	0.0015
	TS7	TS8	TS9	TS10	TS11	TS12
E_{TSi}	-14.307	-14.280	-14.271	-14.051	-14.021	-14.004
p_i	0.0037	0.0011	0.0018	0.111	0.0185	0.127

Table 5: Saddle point energies and prefactors for transitions from Minimum 4 (C2) to the surrounding saddle points. The error bars are the same as in Table II.

	TS1	TS2	TS3	TS4	TS5	TS6	TS7	TS8	TS9
E_{TSi}	-15.283	-15.022	-14.816	-14.811	-14.723	-14.555	-14.549	-14.447	-14.381
p_i	0.145	0.0023	0.015	0.314	0.191	0.070	0.186	0.0023	0.00188
	TS10	TS11	TS12	TS13	TS14	TS15	TS16	TS17	TS18
E_{TSi}	-14.380	-14.354	-14.348	-14.307	-14.285	-14.266	-14.198	-14.021	-14.004
p_i	0.0097	0.0091	0.00064	0.00097	0.0018	0.0395	0.00091	0.010	0.00011

4. Summary

In summary, we have applied the theory of Ref.5 to the study of Lennard-Jones (LJ) clusters. For an LJ cluster of seven particles in three dimensions, a system of 21 degrees of freedom, we have thoroughly examined transitions out of each of the four minima. All relevant (low lying) saddle points in the 21-dimensional PES of the LJ cluster, which can be reached directly from a minimum without passing through another minimum, are identified without any presumption of their characteristics, nor of the product states they lead to. The probabilities, p_i , are computed for paths going from an equilibrium ensemble of states near a given minimum to the surrounding saddle points. This determination of the rate prefactor probabilities includes all anharmonicities, close to or

far from saddle points, which are pertinent in the very sophisticated energy landscape of LJ clusters and in many other complex systems.

It should be pointed out that the LJ cluster of seven particles has been well studied in the literature. This analysis provides the determination of prefactor probabilities for transitions through low-lying saddle points. These prefactor probabilities are computed beyond the local anharmonicity correction to the harmonic TST approximation. Finally, this study can be extended in a straightforward way to study much larger LJ clusters and other complex systems such as biomolecules.

Acknowledgements

This work has been supported in part by an NIH/MBRS/SCORE grant (GM008194), by an NCSA/TeraGrid grant (TG-MCB060002T), and by the U.S. Department of Defense (DAAD #19-01-1-0592) through the DURINT program of the Army Research Office.

References

1. C. J. Tsai and K. D. Jordan, *J. Phys. Chem.* **97**, 11227 (1993).
2. M. A. Miller and D. J. Wales, *J. Chem. Phys.* **107**, 8568 (1997).
3. E. Yurtsever, A. Palazoglu, and Y. Arkun, *J. Phys. Chem. A* **107**, 6025 (2003).
4. M. P. Dawn and E. Curotto, *J. Chem. Phys.* **123**, 144301 (2005).
5. L. Y. Chen and N. J. M. Horing, *J. Chem. Phys.* **124**, 164102 (2006).
6. For a comparison of various methods of finding saddle points without knowledge of the final states, see R. A. Olsen, G. J. Kroes, G. Henkelman, A. Arnaldsson, and H. Jonsson, *J. Chem. Phys.* **121**, 9776 (2004).
7. G. M. Crippen and H. A. Scheraga, *Arch. Biochem. Biophys.* **144**, 462 (1971).

VAN DER WAALS INTERACTION BETWEEN ADSORBATE LAYERS/MOLECULES AND A SUBSTRATE

NORMAN J. MORGENSTERN HORING

*Department of Physics and Engineering Physics, Stevens Institute of Technology
Hoboken, New Jersey 07030, USA*

We present a theoretical formulation of van der Waals (VdW)¹⁻¹⁹ molecule-substrate attraction in which the second order vdW energy is explicitly exhibited as a correlation/self-energy of the molecular/atomic electrons generated by a dynamic nonlocal image potential due to polarization of the electrons of the bounded metal/semiconductor substrate in the electrostatic limit.²⁰⁻²³ This formulation can also be applied to the vdW interaction between an adsorbate layer and the substrate, as well as the interaction between layers. We have already applied it in the case of atom-surface vdW attraction in the presence of a normal magnetic field which induces both classical and Landau quantization magnetic effects,²⁴ incorporating the role of dynamic and nonlocal plasma effects. The extension to multiple adsorbate layers and their mutual interactions as well as their attraction to the substrate is straightforward.²⁵ The dependence of the atom/molecule-surface vdW energy on magnetic field strength provides an adjustable parametrization of the underlying zero-point photon energy (represented in terms of the nonretarded longitudinal plasmon-photons of the Coulomb interaction), opening the possibility of analyzing the concomitant fundamental quantum phenomenology in detail with material parameters that can be examined experimentally.

Keywords: Van der Waals Interaction; Image Potential; Screening Function.

1. Introduction

The theory of van der Waals¹⁻¹⁹ (vdW) surface interactions is presented here in terms of correlation-self energies of the constituent parts involved in the interaction due to their mutual polarization in the electrostatic limit.²⁰⁻²³ In this description the van der Waals interactions are exhibited using the dynamic, nonlocal and inhomogeneous screening functions of the constituent parts. In regard to the van der Waals interaction of a single molecule and a substrate, this problem is substantially the same as that of the van der Waals interaction of an atom and a substrate, in which the atomic aspects of the problem are subsumed in a multipole expansion based on spatial localization of the atom/molecule. As we (and others) have treated this in detail in the past²⁰⁻²⁵ we will not discuss it further in this paper. Here, our attention will be focussed on the van der Waals interaction of an adsorbate layer with a substrate, with the dielectric properties of the adsorbate layer modeled as a two-dimensional plasma sheet, and those of the substrate modeled by a semi-infinite bulk plasma. This formulation can be easily adapted to an

adsorbate layer of any given dynamic, nonlocal polarizability, or a wire-like structure, dot-like structure, and can accommodate nonlocality, multiple-layer structure of the adsorbate and of the substrate, as well as phonons, magnetic field, etc.

2. Formulation of van der Waals Surface Interactions: Green's Functions and Screening Functions

In earlier work,²⁰⁻²⁵ we have shown that the van der Waals surface interaction energy of two distinct neutral systems, A and B, which have no wave function overlap and/or charge exchange, is given to second order in the Coulomb interaction between its constituent parts as ($v(\vec{r}_1 - \vec{r}_2)$ is the Coulomb potential; $\hbar \rightarrow 1$)

$$E_{int} = -i \int_{-\infty}^0 dt \int d\vec{r}_1 \int d\vec{r}_2 \int d\vec{r}_3 \int d\vec{r}_4 v(\vec{r}_1 - \vec{r}_2) v(\vec{r}_3 - \vec{r}_4) \quad (1)$$

$$\times G_2^A(\vec{r}_1 0, \vec{r}_3 t; \vec{r}_1 0^+, \vec{r}_3 t^+) G_2^B(\vec{r}_2 0, \vec{r}_4 t; \vec{r}_2 0^+, \vec{r}_4 t^+).$$

Here, $G_2^{A,B}$ are the two-electron equilibrium Green's functions of the disjoint subsystems A, B, respectively, and we have used overall charge neutrality to exactly eliminate a number of terms involving one-electron equilibrium Green's functions, $G_1^{A,B}$. The interpretation in terms of self-energy can be advanced by employing a Kubo-type linear response relation for the screening function $K(\vec{r}_1 t_1, \vec{r}_2 t_2)$ that is the space-time matrix inverse of the direct dielectric function $\varepsilon(\vec{r}_1 t_1, \vec{r}_2 t_2)$, written for subsystem A as

$$\int d\vec{r}_3 \int dt_3 \varepsilon_A(\vec{r}_1 t_1, \vec{r}_3 t_3) K_A(\vec{r}_3 t_3, \vec{r}_2 t_2) = \delta^{(3)}(\vec{r}_1 - \vec{r}_2) \delta(t_1 - t_2). \quad (2)$$

The linear response relation²⁰⁻²⁵ for A states that K_A relates to the density-density correlation function of A as ($v(1-3) = v(\vec{r}_1 - \vec{r}_3) \delta(t_1 - t_3)$; $1 = \vec{r}_1, t_1$, etc.)

$$K_A(1,2) = \delta^{(4)}(1-2) + \int \delta^{(4)}(3) v(1-3) \quad (3)$$

$$\times [G_2^A(2,3;2^+,3^+) - G_1^A(2,2^+) G_1^A(3,3^+)].$$

Explicitly recognizing its role in a screening capacity, it generates the effective potential due to Coulomb interaction, V_{eff} , and its image part, V_{image} , in accordance with

$$V_{eff}^A(3,1) = \int d^{(4)}2 K(3,2)v(2-1).$$

and

$$V_{image}^A(3,1) \equiv V_{eff}^A(3,1) - v(3-1). \tag{4}$$

Employing this in Eq(1), we have

$$E_{int} = - \int_{-\infty}^0 dt \int d\vec{r}_1 \int d\vec{r}_3 G_2^B(\vec{r}_1 0, \vec{r}_3 t; \vec{r}_1 0^+, \vec{r}_3 t^+) V_{image}^A(\vec{r}_3 t; \vec{r}_1 0). \tag{5}$$

In all of these considerations $G_2^{A,B}$ carry the full correlations internal to their individual subsystems, A , B , respectively, with no relation to the other subsystem, B , A , respectively. As the operators internal to the statistical trace matrix elements of G_2^B in Eq.(5) involve the charge density operators of subsystem B, $\hat{\rho}_B(\vec{r}_1 0), \hat{\rho}_B(\vec{r}_3, t)$, it is immediately apparent that $G_2^B V_{image}^A$ as it occurs in E_{int} in Eq(5) represents a correlation self-energy of subsystem B electrons mediated by the potential V_{image} induced through the polarization of subsystem A (which occurs mutually with the polarization of subsystem B when A and B are in proximity). Of course, this image potential is dynamic, nonlocal and inhomogeneous.

As attractive as the foregoing interpretation is, it is preferable to use a more symmetric formulation for the problem of van der Waals interaction between a 2D adsorbate layer and a planar substrate. To this end, we reconsider Eq(1) using Eq(3), obtaining

$$E_{int} = -i \int_{-\infty}^0 dt \int d\vec{r}_2 \int d\vec{r}_3 \{ [K_A(\vec{r}_2, 0; \vec{r}_3, t) - \delta^{(3)}(\vec{r}_2 - \vec{r}_3)\delta(t)] \times [K_B(\vec{r}_3, t; \vec{r}_2, 0) - \delta^{(3)}(\vec{r}_3 - \vec{r}_2)\delta(t)] \}, \tag{6}$$

where irrelevant terms involving one-electron Green's functions have again been eliminated. The screening functions K_A and K_B have been determined in our earlier work, using the random phase approximation (RPA) for each subsystem separately.

Modeling the substrate, "A", as a local semi-infinite planar plasma-like medium ($z > 0$) interfacing vacuum at $z = 0$, the screening function, K_A , is given by^{26,27,28}

$[\Theta_+(z)$ is the Heaviside unit step function and we Fourier transform in the plane of translational symmetry ($x, y \rightarrow q_x, q_y = \bar{q}$) and in time ($t_1 - t_2 \rightarrow \omega$)],

$$K_A(z_1, z_2) = \frac{1}{\eta_+(z_1)} \delta(z_1 - z_2) - \delta(z_2) \Gamma \frac{\eta_-(z_1)}{\eta_+(z_1)} e^{-|z_1| |\bar{q}|} \tag{7}$$

where Γ is the image strength factor (and \bar{q}, ω are suppressed)

$$\Gamma = \frac{1 - \varepsilon}{1 + \varepsilon} \tag{8}$$

with

$$\varepsilon = \varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2} \tag{9}$$

(ω_p is the classical bulk plasma frequency) and

$$\eta_-(z) = \Theta(z) - \Theta(-z) ; \eta_+(z) = \varepsilon \Theta(z) + \Theta(-z). \tag{10}$$

In regard to the adsorbate layer, "B", outside the substrate, we treat it as a sheet of 2D mobile plasma at $z_0 < 0$. We have exhibited the associated RPA screening function as²⁹

$$K_B(z_1, z_2) = \delta(z_1 - z_2) + \delta(z_2 - z_0) e^{-|\bar{q}| |z_1 - z_0|} [\tilde{K}_{2D} - 1] \tag{11}$$

where

$$\tilde{K}_{2D} = (1 + \alpha_{2D})^{-1} = [\tilde{\varepsilon}_{2D}]^{-1} \tag{12}$$

and $\alpha_{2D} = \alpha_{2D}(\bar{q}, \omega)$ was determined in the random phase approximation (RPA) by Stern³⁰ in the absence of a magnetic field, and with a magnetic field by Horing, et al.^{31,32} While we could employ the fully detailed nonlocal structures of K_A and K_B , we illustrate here using just the local limit with no magnetic field or phonons. In this case (n_{2D} is the 2D electron sheet density),

$$\alpha_{2D} = -\omega_{2D}^2/\omega^2 \quad \text{and} \quad \omega_{2D}^2 = 2\pi e^2 \frac{n_{2D}}{m} |\bar{q}|. \tag{13}$$

3. Applications: Two Bulk Dielectrics with Gap and Adsorbate Layer with Substrate

To compare the physically transparent formulation of Eq.(6) with known results, we consider the case of two dielectric half-spaces separated by a gap from $-z_0$ to z_0 (Figure 1, planar geometry). Fourier transforming in the plane of translational symmetry, $(x, y \rightarrow p_x, p_y = \bar{p})$ and time $(t \rightarrow \omega - i0^+$ for convergence), we employ the notation $\overline{\overline{K}} \equiv K - \delta$, obtaining

$$\frac{E_{int}}{Area} = - \int dz_2 \int dz_3 \int \frac{d^2 \bar{q}}{(2\pi)^2} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \int_{-\infty}^{\infty} \frac{d\Omega}{2\pi} \frac{\overline{\overline{K}}_A(z_2, z_3; \bar{q}, \omega) \overline{\overline{K}}_B(z_3, z_2; \bar{q}, \Omega)}{(\omega - \Omega - i0^+)}. \tag{14}$$

Manipulation of the frequency integrals and deforming their contour to the imaginary axis,^{33,34} with $\omega = iu$, yields

$$\frac{E_{int}}{Area} = \frac{1}{\pi} \int_0^{\infty} du \int \frac{d^2 \bar{q}}{(2\pi)^2} \int dz_2 \int dz_3 \overline{\overline{K}}_A(z_2, z_3; \bar{q}, u) \overline{\overline{K}}_B(z_3, z_2; \bar{q}, u). \tag{15}$$

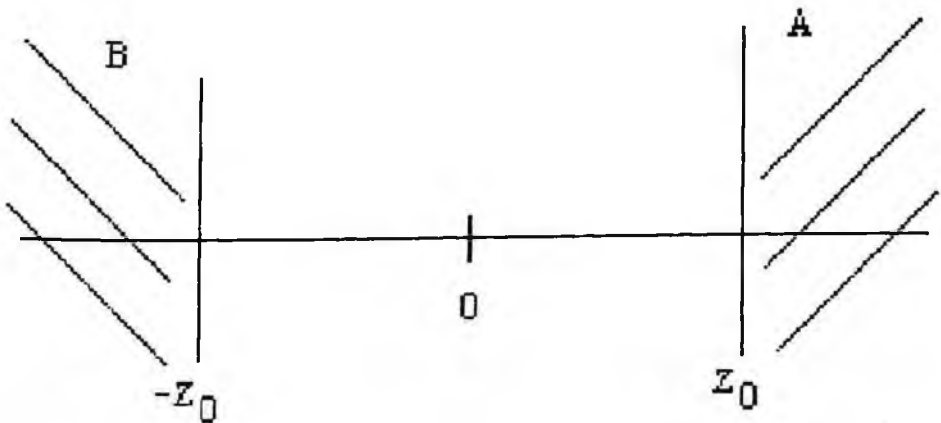


Figure 1. Schematic of two semi-infinite bulk dielectrics, A,B, with gap width of $2z_0$

Adjusting Eq.(7) in accordance with the two separated semi-infinite bulks of Fig. 1, the z_2 - and z_3 - integrals of Eq.(15) are given by

$$\begin{aligned} \int dz_2 \int dz_3 \dots &= \int dz_2 \int dz_3 \delta(z_3 - z_0) \delta(z_2 + z_0) \Gamma_A(iu) \Gamma_B(iu) \\ &\times e^{-|z_2 - z_0| |\bar{q}|} e^{-|z_3 + z_0| |\bar{q}|} \\ &= e^{-4|z_0| |\bar{q}|} \Gamma_A(iu) \Gamma_B(iu). \end{aligned} \tag{16}$$

Evaluating the \bar{q} -integral of Eq.(15) in circular coordinates, we obtain

$$\frac{E_{int}}{Area} = \frac{1}{8\pi^2 |2z_0|^2} \int_0^\infty du \Gamma_A(iu) \Gamma_B(iu). \tag{17}$$

This nonretarded result is similar to that of the retarded calculations by Lifshitz³⁵ and Renne^{36,37} to second order.

Considering the problem of a thin 2D adsorbate layer at $z_0 < 0$ and a semi-infinite substrate terminating at the plane $z = 0$ (Figure 2), we substitute Eqs.(7) and (11) in Eq.(15), obtaining

$$\begin{aligned} \frac{E_{int}}{Area} &= \frac{1}{\pi} \int_0^\infty du \int \frac{d^2 \bar{q}}{(2\pi)^2} \Gamma_A(iu) [\tilde{K}_{2D}(iu) - 1] \\ &\times \int dz_2 \int dz_3 \delta(z_3) \delta(z_2 - z_0) e^{-|\bar{q}| |z_2|} e^{-|\bar{q}| |z_3 - z_0|}, \end{aligned} \tag{18}$$

or

$$\frac{E_{int}}{Area} = \frac{1}{\pi} \int_0^\infty du \Gamma_A(iu) \int \frac{d^2 \bar{q}}{(2\pi)^2} [\tilde{K}_{2D}(\bar{q}, iu) - 1] e^{-2|\bar{q}| |z_0|}. \tag{19}$$

Employing Eqns.(12),(13) with $C = 2\pi e^2 n_{2D}/m$, we have

$$\frac{E_{int}}{Area} = \frac{C}{2\pi^2} \int_0^\infty du \Gamma_A(iu) \int \frac{d|\bar{q}| |\bar{q}|^2 e^{-2|\bar{q}| |z_0|}}{u^2 + C|\bar{q}|}. \tag{20}$$

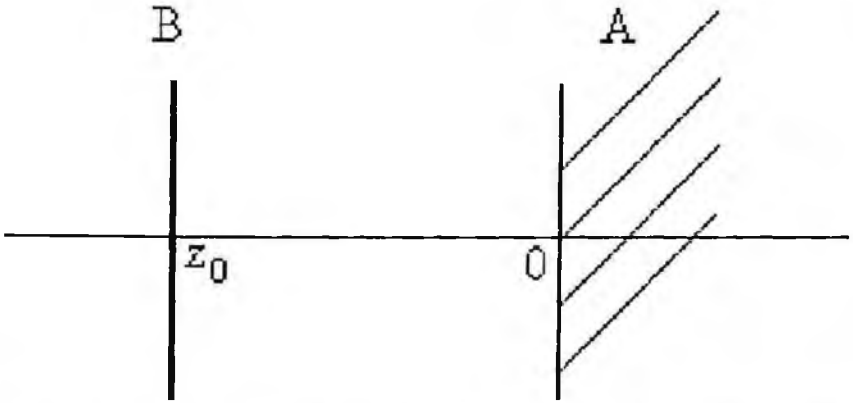


Figure 2: Schematic of a semi-infinite bulk substrate A, and a planar 2D layer B separated by a gap of width $|z_0|$.

Introducing the image strength factor Γ of Eq.(8) with Eq.(9) and performing the u -integral as

$$\int du \frac{1}{(u^2 + a^2)(u^2 + b^2)} = \frac{\pi}{2ab} \frac{1}{(a+b)}, \quad (21)$$

we find

$$\begin{aligned} \frac{E_{int}}{Area} &= \frac{\omega_p C^{1/2}}{32\pi |z_0|^{5/2}} \int dx \frac{x^{3/2} e^{-x}}{\omega_s + x^{1/2} \left[\frac{C^{1/2}}{(2|z_0|)^{1/2}} \right]} \\ &= \hbar \frac{|e| n_{2D}^{1/2} \omega_p}{8\sqrt{2} \pi^{1/2} m^{1/2} |z_0|^{5/2}} \int dy \frac{y^4 e^{-y^2}}{\omega_s + y (\pi e^2 n_{2D} / m |z_0|)^{1/2}} \end{aligned} \quad (22)$$

where $\omega_s = \omega_p / \sqrt{2}$ is the classical surface plasmon frequency ($y = x^{1/2}$ and \hbar is restored). If the 2D density, n_{2D} , of the adsorbate layer is small relative to the bulk density of the semi-infinite substrate, ρ_{3D} , in the sense

$$\frac{n_{2D}}{\rho_{3D} |z_0|} \ll 1, \quad (23)$$

or if $|z_0|$ is sufficiently large, then we obtain

$$\frac{E_{int}}{Area} = \hbar \frac{|e|n_{2D}^{1/2}}{8\pi^{1/2}m^{1/2}|z_0|^{5/2}} \int_0^\infty dy y^4 e^{-y^2} = \hbar \frac{3|e|n_{2D}^{1/2}}{128\pi m^{1/2}|z_0|^{5/2}}. \quad (24)$$

4. Conclusions

We have interpreted the second order van der Waals interaction energy between two distinct subsystems, A and B , in terms of a correlation self-energy (Eq.(5)) of the electrons of B mediated by the image potential induced through the polarization of subsystem A (which occurs mutually with the polarization of subsystem B when A and B are in proximity). This image potential is dynamic, nonlocal and inhomogeneous. Of course, this relation is reflexive, and subsystems A and B can be interchanged in this interpretation. We have also re-interpreted the second order vdW energy in terms of the dynamic, nonlocal and inhomogeneous inverse dielectric screening functions, K_A , K_B , of the two subsystems (Eq.(6)). This has been applied to two special cases,

- (1) two semi-infinite bulk dielectrics separated by a gap, to compare with earlier results, and
- (2) a semi-infinite bulk substrate and a planar 2D adsorbate layer, to illustrate the ease of application to new problems of this formulation of the vdW interaction.

Although we have limited this presentation to the local plasma limit we have already applied it to study the effects of spatial nonlocality and an ambient magnetic field.²⁰⁻²⁵ It can also be easily extended to include the roles of phonons, multiple layers, wire- and dot-like structures, lattices of such, as well as a few molecules.

Acknowledgments

This work was supported in part by the U.S. Department of Defense (DAAD # 19-01-1-0592) through the DURINT program of the Army Research Office.

References

1. U. Mohideen and Anushree Roy, Phys. Rev. Lett. **81**, 4549 (1998).
2. K.A. Milton, *The Casimir Effect* (World Scientific, Singapore, 2001).
3. *The Casimir Effect 50 Years Later*, edited by Michael Bordag (World Scientific, Singapore, 1999).
4. Peter W. Milonni, *The Quantum Vacuum* (Academic, New York, 1994).
5. W. Dittrich and H. Gies, *Probing the Quantum Vacuum*, Vol. 166 of Tracts in Modern Physics (Springer-Verlag, Berlin, 2000)
6. Paul R. Berman, *Cavity Quantum Electrodynamics* (Academic, New York, 1994).

7. *Long Range Casimir Forces* edited by F. S. Levin and D. A. Micha (Plenum Press, New York, 1993).
8. V. M. Mostrepanenko and N. N. Trunov, *The Casimir Effect and its Applications* (Clarendon Press, Oxford 1997); *Sov. Phys. Usp.* **31**, 965 (1988).
9. G. Plunien, B. Müller, and W. Greiner, *Phys. Rep.* **134**, 87 (1986).
10. J. Mahanty and B. W. Ninham, *Dispersion Forces* (Academic, New York, 1976).
11. D. Langbein, *Theory of van der Waals Attraction*, Tracts in Modern Physics Vol. **72** (Springer-Verlag, Berlin, 1974).
12. J. N. Israelachvili and D. Tabor, *Prog. Surf. Membrane Phys.* **7**, 1 (1973).
13. G. Feinberg, J. Sucher, and C.K. Au, *Phys. Rep.* **180**, 83 (1989).
14. I. E. Dzyaloshinskii, E. M. Lifshitz and L. P. Pitaevskii, *Adv. Phys.* **10**, 165 (1961).
15. E. A. Power and T. Thirunamachandran, *Phys. Rev. A* **50**, 3929 (1994).
16. M. Hawton, *Phys. Rev. A* **46**, 6846 (1992).
17. Yu. S. Barash and V. L. Ginzburg, *Sov. Phys. Usp.* **18**, 306 (1975).
18. P. W. Milonni and M. L. Shih, *Phys. Rev. A* **45**, 4241 (1992).
19. J. Bardeen, *Phys. Rev.* **58**, 727 (1940).
20. N. J. M. Horing and S. Silverman, *Il Nuovo Cimento* **38**, 396 (1977).
21. N. J. M. Horing and S. Silverman, *Proc 4th Intl. Conf. Solid Surfaces*, Supp. Rev. La Vide, les Courches Mince **201**, 83 (1980).
22. N. J. M. Horing and L. Y. Chen, *Phys. Rev A* **66**, 042905 (2002).
23. N. J. M. Horing, in "*Quantum Field Theory Under the Influence of External Conditions*", Ed: K. A. Milton, Rinton Press, 304-316 (2004).
24. N. J. M. Horing, *Phys. Rev. A* **69**, 032901 (2004).
25. N. J. M. Horing, *J. Phys. A*, accepted and in press (2006).
26. N. J. M. Horing, E. Kamen and H. L. Cui, *Phys. Rev.* **B32**, 2184 (1985).
27. N. J. M. Horing, T. Jena and H. L. Cui, *Phys. Rev.* **B54**, 2785 (1996).
28. N. J. M. Horing, et al., *Proc. ICPIG22*, Ed: E. Kunhardt, Vol. **3**, 27 (1995).
29. N. J. M. Horing, H. C. Tso and G. Gumbs, *Phys. Rev.* **B36**, 1588-Appendix (1987).
30. T. Ando, A. Fowler and F. Stern, *Rev. Mod. Phys.* **54**, 437 (1982).
31. N. J. M. Horing, et al., *Phys. Lett.* **48A**, 7 (1974).
32. N. J. M. Horing and M. M. Yildiz, *Ann. of Phys. (NY)*, **97**, 216 (1976).
33. A.D. McLachlan, *Proc. Roy. Soc. A, London*, **271**, 387 (1963).
34. C. Mauroyannis, *Mol. Phys.* **6**, 593 (1963).
35. E.M.Lifshitz, *Soviet Phys. JETP* **2**, 73 (1956).
36. M.J.Renne, *Physica* **56**, 125 (1971).
37. M.J.Renne and B.R.A. Nijboer, *Chem. Phys. Lett.* **1**, 317 (1967).

THE SPIN-HALL EFFECT IN p -TYPE BULK SEMICONDUCTORS

S. Y. LIU

*Department of Physics and Engineering Physics, Stevens Institute of Technology,
Hoboken, New Jersey 07030, USA
and Department of Physics, Shanghai Jiaotong University,
1954 Huashan Road, Shanghai 200030, China*

N. J. M. HORING

*Department of Physics and Engineering Physics, Stevens Institute of Technology,
Hoboken, New Jersey 07030, USA*

X. L. LEI

*Department of Physics, Shanghai Jiaotong University,
1954 Huashan Road, Shanghai 200030, China*

The spin Hall effect provides a new possible way to effectively inject spins into paramagnetic semiconductors. Here, we investigate the spin-Hall effect in a p -type Luttinger semiconductor employing a two-band kinetic equation analysis. The long-range disorder effect on spin-Hall current (SHC) is considered within the self-consistent Born approximation. We find that in addition to the intrinsic SHC proposed previously, there is a nonvanishing SHC that originates from long-range electron-impurity scattering, but which is independent of impurity density in the diffusive regime. This SHC has an opposite sign from the intrinsic one, leading to a significant reduction of the total SHC. We also carry out a numerical analysis of the hole density dependencies of SHC and spin mobility, finding that with increasing hole density, the SHC first increases and then falls, while the spin mobility monotonically decreases.

Keywords: Spin-Hall Effect; p -type Bulk Semiconductors; Magnetoresistive Sensors.

1. Introduction

Magetoresistivity-based sensors, which are based on a resistivity change induced by a magnetic field, have found wide application in industry. However, such sensors are usually fabricated of metals and hence it is difficult to integrate them with conventional semiconductor devices. To develop magetoresistive sensors in semiconductors, one of the key problems is how to effectively inject spin-polarized electrons into paramagnetic semiconductors.¹ The spin Hall effect (SHE), which involves a spin-current transverse to an applied electric field,^{2,3,4,5,6,7} potentially provides a new way to resolve this issue.

It has been shown that the spin-Hall effect may arise from various spin-orbit couplings, such as a spin-orbit (SO) interaction induced by the electron-impurity scattering potential,^{2,3} a Rashba SO coupling in two-dimensional systems,⁵ *etc.* Murakami *et al.* also predicted a nonvanishing spin-Hall current (AHC) in a perfect Luttinger bulk *p*-type semiconductors (no impurities or defects).⁴ Experimental observations of the spin-Hall effect have been reported recently in a *n*-type bulk semiconductor⁶ and in a two-dimensional heavy-hole system.⁷

In this paper, we employ a helicity-basis kinetic equation approach to investigate the spin-Hall effect in a *p*-type Luttinger semiconductor. Considering long-range electron-impurity scattering, we find that in addition to the intrinsic SHE proposed by Murakami *et al.*,⁴ long-range electron-impurity scattering gives rise to a nonvanishing SHC. This disorder-related SHC is independent of impurity density in the diffusive regime and has an opposite sign from the intrinsic one, leading to a significant reduction of the total SHC. A numerical analysis of the hole density dependencies of SHC and spin mobility indicates that with increasing hole density, the SHC first increases and then falls, while the spin mobility monotonically decreases.

2. Formalism

In semiconductors with diamond structure (e.g. Si, Ge) or zinc blende structure (e.g. GaAs), the electronic structure near the top of valence bands having $S = 3/2$ (S is the total angular momentum of the atomic orbital) can be described by the Luttinger Hamiltonian,⁸

$$\tilde{h}_0(\mathbf{p}) = \frac{1}{2m} \left[\left(\gamma_1 + \frac{5}{2} \gamma_2 \right) p^2 - 2\gamma_2 (\mathbf{p} \cdot \mathbf{S})^2 \right], \quad (1)$$

where, $\mathbf{p} \equiv (p_x, p_y, p_z) \equiv (p \sin \theta_p \cos \phi_p, p \sin \theta_p \sin \phi_p, p \cos \theta_p)$ is the

three-dimensional (3D) hole momentum, m is the free electron mass, $\mathbf{S} \equiv (S_x, S_y, S_z)$

are the spin-3/2 matrices, and γ_1 and γ_2 are material constants. By a local unitary

spinor transformation, $U_p = \exp(-iS_z \phi_p) \exp(-iS_y \theta_p)$, Hamiltonian (1) can be

diagonalized as $\hat{h}_0(\mathbf{p}) = U_p^\dagger \tilde{h}_0(\mathbf{p}) U_p = \text{diag}[\epsilon_H(p), \epsilon_L(p), \epsilon_L(p), \epsilon_H(p)]$, with

$\epsilon_H(p) = (\gamma_1 - 2\gamma_2)p^2 / (2m)$ and $\epsilon_L(p) = (\gamma_1 + 2\gamma_2)p^2 / (2m)$ as the dispersion

relations of the heavy- and light-hole bands. Physically, this transformation corresponds to a change from a spin basis to a helicity basis. In such a *p*-type semiconductor, the interaction between holes and impurities can be described by an isotropic potential, $V(|\mathbf{p} - \mathbf{k}|)$, which, in the helicity basis, reduces to $\hat{T}(\mathbf{p}, \mathbf{k}) = U_p^\dagger V(|\mathbf{p} - \mathbf{k}|) U_k$.

In this paper, we study the spin-Hall current in a bulk hole system driven by a dc electric field \mathbf{E} along the *z* axis. In Coulomb gauge, this electric field can be described by a scalar potential, $V \equiv e\mathbf{E} \cdot \mathbf{r}$, with \mathbf{r} as the hole coordinate. Using the expression for the conserved spin-Hall operator in spin basis proposed by Murakami *et al.*,⁹ and applying the unitary transformation U_p , we find that the observed net spin-Hall current,

J_y^x , which is polarized along the *x* axis and flows along the *y* axis, is determined in helicity basis by

$$\begin{aligned}
 J_y^x = & \frac{\sqrt{3}\gamma_2}{m} \sum_{\mathbf{p}} p \left\{ 4 \cos^2 \phi_p \sin \theta_p \text{Im}[\hat{\rho}_{12}(\mathbf{p}) + \hat{\rho}_{34}(\mathbf{p})] \right. \\
 & - \sin(2\phi_p) \sin(2\theta_p) \text{Re}[\hat{\rho}_{12}(\mathbf{p}) + \hat{\rho}_{34}(\mathbf{p})] \\
 & + 2 \cos(2\phi_p) \cos \theta_p \text{Im}[\hat{\rho}_{13}(\mathbf{p}) - \hat{\rho}_{24}(\mathbf{p})] \\
 & \left. - \sin(2\phi_p) [1 + \cos^2 \theta_p] \text{Re}[\hat{\rho}_{13}(\mathbf{p}) - \hat{\rho}_{24}(\mathbf{p})] \right\}
 \end{aligned} \quad (2)$$

Here, $\hat{\rho}_{\mu\nu}(\mathbf{p})$ ($\mu, \nu = 1, 2, 3, 4$) are the matrix elements of the helicity-basis distribution function $\hat{\rho}(\mathbf{p})$. From Eq. (2) it is obvious that contributions to the spin-Hall current arise only from those matrix elements of the distribution function which describe interband polarization, such as $\hat{\rho}_{12}(\mathbf{p})$, $\hat{\rho}_{13}(\mathbf{p})$, $\hat{\rho}_{23}(\mathbf{p})$ and $\hat{\rho}_{24}(\mathbf{p})$. This implies that the spin-Hall effect in helicity basis is just the result of polarization between the heavy and light hole bands.

To obtain the spin-Hall current, it is necessary to analyze the hole distribution function. Within the linear response regime, we find that the linear electric field part of the helicity-basis distribution function, $\hat{\rho}_1(\mathbf{p})$, can be written as a sum of two terms,

$\hat{\rho}_1(\mathbf{p}) = \hat{\rho}_1'(\mathbf{p}) + \hat{\rho}_1''(\mathbf{p})$ with $\hat{\rho}_1'(\mathbf{p})$ and $\hat{\rho}_1''(\mathbf{p})$ determined by ($\hat{I}^{(1)}$ is the linear electric field part of the scattering term)

$$e\mathbf{E} \cdot \nabla_{\mathbf{p}} \hat{\rho}_0(\mathbf{p}) + i[\hat{h}_0(\mathbf{p}), \hat{\rho}_1'(\mathbf{p})] = -\hat{I}^{(1)}, \quad (3)$$

$$-e\mathbf{E} \cdot [\hat{\rho}_0(\mathbf{p}), U_p^+ \nabla_p U_p] + i[\hat{h}_0(p), \hat{\rho}_1''(\mathbf{p})] = 0. \tag{4}$$

Ignoring the higher order contributions to $\hat{I}^{(1)}$ from the off-diagonal elements of distribution function in the impurity-density expansion, $\hat{\rho}'_1(\mathbf{p})$ and $\hat{\rho}''_1(\mathbf{p})$ can be approximately determined independently. Correspondingly, the spin-Hall current, which is given by Eq. (2), can be written as $J_y^x = J_y^x|' + J_y^x|''$ with $J_y^x|'$ and $J_y^x|''$ as the contributions from $\hat{\rho}'_1(\mathbf{p})$ and $\hat{\rho}''_1(\mathbf{p})$, respectively.

It is evident that the solution of Eq. (4), $\hat{\rho}''_1(\mathbf{p})$, has null diagonal elements. Its nonvanishing off-diagonal elements are given by

$$\begin{aligned} (\hat{\rho}''_1)_{12}(\mathbf{p}) &= -(\hat{\rho}''_1)_{21}(\mathbf{p}) = (\hat{\rho}''_1)_{34}(\mathbf{p}) = -(\hat{\rho}''_1)_{43}(\mathbf{p}) \\ &= \frac{\sqrt{3}m}{4\gamma_2 p^3} ieE \sin \theta_p [f_0^H(p) - f_0^L(p)], \end{aligned} \tag{5}$$

with $f_0^H(p) = n_F[\varepsilon_H(p)]$ and $f_0^L(p) = n_F[\varepsilon_L(p)]$, while its remaining elements, such as $(\hat{\rho}''_1)_{13}(\mathbf{p})$, $(\hat{\rho}''_1)_{24}(\mathbf{p})$, etc. vanish. Substituting $\hat{\rho}''_1(\mathbf{p})$ into Eq. (2), $J_y^x|''$ can be written as

$$J_y^x|'' = \frac{eE}{6\pi^2} \int_0^\infty [f_0^H(p) - f_0^L(p)] dp, \tag{6}$$

in agreement with that obtained in Ref.9. Obviously, $J_y^x|''$ is independent of any electron-impurity scattering.

To obtain the other component of the spin-Hall current, we analyze the solution of Eq. (3), $\hat{\rho}'_1(\mathbf{p})$. It is obvious that the diagonal elements of $\hat{\rho}'_1(\mathbf{p})$, which describe the distribution of heavy holes in same band, but having opposite spins, should be the same since the action of the dc electric field on these hole populations is same. Hence, we get $(\hat{\rho}'_1)_{22}(\mathbf{p}) = (\hat{\rho}'_1)_{33}(\mathbf{p})$ and $(\hat{\rho}'_1)_{11}(\mathbf{p}) = (\hat{\rho}'_1)_{44}(\mathbf{p})$. Using these equalities, we find that

there are relations between the off-diagonal elements of the scattering term $\hat{I}^{(1)}$: $\hat{I}_{12}^{(1)} = -\hat{I}_{34}^{(1)}$ and $\hat{I}_{13}^{(1)} = \hat{I}_{24}^{(1)}$, which result in symmetry relations for the $\hat{\rho}'_1(\mathbf{p})$ elements as: $(\hat{\rho}'_1)_{12}(\mathbf{p}) = (\hat{\rho}'_1)_{34}(\mathbf{p})$ and $(\hat{\rho}'_1)_{13}(\mathbf{p}) = -(\hat{\rho}'_1)_{24}(\mathbf{p})$. Hence, considering the Hermitian character of the distribution function, we point out that the determination of the disorder-related spin-Hall effect only requires evaluation of the diagonal elements, $(\hat{\rho}'_1)_{11}(\mathbf{p})$ and $(\hat{\rho}'_1)_{22}(\mathbf{p})$, and the off-diagonal elements, $(\hat{\rho}'_1)_{12}(\mathbf{p})$ and $(\hat{\rho}'_1)_{13}(\mathbf{p})$ (the other off-diagonal elements, which describe the *intraband* polarization, such as $(\hat{\rho}'_1)_{14}(\mathbf{p})$ and $(\hat{\rho}'_1)_{23}(\mathbf{p})$, vanish).

From Eq. (3), it follows that the diagonal $\hat{\rho}'_1(\mathbf{p})$ elements are determined by the integral equation

$$\begin{aligned}
 -e\mathbf{E} \cdot \nabla_{\mathbf{p}} n_{\mathbb{F}}[\varepsilon_{\mu}(p)] &= \pi \sum_{\mathbf{k}} |V(\mathbf{p}-\mathbf{k})|^2 \{a_1(\mathbf{p}, \mathbf{k})[(\hat{\rho}'_1)_{\mu\mu}(\mathbf{p}) \\
 &\quad - (\hat{\rho}'_1)_{\mu\mu}(\mathbf{k})]\Delta_{\mu\mu} + a_2(\mathbf{p}, \mathbf{k})[(\hat{\rho}'_1)_{\mu\mu}(\mathbf{p}) \\
 &\quad - (\hat{\rho}'_1)_{\bar{\mu}\bar{\mu}}(\mathbf{k})\Delta_{\bar{\mu}\bar{\mu}}]\}.
 \end{aligned} \tag{7}$$

Here, $\mu = 1, 2$, respectively, correspond to the heavy- and light-hole bands: $\varepsilon_1(p) \equiv \varepsilon_H(p)$, $\varepsilon_2(p) \equiv \varepsilon_L(p)$, $\bar{\mu} = 3 - \mu$, $\Delta_{\mu\nu} = \delta[\varepsilon_{\mu}(p) - \varepsilon_{\nu}(k)]$. The factors $a_1(\mathbf{p}, \mathbf{k})$ and $a_2(\mathbf{p}, \mathbf{k})$ are associated only with the momentum angles:

$$\begin{aligned}
 a_1(\mathbf{p}, \mathbf{k}) &= \frac{1}{4} \{2 + 6 \cos^2 \phi_{\mathbf{p}\mathbf{k}} [\sin^2 \theta_{\mathbf{p}} - \cos^2 \theta_{\mathbf{k}}] + 6 \cos^2 \theta_{\mathbf{p}} \cos^2 \theta_{\mathbf{k}} [1 + \cos^2 \phi_{\mathbf{p}\mathbf{k}}] \\
 &\quad + 3 \cos \phi_{\mathbf{p}\mathbf{k}} \cos(2\theta_{\mathbf{p}}) \cos(2\theta_{\mathbf{k}})\},
 \end{aligned} \tag{8}$$

$$a_2(\mathbf{p}, \mathbf{k}) = 2 - a_1(\mathbf{p}, \mathbf{k}), \tag{9}$$

where $\phi_{\mathbf{p}\mathbf{k}} \equiv \phi_{\mathbf{p}} - \phi_{\mathbf{k}}$. From Eq. (7), we see that $(\rho'_1)_{\mu\mu}(\mathbf{p})$ is independent of the momentum angle, $\phi_{\mathbf{p}}$, since the potential $V(\mathbf{p}-\mathbf{k})$, as well as the factors $a_1(\mathbf{p}, \mathbf{k})$ and $a_2(\mathbf{p}, \mathbf{k})$, depends on $\phi_{\mathbf{p}}$ and $\phi_{\mathbf{k}}$ only through the combination $\phi_{\mathbf{p}\mathbf{k}}$.

From the off-diagonal parts of Eq. (3), it is clear that $(\hat{\rho}'_1)_{12}(\mathbf{p})$ and $(\hat{\rho}'_1)_{13}(\mathbf{p})$ are similarly effectively independent of ϕ_p . In connection with this, contributions to the disorder-related spin-Hall current, $J_y^x|'$, from $(\hat{\rho}'_1)_{13}(\mathbf{p})$ and $\text{Re}[(\hat{\rho}'_1)_{12}(\mathbf{p})]$ vanish under the ϕ_p -integration in Eq. (2), and only the imaginary part of $(\hat{\rho}'_1)_{12}(\mathbf{p})$ makes a nonvanishing contribution to $J_y^x|'$. Hence,

$$J_y^x|' = \frac{8\sqrt{3}\gamma_2}{m} \sum_{\mathbf{p}} p \left\{ \cos^2 \phi_p \sin \theta_p \text{Im}[(\hat{\rho}'_1)_{12}(\mathbf{p})] \right\}, \quad (10)$$

with

$$\begin{aligned} \text{Im}[(\hat{\rho}'_1)_{12}(\mathbf{p})] &= \frac{\sqrt{3}mm}{4\gamma_2 p^2} \sum_{\mathbf{k}, \mu=1,2} |V(\mathbf{p}-\mathbf{k})|^2 a_3(\mathbf{p}, \mathbf{k}) \\ &\times (-1)^\mu \{ \Delta_{\mu\mu}[(\hat{\rho}'_1)_{\mu\mu}(\mathbf{p}) - (\hat{\rho}'_1)_{\mu\mu}(\mathbf{k})] \\ &- \Delta_{\mu\bar{\mu}}[(\hat{\rho}'_1)_{\mu\mu}(\mathbf{p}) - (\hat{\rho}'_1)_{\mu\bar{\mu}}(\mathbf{k})] \}, \end{aligned} \quad (11)$$

and

$$\begin{aligned} a_3(\mathbf{p}, \mathbf{k}) &= -\frac{1}{2} \{ \sin(2\theta_p) [\cos^2 \theta_k - \sin^2 \theta_k \cos^2 \phi_{\mathbf{p}\mathbf{k}}] \\ &+ \sin(2\theta_k) \cos \phi_{\mathbf{p}\mathbf{k}} [1 - 2 \cos^2 \theta_p] \}. \end{aligned} \quad (12)$$

3. Results and discussions

We first study the disorder effect on the spin-Hall current considering a short-range electron-impurity scattering potential: $V(\mathbf{p}-\mathbf{k}) \equiv u$, with u as a constant. Substituting Eq. (11) into Eq. (2) and performing integration with respect to the angles of \mathbf{p} or \mathbf{k} , respectively, for terms involving $(\hat{\rho}'_1)_{\mu\mu}(\mathbf{k})$ or $(\hat{\rho}'_1)_{\mu\mu}(\mathbf{p})$, we find that the contribution from short-range electron-impurity scattering to the spin-Hall current vanishes, *i.e.* $J_y^x|' = 0$ and the total spin-Hall current is just the intrinsic one, $J_y^x = J_y^x|''$. This result agrees with that obtained in Ref. 10.

Furthermore, we perform a numerical calculation to investigate the effect of long-range electron-impurity collisions on the spin-Hall current in a GaAs bulk semiconductor at $T = 0$ K. The long-range electron-impurity scattering potential is taken from Ref. 11.

From Eq. (6) it follows that the disorder-independent intrinsic spin-Hall current, $J_y^x|^{II}$, can be obtained analytically: $J_y^x|^{II} = eE[k_F^H - k_F^L]/(6\pi^2)$, with k_F^H and k_F^L as the Fermi momenta for heavy- and light-hole bands, respectively. In order to investigate the disorder-related spin-Hall effect, we compute the distribution function $\hat{\rho}_i^l(\mathbf{p})$ at the Fermi surface by solving the integral equation, Eq. (7), for the diagonal $\hat{\rho}_i^l(\mathbf{p})$ elements and determining the off-diagonal elements by Eq. (11). Following that, we can obtain the disorder-related spin-Hall current from Eq. (10), performing the momentum integration.

In Fig. 1, the calculated total and intrinsic spin-Hall conductivities, $\sigma_{yz}^x = J_y^x/E$ and $(\sigma^{II})_{yz}^x = J_y^x|^{II}/E$, and total and intrinsic spin-Hall mobilities, $\mu_{yz}^x = \sigma_{yz}^x/N_p$ and $(\mu^{II})_{yz}^x = (\sigma^{II})_{yz}^x/N_p$, are shown as functions of the hole density. From Fig. 1, it is obvious that the total spin-Hall conductivity is significantly smaller than the intrinsic one: for a typical hole density, $N_p = 1 - 100 \times 10^{21} \text{ m}^{-3}$, the total spin-Hall conductivity σ_{yz}^x varies in the region $0.84 - 1.54 \times 10^{-21} \text{ m}^{-1} \text{ V}^{-1} \text{ s}^{-1}$, whereas the intrinsic one, $(\sigma^{II})_{yz}^x$, ranges from 1.0×10^{-21} to $4.5 \times 10^{-21} \text{ m}^{-1} \text{ V}^{-1} \text{ s}^{-1}$. The contribution to the spin Hall current from electron-impurity scattering always has its sign opposite to that of the intrinsic mechanism. In the region of hole density examined here, the effect of disorder on spin-Hall mobility is relatively small. With increasing hole density, the spin-Hall conductivity first increases and then falls, whereas the spin-Hall mobility monotonically decreases.

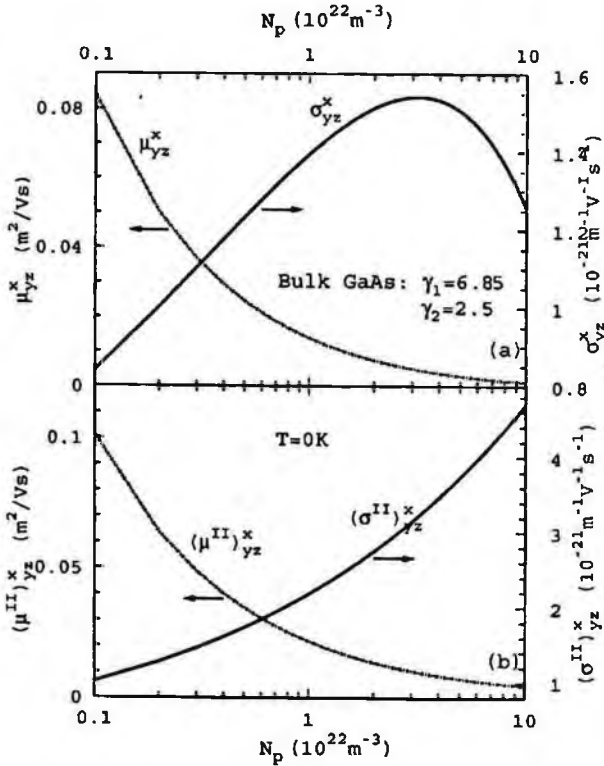


Fig. 1 Dependencies of (a) total σ_{yz}^x and μ_{yz}^x , and (b) intrinsic $(\sigma^{II})_{yz}^x$ and $(\mu^{II})_{yz}^x$, on the hole density in a bulk GaAs semiconductor. The material parameters for GaAs are: $\gamma_1 = 6.85$ and $\gamma_2 = 2.5$. The lattice temperature is $T = 0$ K.

We know that the spin-Hall conductivity in 2D hole systems is of order of e/π .¹² For a typical 2D hole density, $n_p^{(2D)} = 1 \times 10^{12} \text{ cm}^{-2}$, the corresponding spin-Hall mobility is about $0.05 \text{ m}^2/\text{Vs}$. Hence, from Fig. 1, it follows that the spin-Hall mobility in bulk systems has the same order of magnitude as that in 2D hole systems.

4. Conclusions

Employing a nonequilibrium Green's function kinetic equation approach, disorder effects on the spin-Hall current in p -type bulk Luttinger semiconductors have been investigated within the diffusive regime. Considering long-range electron-impurity scattering in the self-consistent Born approximation, we have found that, in contrast to the null effect of short-range disorder on the spin-Hall current, long-range scattering produces a

nonvanishing contribution to the spin-Hall current, independent of impurity density in the diffusive regime. This contribution has its sign opposite to that of the intrinsic one, leading to a significant reduction of the total spin-Hall current. The numerical calculation indicates that, with increasing hole density, the total spin-Hall mobility monotonically decreases, whereas the spin-Hall conductivity first increases and then falls.

Acknowledgments

This work was supported by the Department of Defense through the DURINT program administered by the US Army Research Office, DAAD Grant No. 19-01-1-0592, and by projects of the National Science Foundation of China and the Shanghai Municipal Commission of Science and Technology.

References

1. S. A. Wolf, D. D. Awschalom, R. A. Buhrman, J. M. Daughton, S. von Molnár, M. L. Roukes, A. Y. Chtchelkanova, and D. M. Treger, *Science* **294**, 1488 (2001).
2. M. I. Dyakonov and V. I. Perel, *Phys. Lett. A* **35**, 459 (1971).
3. J. E. Hirsch, *Phys. Rev. Lett.* **83**, 1834 (1999).
4. S. Murakami, N. Nagaosa, and S. C. Zhang, *Science* **301**, 1348 (2003).
5. J. Sinova, D. Culcer, Q. Niu, N. A. Sinitsyn, T. Jungwirth, and A. H. MacDonald, *Phys. Rev. Lett.* **92**, 126603 (2004).
6. Y. K. Kato, R. C. Myers, A. C. Gossard, D. D. Awschalom, *Science* **306**, 1910 (2004).
7. J. Wunderlich, B. Kaestner, J. Sinova, and T. Jungwirth, *Phys. Rev. Lett.* **94**, 047204 (2005).
8. J. M. Luttinger, *Phys. Rev.* **102**, 1030 (1956).
9. S. Murakami, N. Nagaosa, and S. C. Zhang, *Phys. Rev. B* **69**, 235206 (2004).
10. S. Murakami, *Phys. Rev. B* **69**, 241202(R) (2004).
11. R. Grill, *Phys. Rev. B* **46**, 2092 (1992).
12. S. Y. Liu and X. L. Lei, *Phys. Rev. B* **72**, 155314 (2005).

TUNABLE GRID GATED DOUBLE-QUANTUM-WELL FET TERAHERTZ DETECTOR

N. J. M. HORING

*Department of Physics and Engineering Physics, Stevens Institute of Technology,
Hoboken, New Jersey 07030, USA*

S. Y. LIU

*Department of Physics and Engineering Physics, Stevens Institute of Technology,
Hoboken, New Jersey 07030, USA
and Department of Physics, Shanghai Jiaotong University,
1954 Huashan Road, Shanghai 200030, China*

V. V. POPOV

*Institute of Radio Engineering and Electronics (Saratov Division),
Russian Academy of Sciences, 41009 Saratov, Russia*

H. L. CUI

*Department of Physics and Engineering Physics, Stevens Institute of Technology,
Hoboken, New Jersey 07030, USA*

Several aspects of the theory of plasmon resonant DC photoconduction are discussed here, in connection with recent observations involving a THz-irradiated grid-gated double-quantum-well FET.¹ In this, we construct a classical model of nonlinear polarizability to second order in the THz field using a "hydrodynamic" type formulation including the roles of a stress-tensor and friction/viscosity. The resulting second order polarizability exhibits resonant behavior when the THz frequency matches plasmon frequencies of the system, sharply reducing the effectiveness of screened impurity scattering potentials which can admit resonant DC photoconduction. Furthermore, we also show that an asymmetric double-quantum-well system with lateral periodicity can mix optical and acoustic plasmons, giving rise to an interlayer THz field which becomes very strong when tuned by gate voltage into the "mode-mode-repulsion" regime wherein the optical and acoustic modes equally share amplitude. This can enhance interlayer electron tunneling and may contribute to photoconductivity.

Keywords: Terahertz Detector; Resonant Photoconduction; Double Quantum Well.

1. Introduction

This paper is concerned with several aspects of the theory of plasmon-resonant dc photoconduction in a grid-gated double-quantum-well field effect transistor subject to THz radiation observed by Peralta and Allen, *et al.*¹ In earlier electrodynamic absorption studies,² we definitively established that the dc photoconductance resonances occur when the incident terahertz frequencies match those of the plasmons of the grid-gated double-quantum-well system. On a fundamental level, the involvement of the THz field and plasmon resonant phenomenology in the source-drain dc steady state conduction clearly indicates nonlinear conditioning of the system. In this connection, we have examined the nonlinear polarizability of an electron plasma in a terahertz field in an iterated shielded potential approximation, finding it to exhibit resonant behavior when the incident THz frequency matches the plasma frequency. Our first calculation exhibited this behavior to first order in the THz field,³ and, recently, we have further supported this with a nonequilibrium Green's function calculation of the polarizability to *second order*⁴ in the THz field. This matter is further studied below using a classical model to calculate the second order polarizability and show its resonant features. Such resonant behavior of the polarizability and hence the direct dielectric function, can sharply reduce its inverse, the dynamic, nonlocal and spatially inhomogeneous screening function of the system. In turn, this reduces the effectiveness of the screened resistive impurity scattering potentials, admitting resonant increase of dc conduction when the THz frequency matches the plasma frequency.

The Peralta-Allen experiment also exhibited substantial dependence of the resonance frequencies on a magnetic field parallel to the quantum wells, which might be used as a tuning agent (in addition to tuning by gate voltage which controls population of the quantum wells). To explore the role of a parallel magnetic field, we have derived an exact closed form analytic solution for the Green's function describing the dynamics of electrons in a symmetric narrow double quantum well system in a parallel magnetic field.⁵ The Landau quantized electrons are in skipping states of motion between the walls of the individual quantum wells with in-plane translation parallel to the well walls and tunnelling between the wells, coupled with the zero-field lowest-subband energy states of the double well system. This Green's function will facilitate calculation of the "ring diagram" polarizability of the double well system in a parallel magnetic field, which determines the plasmon resonance spectrum as a function of the magnetic field.

2. Polarizability to second order in THz field: Classical model

In so far as the dielectric response of a solid-state plasma is concerned, we have already provided explicit results for the nonlinear polarizability in an iterated shielded potential approximation to first order in a THz field, both quantum mechanically in terms of nonequilibrium Green's functions, and also classically.³ Our recent second order quantum mechanical calculation is to be presented elsewhere.⁴ Here, we discuss our classical analysis of the nonlinear polarizability to second order in the THz field. In this, we employ a classical "hydrodynamic"-type formulation which includes the role of a viscous-friction coefficient γ and the stress tensor $\vec{\Pi}$ in the determination of the density perturbation to third order in the THz field (which yields the polarizability to second order). Thus, we employ the equation

$$\rho \frac{\partial \vec{v}}{\partial t} + \gamma \rho \vec{v} + \rho [\vec{v} \cdot \vec{\nabla}] \vec{v} = -c \rho \vec{\nabla} V - \frac{1}{m} \vec{\nabla} \cdot \vec{\Pi}, \quad (1)$$

where V is the shielded(dynamic, nonlocal) potential, $c = e/m$ and

$$\vec{\nabla} \cdot \vec{\Pi} = m \beta^2 \vec{\nabla} \rho, \quad (2)$$

taken jointly with the equation of continuity relating density, ρ , and velocity field, \vec{v} ,

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{v}) = 0. \quad (3)$$

β is a velocity parameter, proportional to the Fermi velocity.

The calculated results below determine the polarizability at space-time point $\mathbf{l} = \vec{r}_1, t_1$, $\alpha(1,2)$, through the relation

$$\alpha(1,2) = - \int d^4 3\nu_c (1-3)R(3,2), \quad (4)$$

where ν_c is the Coulomb potential and R is given by the variational derivative of density with respect to V as

$$R(3,2) = \frac{\delta\rho(3)}{\delta V(2)}. \quad (5)$$

At this stage we assume translational invariance and we Fourier transform in position-space and time. Sorting out the first, second and third orders of $\rho = \rho_0 + \rho_1 + \rho_2 + \rho_3$ and $\bar{v} = \bar{v}_1 + \bar{v}_2$ in the shielded THz field (respectively), $V(\bar{r}, t) \rightarrow V(\bar{p}, \omega)$, we obtain

$$\rho_1 = \rho_0 c p^2 A(\bar{p}, \omega) V(\bar{p}, \omega), \quad (6)$$

$$\bar{v}_1 = \omega c \bar{p} A(\bar{p}, \omega) V(\bar{p}, \omega), \quad (7)$$

and

$$\begin{aligned} \rho_2 &= \rho_0 c^2 A(\bar{p}, \omega) \int \frac{d\bar{p}' d\omega'}{(2\pi)^4} \\ &\times A(\bar{p}', \omega') A(\bar{p} - \bar{p}', \omega - \omega') r_2(\bar{p}, \omega; \bar{p}', \omega') V(\bar{p}', \omega') V(\bar{p} - \bar{p}', \omega - \omega'), \end{aligned} \quad (8)$$

$$\begin{aligned} \bar{v}_2 &= \frac{c^2}{\omega + i\gamma} \int \frac{d\bar{p}' d\omega'}{(2\pi)^4} \\ &\times A(\bar{p}', \omega') A(\bar{p} - \bar{p}', \omega - \omega') \bar{w}(\bar{p}, \omega; \bar{p}', \omega') V(\bar{p}', \omega') V(\bar{p} - \bar{p}', \omega - \omega'), \end{aligned} \quad (9)$$

and

$$\begin{aligned} \rho_3 &= \rho_0 c^3 A(\bar{p}, \omega) \int \frac{d\bar{p}' d\omega' d\bar{p}'' d\omega''}{(2\pi)^8} A(\bar{p}'', \omega'') A(\bar{p}' - \bar{p}'', \omega' - \omega'') A(\bar{p} - \bar{p}', \omega - \omega') \\ &\times r_3(\bar{p}, \omega; \bar{p}', \omega'; \bar{p}'', \omega'') V(\bar{p}'', \omega'') V(\bar{p}' - \bar{p}'', \omega' - \omega'') V(\bar{p} - \bar{p}', \omega - \omega'), \end{aligned} \quad (10)$$

where

$$A(\bar{p}, \omega) = [\omega^2 + i\omega\gamma - \beta^2 p^2]^{-1}, \quad (11)$$

$$\begin{aligned} \bar{w}(\bar{p}, \omega; \bar{p}', \omega') &= \omega'(\omega - \omega')p'^2(\bar{p} - \bar{p}') - \beta^2 p'^2(\bar{p} - \bar{p}')^2 \bar{p}' \\ &+ \beta^2 \bar{p} A(\bar{p}, \omega) \left\{ \omega'(\omega - \omega') [(\bar{p} - \bar{p}')^2(\bar{p} \cdot \bar{p}') + [\bar{p} \cdot (\bar{p} - \bar{p}')] p'^2] \right. \\ &\left. + (\bar{p} - \bar{p}')^2(\bar{p} \cdot \bar{p}') [\omega'^2 + i\omega'\gamma - \beta^2 p'^2] \right\}, \end{aligned} \quad (12)$$

$$\begin{aligned} r_2(\bar{p}, \omega; \bar{p}', \omega') &= (\bar{p} - \bar{p}')^2(\bar{p} \cdot \bar{p}') [\omega'^2 + i\omega'\gamma - \beta^2 p'^2] \\ &+ \omega'(\omega - \omega') \left\{ (\bar{p} - \bar{p}')^2(\bar{p} \cdot \bar{p}') + [\bar{p} \cdot (\bar{p} - \bar{p}')] p'^2 \right\}, \end{aligned} \quad (13)$$

and

$$\begin{aligned} r_3(\bar{p}, \omega; \bar{p}', \omega') &= \frac{(\omega - \omega')}{\omega' + i\gamma} \left\{ 2(\bar{p} - \bar{p}')^2 [\bar{p} \cdot \bar{w}(\bar{p}', \omega'; \bar{p}'', \omega'')] \right. \\ &+ [\bar{p} \cdot (\bar{p} - \bar{p}')] [\bar{p}' \cdot \bar{w}(\bar{p}', \omega'; \bar{p}'', \omega'')] \left. \right\} \\ &+ p''^2 [\bar{p} \cdot (\bar{p}' - \bar{p}'')] (\omega' - \omega'') (\omega - \omega') (\bar{p} - \bar{p}')^2 \\ &+ [\bar{p} \cdot (\bar{p} - \bar{p}')] [(\omega - \omega') \omega' + A^{-1}(\bar{p} - \bar{p}', \omega - \omega')] \\ &\times A(\bar{p}', \omega') r_2(\bar{p}', \omega'; \bar{p}'', \omega''). \end{aligned} \quad (14)$$

Considering the externally impressed THz field to have frequency ω_E ,

$$U(\vec{r}, t) = U_0 \cos \omega_E t, \quad (15)$$

we obtain the shielded, effective THz field in terms of the lowest order screening function K evaluated at the incident frequency, ω_E , as

$$V(\bar{p}, \omega) = \pi U_0 [\delta(\omega - \omega_E) + \delta(\omega + \omega_E)] K(\bar{p}, \omega_E). \quad (16)$$

Thus, we have

$$R(\vec{r}_1, t_1; \vec{r}_2, t_2) = \frac{\delta \rho_1(\vec{r}_1, t_1)}{\delta V(\vec{r}_2, t_2)} + \frac{\delta \rho_2(\vec{r}_1, t_1)}{\delta V(\vec{r}_2, t_2)} + \frac{\delta \rho_3(\vec{r}_1, t_1)}{\delta V(\vec{r}_2, t_2)} \quad (17)$$

with the following:

$$\frac{\delta \rho_1(\vec{r}_1, t_1)}{\delta V(\vec{r}_2, t_2)} = \rho_0 c \int \frac{d\vec{p} d\omega}{(2\pi)^4} e^{i\vec{p} \cdot (\vec{r}_1 - \vec{r}_2)} e^{-i\omega(t_1 - t_2)} \bar{p}^2 A(\bar{p}, \omega), \quad (18)$$

$$\begin{aligned} \frac{\delta\rho_2(\bar{r}_1, t_1)}{\delta V(\bar{r}_2, t_2)} &= \rho_0 c^2 \pi \sum_{\pm} \int \frac{d\bar{p} d\bar{p}' d\omega}{(2\pi)^8} e^{i(\bar{p}\bar{r}_1 - \omega t_1)} e^{i\bar{p}'(\omega \mp \omega_E)} A(\bar{p}, \omega) \\ &\times \left\{ A(\bar{p} - \bar{p}', \omega \mp \omega_E) A(\bar{p}', \pm \omega_E) r_2(\bar{p}, \omega; \bar{p}', \pm \omega_E) K(\bar{p}', \omega_E) e^{-i\bar{r}_2 \cdot (\bar{p} - \bar{p}')} \right. \\ &\left. + A(\bar{p} - \bar{p}', \pm \omega_E) A(\bar{p}', \omega \mp \omega_E) r_2(\bar{p}, \omega; \bar{p}', \omega \mp \omega_E) K(\bar{p} - \bar{p}', \omega_E) e^{-i\bar{r}_2 \cdot \bar{p}'} \right\}, \end{aligned} \quad (19)$$

$$\begin{aligned} \frac{\delta\rho_3(\bar{r}_1, t_1)}{\delta V(\bar{r}_2, t_2)} &= \frac{1}{4} \rho_0 c^3 \sum_{\pm} \int \frac{d\bar{p} d\omega}{(2\pi)^4} \int \frac{d\bar{p}'}{(2\pi)^3} \int \frac{d\bar{p}''}{(2\pi)^3} A(\omega, \bar{p}) e^{i(\bar{p}\bar{r}_1 - \omega t_1)} \\ &\left\{ e^{i\omega t_1} \left[A(\bar{p} - \bar{p}', \omega) A(\bar{p}' - \bar{p}'', \pm \omega_E) A(\bar{p}'', \mp \omega_E) r_2(\bar{p}, \omega; \bar{p}', 0; \bar{p}'', \mp \omega_E) e^{-i\bar{r}_2 \cdot (\bar{p} - \bar{p}')} \right. \right. \\ &+ A(\bar{p} - \bar{p}', \pm \omega_E) A(\bar{p}' - \bar{p}'', \omega) A(\bar{p}'', \mp \omega_E) r_2(\bar{p}, \omega; \bar{p}', \omega \mp \omega_E; \bar{p}'', \mp \omega_E) e^{-i\bar{r}_2 \cdot (\bar{p}' - \bar{p}'')} \\ &+ A(\bar{p} - \bar{p}', \pm \omega_E) A(\bar{p}' - \bar{p}'', \mp \omega_E) A(\bar{p}'', \omega) r_2(\bar{p}, \omega; \bar{p}', \omega \mp \omega_E; \bar{p}'', \omega) e^{-i\bar{r}_2 \cdot \bar{p}''} \left. \right] \\ &+ e^{i(\omega \mp 2\omega_E)t_1} \left[A(\bar{p} - \bar{p}', \omega \mp 2\omega_E) A(\bar{p}' - \bar{p}'', \pm \omega_E) A(\bar{p}'', \pm \omega_E) \right. \\ &\times r_2(\bar{p}, \omega; \bar{p}', \pm 2\omega_E; \bar{p}'', \pm \omega_E) e^{-i\bar{r}_2 \cdot (\bar{p} - \bar{p}'')} \\ &+ A(\bar{p} - \bar{p}', \pm \omega_E) A(\bar{p}' - \bar{p}'', \omega \mp 2\omega_E) A(\bar{p}'', \pm \omega_E) \\ &\times r_2(\bar{p}, \omega; \bar{p}', \omega \mp \omega_E; \bar{p}'', \pm \omega_E) e^{-i\bar{r}_2 \cdot (\bar{p}' - \bar{p}'')} \\ &+ A(\bar{p} - \bar{p}', \pm \omega_E) A(\bar{p}' - \bar{p}'', \pm \omega_E) A(\bar{p}'', \omega \mp 2\omega_E) \\ &\left. \times r_2(\bar{p}, \omega; \bar{p}', \omega \mp \omega_E; \bar{p}'', \omega \mp 2\omega_E) e^{-i\bar{r}_2 \cdot \bar{p}''} \right] \left. \right\}. \end{aligned} \quad (20)$$

These results determine $R(3,2)$ in Eq. (5), hence they yield $\alpha(1,2)$ upon integration with the Coulomb potential (Eq. (4)). Since the lowest order screening function, K , is evaluated at ω_E in the second order results for $\delta\rho'\delta V$ above, both the polarizability and the direct dielectric function have resonances when ω_E matches a plasmon frequency.

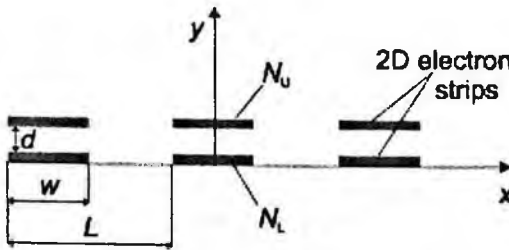


Fig. 1 Schematic of a bilayer of two-dimensional electron strips and coordinate system.

3. Tunability Considerations; Interlayer electric field

In regard to tunability by gate voltage, we have also analyzed the electromagnetic terahertz (THz) response of a bilayer of density-modulated two-dimensional electron gases used to model the actual double-quantum-well electron channel of the grid-gated field-effect transistor, in which strong THz photoresponse was observed (Figure 1). We found that vertical asymmetry and lateral periodicity of the double quantum well mix the optical and acoustic plasmons, hybridizing them and giving rise to an intense THz electric field between the 2D quantum well layers when the interlayer separation, d , causes the two mixed modes (optical-like with some admixture of acoustic, and acoustic-like with some admixture of optical) to cross. Figure 2 shows the hybridized modes and their crossing ("anticrossing") in terms of transmission spectra for various values of d and Figure 3 shows the amplitude of the interlayer THz electric field between the two wells as a function

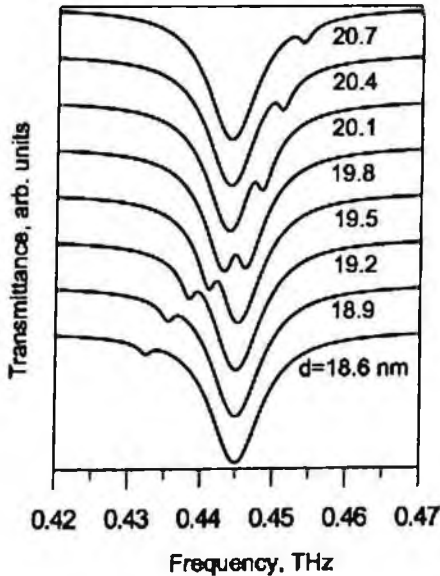


Fig. 2 Terahertz transmission spectra for various values of interlayer separation d . Anticrossing of the optical-like (stronger) and acoustic-like (weaker) resonances takes place at $d = 19.8$ nm.

of position along one period of the double quantum well channel, which increases hugely in the anticrossing regime (when $d = 19.8$ nm) to become two and a half orders of magnitude greater than the amplitude of the incoming THz electric field.⁶ This can significantly enhance tunneling between the two quantum wells and thereby influence the source-drain dc transport.

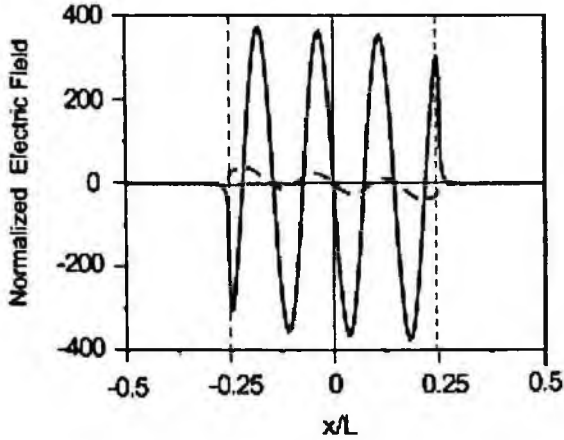


Fig. 3 Amplitudes of normal-to-plane interwell electric field oscillations at the center plane between the two layers of 2D electron strips in the anticrossing regime ($d=19.8$ nm, solid curve) and far from the anticrossing regime ($d=27$ nm, dashed curve). The amplitude of the interwell electric field is normalized to the amplitude of the electric field of the incident terahertz wave.

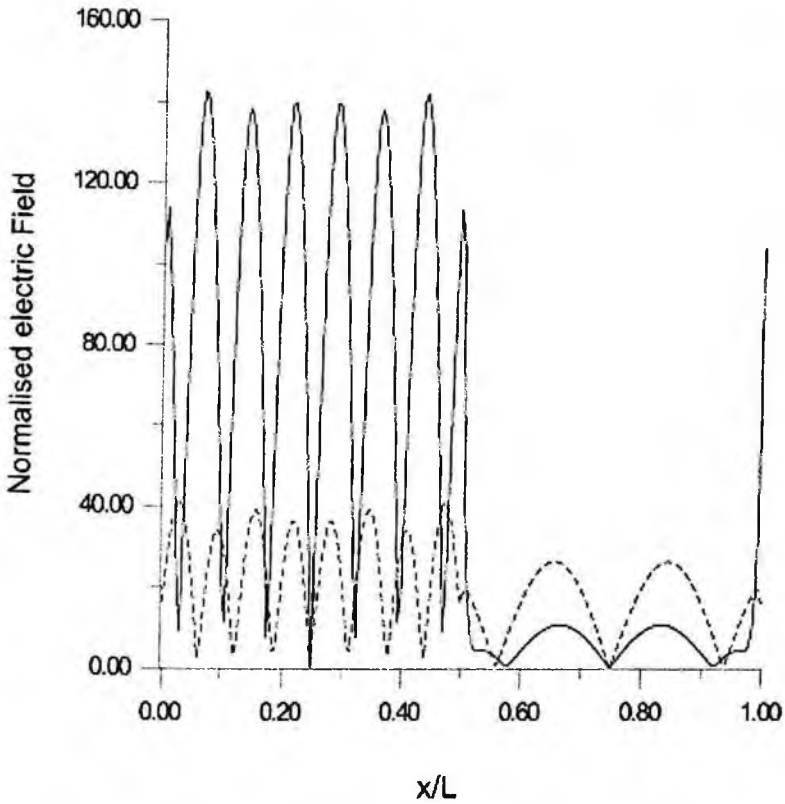


Fig. 4 Amplitudes of the normal-to-plane interlayer electric field oscillations at the center plane between the two of 2D electron layers in the anticrossing regime ($U_g = -1.83$ V, solid curve) and far away from the anticrossing regime ($U_g = -1.5$ V, dashed curve). The amplitude of the interwell electric field is normalized to the amplitude of electric field in the incident terahertz wave.

We have now found that the anticrossing regime, which is characterized by the huge interlayer electric field and by the fact that the very strong optical-like plasmon shares its amplitude equally with the otherwise weak acoustic-like plasmon, can be accessed by tuning the gate voltage, U_g .⁷ This is a much more useful mechanism than the earlier mechanism we examined in terms of variation of the interlayer distance, d , between the quantum wells. In Figure 4, we plot the amplitude of the interlayer electric field at the center plane between the quantum wells as a function of position along one period of the double quantum well channel for two gate voltages, $U_g = -1.83$ V (solid curves) and

$U_g = -1.5$ V (dashed curves). The anticrossing regime is accessed at $U_g = -1.83$ V, with dramatic increase of the amplitude of the interlayer electric field in the ungated positions of the channel. The parameters involved are typical for the double quantum well FET exhibiting strong THz photoresponse.^{1,2} Full detail will be provided in a forthcoming publication.⁷

4. Conclusions

We have examined several aspects of the theory of THz response of a grid-gated double quantum well FET exhibiting resonant dc photoconductivity when the incident THz frequency matches plasmon frequencies of the system. Analyzing the effect of the THz field on the nonlinear polarizability, we have calculated its role to second order using a simple "hydrodynamic" classical model. In this, we explicitly exhibit resonant behavior in the second order nonlinear polarizability, which sharply reduces the nonlinear screening function at the plasmon-resonant frequencies. This, in turn, reduces the effectiveness of screened scattering potentials, partially disabling them and admitting resonant dc current increase.

We have also shown that such a system can be driven into the anticrossing regime between optical and acoustic plasma resonances by tuning the gate voltage. The amplitude of the interlayer THz electric field in the ungated portions of the channel increases dramatically in the anticrossing regime of "mode-mode" repulsion. This strong interlayer THz electric field may substantially enhance interlayer electron tunneling which, in turn, may contribute to the physical mechanism underlying the strong THz photoresponse in the dc source-drain conductance. Finally, though it is not discussed here, we have also recently completed the derivation of an exact Green's function for a double quantum well in a parallel magnetic field in the case of two non-symmetric narrow wells. The resulting Green's function is obtained in an analytical, closed-form expression that is highly tractable, describing Landau-quantized electrons in skipping states of motion between the walls of the individual, inequivalent quantum wells with in-plane translation parallel to the well walls and tunneling between the wells, coupled with the zero-field lowest-subband energy states (involving bonding and antibonding features) of the double well system.

Acknowledgments

This work was supported by the U.S. Department of Defense through the DURINT program administered by the US Army Research Office, DAAD Grant No. 19-01-1-05 92.

References

1. X.G. Peralta, S.J. Allen, M.C. Wanke, N.B. Harff, J.A. Simmons, M.P. Lilly, J.L. Reno, P.J. Burke and J.P. Eisenstein, *Appl. Phys. Lett.* **81**, 1627 (2002).
2. V.V. Popov, O. V. Polischuk, T.V. Teperik, X.G. Peralta, S.J. Allen, N.J.M. Horing, and M.C. Wanke, *J. Appl. Phys.* **94**, 3556 (2003).
3. N. J. M. Horing, *Solid State Comm.* **137**, 338 (2006).
4. N. J. M. Horing, S. Y. Liu, and H. L. Cui, *Proc. IEEE-NANO-2006*, accepted.
5. N. J. M. Horing, B. Dong, and H. L. Cui, *Proc. IEEE-NANO-2005*, 229 (2005).
6. V.V. Popov, G. M. Tsymbalov, and N. J. M. Horing, *J. Appl. Phys.*, accepted for publication.
7. V.V. Popov, G. M. Tsymbalov, and N. J. M. Horing, submitted for publication.

LASER-IONIZATION MASS SPECTROMETRY OF EXPLOSIVES AND CHEMICAL WARFARE SIMULANTS

DAVID L. HUESTIS

*Molecular Physics Laboratory, SRI International
Menlo Park, CA 94025
david.huestis@sri.com*

CHRISTOPHER MULLEN

*Molecular Physics Laboratory, SRI International
Menlo Park, CA 94025
christopher.mullen@sri.com*

MICHAEL J. COGGIOLA

*Molecular Physics Laboratory, SRI International
Menlo Park, CA 94025
michael.coggiola@sri.com*

HARALD OSER

*Molecular Physics Laboratory, SRI International
Menlo Park, CA 94025
harald.oser@sri.com*

The objective of the present study was to better understand the photophysics of explosives and chemical warfare simulants in order to develop better performing analytical tools. Photoionization mass spectra were taken using three optical schemes. The first was resonance-enhanced multiphoton ionization (REMPI) using few-ns duration 248 or 266 nm laser pulses. The second scheme was non-resonant multiphoton ionization (MPI) using 100 fs duration laser pulses at wavelengths between 325 and 795. The third approach was single photon ionization (SPI) using few-ns duration 118 nm laser pulses. For all the molecules investigated, mass spectra resulting exposure to ns-duration 248 or 266 nm laser pulses consisted of only low molecular weight fragments. Using fs-duration laser pulses produced more complicated, potentially analyzable, fragmentation patterns, usually with some parent peak. Single photon ionization gave the best results, with mass spectra consisting of almost only parent peak, except for the case of TATP.

Keywords: mass spectrometry, photoionization, explosives, chemical warfare simulants, trace species detection

1. Introduction

In this work we examine the applicability of Laser Ionization Time-of-Flight Mass Spectrometry to the detection of explosives, explosive-related compounds (ERCs),

chemical agent simulants and precursors, using three different laser ionization methods. We seek a scheme that has the ability to discriminate between explosives or ERCs, chemical agent simulants, and background signals. We also investigated the limits of detection on the way towards a compact, sensitive, rugged, and fast sensing device.

Laser ionization can provide significant enhancements in selectivity and sensitivity for mass-spectrometric identification and quantification of trace species. The high laser powers available from commercial sources produce efficient ionization, and in many cases the mass spectrum consists of a single mass corresponding to the molecular ion or parent peak. As an example, resonance enhanced multiphoton ionization (REMPI) of jet-cooled dioxins allows for selective detection of individual isomers of the same molecular weight.¹ For molecules lacking conveniently placed intermediate excited states, non-resonant multiphoton ionization (MPI) can be successful, but may lead to fragment ions resulting from absorption of additional photons. Alternatively, one can first use harmonic generation to produce vacuum ultraviolet light with a photon energy sufficient for single photon ionization (SPI).^{2,3}

Laser ionization mass spectrometry of explosives and chemical warfare simulants has been studied using nanosecond laser pulses. Primary ions observed in many of these studies were NO^+ and PO^+ , which are not unique signatures of the parent molecules. It is now widely accepted that after absorption of the first photon, the parent molecule dissociates on a time scale of about 100 femtoseconds (fs). We can attempt to compensate for this rapid dissociation by using ultrafast laser pulses of a corresponding time duration.⁴ Here we compare the nanosecond, ultrafast, and SPI approaches.

2. Experimental

The investigations were conducted using two separate laser photoionization time-of-flight mass spectrometer instruments, both of which are described briefly. The fs system includes a reflectron time of flight mass spectrometer (with a nominal mass resolution of $m/\Delta m \sim 1000$, Stefan Kaesdorf, Munich, Germany) with a capillary inlet system for sample introduction into the ionization region of the mass spectrometer. The fs laser system consists of a Ti:Sapphire oscillator, regenerative amplifier, and tunable OPA (Spectra Physics Tsunami, Spitfire, and OPA 800C), which produced nominally 130 fs pulses with energies of 840 μJ , 30 μJ , and 3 μJ , at wavelengths of 795 nm, 500nm, and 325 nm, respectively, at a repetition rate of 1 kHz. The laser was focused into the ionization region with a 150 mm lens, and the resulting mass spectra were recorded with a 1 GHz digitizer (Acqiris AP235).

The SPI system, shown in Fig. 1, is very similar in all respects. The laser system consists of a Continuum Powerlite Precision 9010 Nd:YAG with a 5 ns pulse width, and a repetition rate of 10 Hz. The 118.2 nm light is produced by frequency tripling the third harmonic output (355 nm) of the Nd:YAG laser. To accomplish this, the 355 nm laser beam (30 mJ pulse energy) is focused into a stainless steel cell filled with Xe gas. A MgF_2 lens serves as the output window for the cell, and also refocuses the 118 nm light into the ion source region. Because the residual 355 nm and product 118 nm beams are

not copropagating, there was very little contribution from the former to the observed ion signals. Ions produced by SPI are detected and mass analyzed using the same compact reflectron time of flight mass spectrometer model as applied in the fs experiments, and recorded by a 1 GHz digitizer (Signatec, Corona, CA). Instrument control and data acquisition were controlled using custom in-house software based on the Labview (National Instruments, Austin, TX) programming environment.

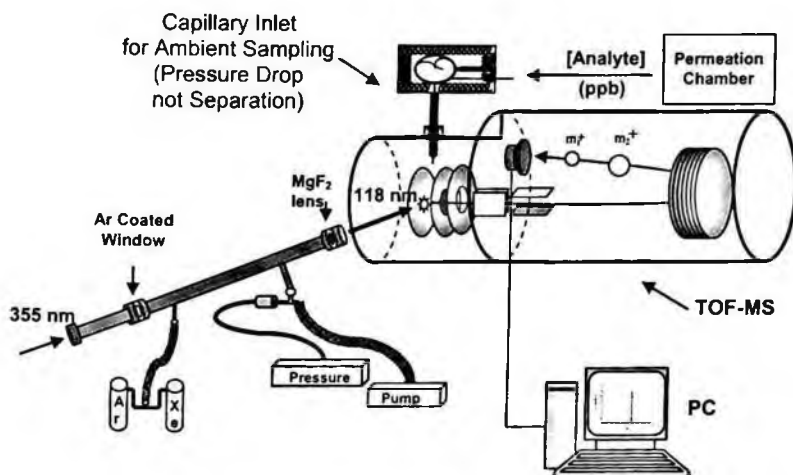


Fig. 1. SPI Apparatus Schematic Diagram

3. Results

Mass spectra of a variety of explosive and chemical warfare agent simulants, including nitrobenzene (NB), dinitrobenzene (DNB), dinitrotoluene (DNT), trinitrotoluene (TNT), triacetone triperoxide (TATP), dimethylmethylphosphonate (DMMP), and chloroethylethylsulfide (CEES) were taken as a functions of laser wavelength, pulse duration, and power density. Fig. 2 shows the mass spectrum for 1,3-dinitrobenzene (DNB) using nanosecond REMPI. Only one mass peak is discernable, $m/c = 40$ amu, corresponding to the nitric oxide fragment ion NO^+ . The generation of nitric oxide ions results from a sequence of photoabsorptions:



Fig. 3 shows the DNB mass spectrum for using femtosecond laser pulses. Throughout our studies, the use of ultrafast ionization produces mass spectra with fragmentation patterns characteristic of individual molecules being investigated. In addition, the presence of the parent molecular ion improves the prospects for molecule-specific identification.

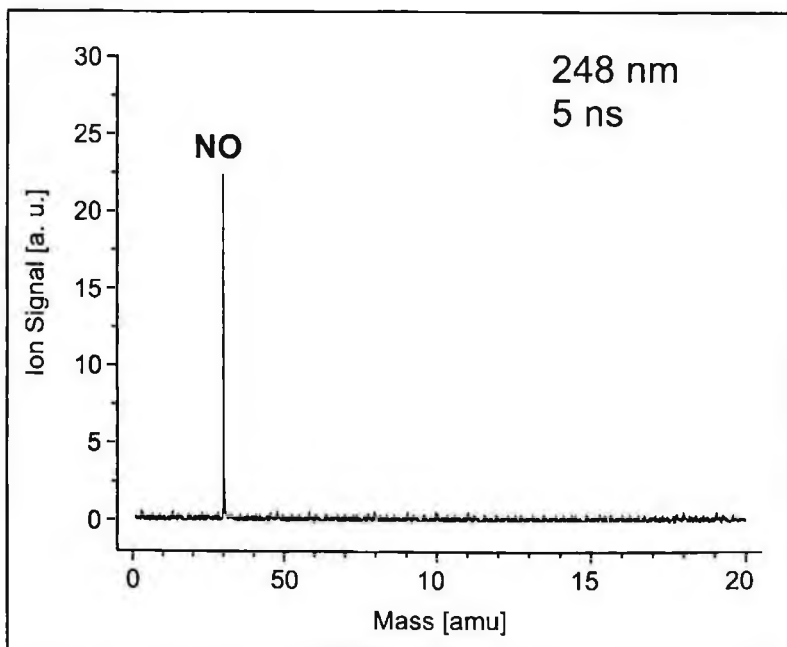


Fig. 2. Photoionization mass spectrum of 1,3-dinitrobenzene (DNB) using 5 ns-duration 248 nm laser pulses

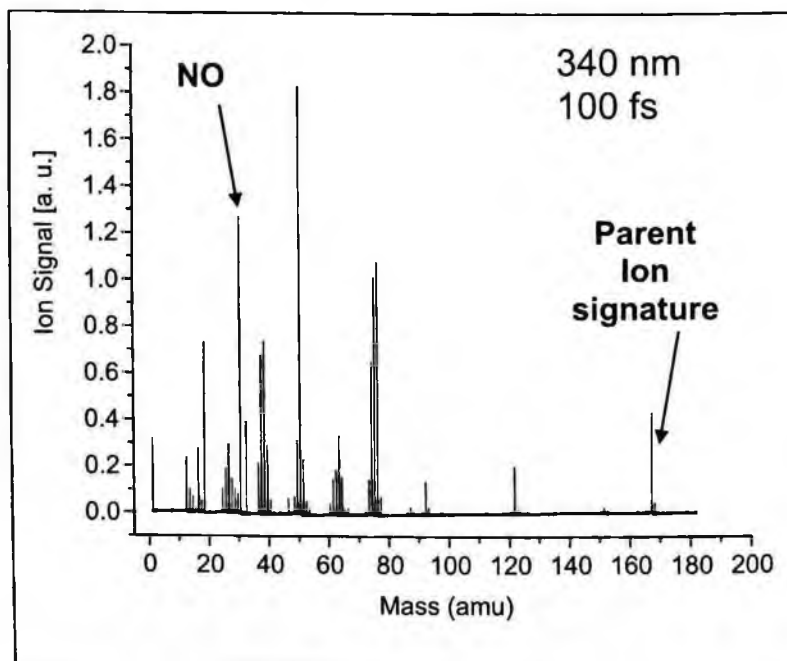


Fig. 3. Photoionization mass spectrum of 1,3-dinitrobenzene (DNB) using 100 fs-duration 340 nm laser pulses

Identification of the same compounds with the SPI method was found to be even more robust because their measured mass spectra are dominated by the parent ion, due to the low excess energy upon ionization. Fig. 4 shows SPI mass spectra for hexane, DNT, and TATP. The corresponding SPI mass spectrum for DMMP is shown in Fig. 5. Only in the case of TATP is there significant fragmentation, which indicates that the TATP parent ion is stable with respect to dissociation, but that photoionization at 118 nm (10.5 eV) leaves most of the ions with enough vibrational energy to dissociate.

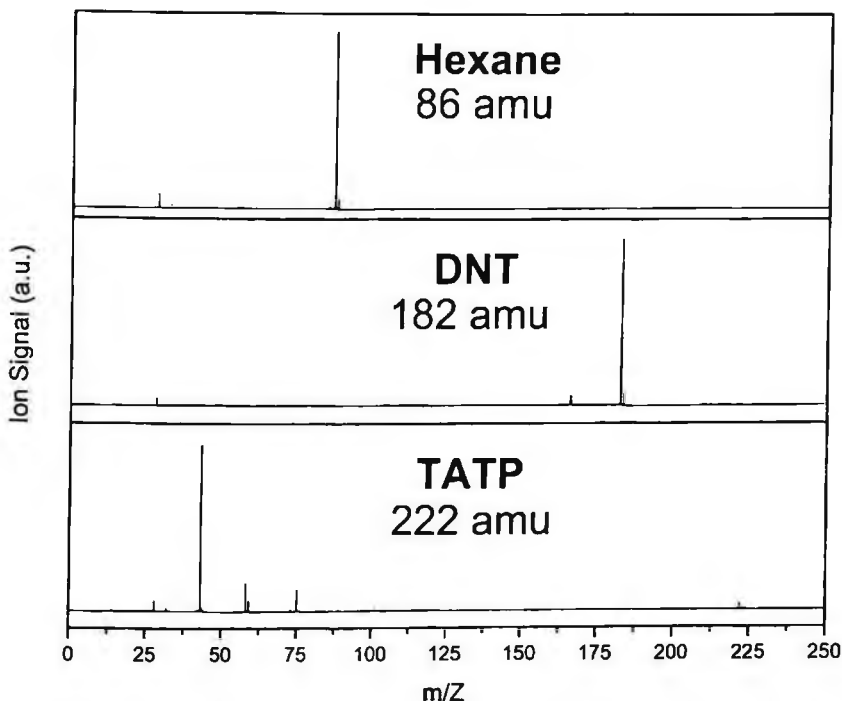


Fig. 4. Single-photon-ionization (SPI) mass spectra of hexane, DNT, and triacetone triperoxide (TATP).

The limits of detection (LODs) for nitrobenzene, dinitrotoluene, and DMMP using SPI were also measured. The LOD for nitrobenzene and dinitrotoluene were determined to be ~ 13 ppb (S:N $\sim 2:1$), ~ 40 ppb (S:N $\sim 2:1$), and ~ 35 ppb (S:N $\sim 3:1$), respectively. In Fig. 6, the calibration curve for the DMMP sensitivity measurements is presented showing a linear response over an order of magnitude in concentration. The ppb-level sensitivity of our method is comparable to the sensitivity measurements that were previously reported for the ns REMPI studies of similar compounds.

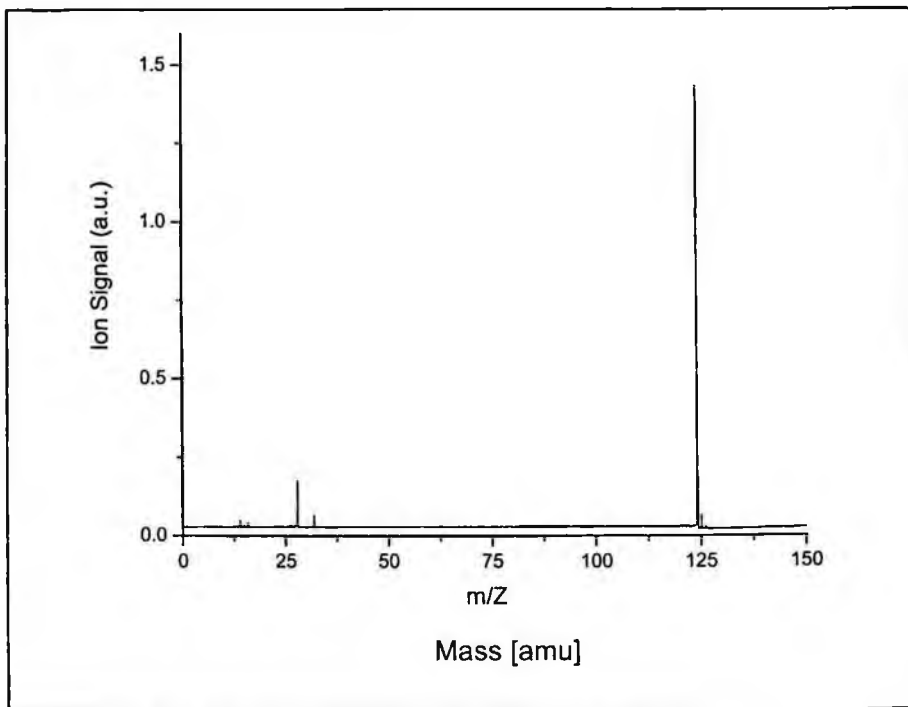


Fig. 5. SPI mass spectrum of dimethylmethylphosphonate (DMMP).

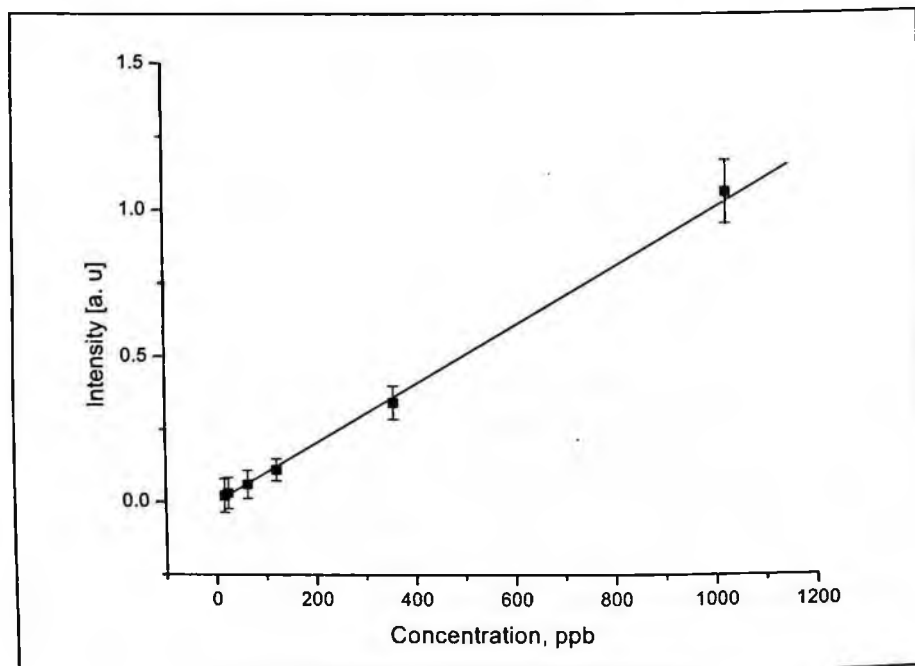


Fig. 6. Limits of detection calibration curve for SPI mass spectrometry of DMMP

4. Future Work

Subsequent investigations with the fs approach will focus on an analysis of the intensity and the mass peaks present as a function of laser wavelength and power density, which should allow for a highly specific detection of this class of compounds. Future work will also include SPI experiments on other explosives and related compounds, as well as chemical warfare agent simulants. Continued, accurate, sensitivity measurements will also be made which will determine the applicability of this technique to explosives detection. Additionally, real world mixtures using both the SPI and MPI techniques need to be examined.

5. Acknowledgements

This work was supported by the National Science Foundation under grant CHE-0321288 and by SRI International internal funds. The authors thank Dr. Jeffrey Bottaro for synthesis of TATP.

References

1. H. Oser, K. Copic, M. J. Coggiola, G. W. Faris and D. R. Crosley, Congener-Specific Detection of Dioxins Using Jet-REMPI, *Chemosphere*, **43**, 469-477 (2001).
2. J. B. Pallix, U. Schüle, C. H. Becker and D. L. Huestis, Advantages of Single-Photon Ionization over Multiphoton Ionization for Mass Spectrometric Surface Analysis of Bulk Organic Polymers, *Anal. Chem.* **61**, 805-811 (1989).
3. C. Mullen, A. Irwin, B. V. Pond, D. L. Huestis, M. J. Coggiola and H. Oser, Detection of Explosives and Explosives-Related Compounds by Single Photon Laser Ionization Time-of-Flight Mass Spectrometry, *Anal. Chem.* **78**, 3807-3814 (2006).
4. C. Mullen, D. Huestis, M. Coggiola and H. Oser, Laser Photon Ionization of Triacetone Triperoxide (TATP) by Femtosecond and Nanosecond Laser Pulses, *Int. J. Mass Spec.* **252**, 69-72 (2006).

IN-LINE OPTICAL FIBER STRUCTURES FOR ENVIRONMENTAL SENSING

A. DHAWAN

*ECE Dept, North Carolina State University, Box 7911
Raleigh, NC 27695, USA
adhawan@ncsu.edu*

M. D. GERHOLD

*Electronics Division, Army Research Office, 4300 S. Miami Blvd,
Durham, NC, 27703, USA
Mike.Gerhold@us.army.mil*

J. F. MUTH*

*ECE Dept, North Carolina State University, Box 7911
Raleigh, NC 27695, USA
muth@ncsu.edu*

Surface plasmon and evanescent wave sensors are attractive for chemical and biological sensing applications. They can work in aqueous media and when used in conjunction with the appropriate surface chemistry they can have high specificity and high sensitivity. However, most surface plasmon sensors are relatively complex as they are based on the use of attenuated total internal reflection to excite surface plasmon resonance in a thin gold film and require light to be incident at the appropriate angle and polarization. Other surface plasmon and optical affinity sensors have used the evanescent waves in planar waveguides to interact with the environment. These devices are sensitive but, have strict optical coupling requirements and are difficult to fabricate. In optical fiber evanescent wave sensors the interaction with the surrounding environment is usually obtained by tapering an optical fiber, which significantly weakens the structure, or by just utilizing the end of the optical fiber. In this paper, in-line optical fiber structures are presented that are mechanically robust, and provide a large interaction length for high sensitivity. They are compatible with standard chemistries for optical affinity sensing of biological compounds.

Keywords: optical fiber sensor; surface plasmon, localized surface plasmon, biomolecules

* Corresponding author

1. Introduction

Conventional surface plasmon sensing uses a prism that is coated with a thin layer of gold.¹⁻⁴ Generally, the Kretschmann configuration, shown in Figure 1 A, is employed to couple the light incident on a prism to a thin gold film on its surface. This enables excitation of surface plasmons on the surface of the gold film. As the angle of incident light is changed, for a fixed wavelength of incident light, a dip in the reflection spectrum appears at an angle when the condition for surface plasmon resonance (SPR) is satisfied. When the refractive index of the medium above the gold film is varied, the position of the dip changes as the condition for surface plasmon resonance is satisfied at a different angle of incidence for the new refractive index of the medium above the gold film. The prism configuration and an example calculation for a change in refractive index are shown in Figure 1A, and 1B respectively. The surface plasmon in a thin film can also propagate in the plane of the film for several microns depending on the conductivity of the material. In the thin film only the TM mode at a specific angle can excite the plasmons. Moreover, surface plasmons can't be excited by propagating light and need evanescent waves to excite them as the propagation constant of propagating light is lower than that required for surface plasmon excitation.

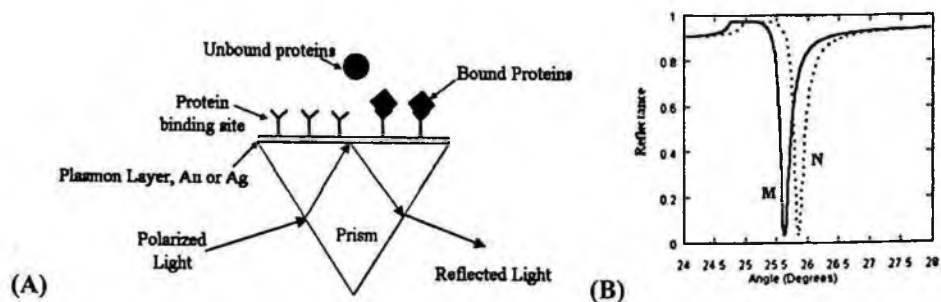


Fig. 1. (A) Prism configuration for detecting refractive index change due to protein binding and (B) Shift in plasmon resonance angle at a given wavelength is employed to sense presence or concentration of chemical and biological molecules; Refractive index of the medium above the gold film changes when biomolecules, e.g. proteins, bind to the binding sites (curve N) as compared to the case with no binding.

In a small conductive particle, the plasmon is localized and one can either describe the situation by Mie scattering or by considering the resonance as an oscillating dipole⁵⁻⁹. The magnitude and of the scattering is strongly dependent on the nanoparticles shape, size and dielectric constant of the surrounding medium. In a particle, while there is polarization dependence based on shape of the particle, the particle can be excited at normal incidence. Moreover, plasmon resonance in the case of metallic nanoparticles can be excited by propagating light. Unlike the thin film case, in a particle the plasmon resonance is localized. Resonant modes associated with the localized surface plasmons (LSPs) can be excited in the nanoparticles when the size of nanoparticles is smaller than wavelength of the incident light. This leads to a significant enhancement in absorbed and scattered light and enhancement of electromagnetic fields inside and near the particles.

LSPs are detected as resonance peaks in the absorption or scattering spectra or as dips in the transmission spectra of the metallic nanoparticles. Nanoparticles of very conductive metals like gold, silver, and copper are ideal materials for excitation of localized surface plasmons due to an extremely high ratio of the modulus of the real (ϵ_r) to the imaginary parts (ϵ_i) of its dielectric constant. Silver and copper nanoparticles are prone to oxidation and therefore often require coatings of protective over layers. Gold nanoparticles are chemically stable and are employed for the development of devices based on plasmon resonances of nanoparticles.

Fiber optic chemical and biological sensors, based on SPR in continuous metallic films deposited on optical fibers, have been previously reported.^{3,4} Sensors based on optical fibers have advantages like high response times, ability to be remotely interrogated, and low electromagnetic interference. In our study, fiber optic devices were developed based on excitation of LSPs in gold nanoparticles, deposited on the surface and tip of different optical fiber structures. As refractive index of the medium around the sensors was varied, or biological molecules attached to the nanoparticles, a shift in the plasmon resonance related dip in the transmission spectrum was observed.

2. Optical fiber structures for sensing applications

In this work, the main objective was to develop surface affinity sensors that employ plasmon resonances of metallic nanoparticles, as a means of transducing the input optical signal. Some of the polarization restrictions inherent in using thin films can be avoided by using localized surface plasmon. Instead of scanning the angle of incidence, one can measure the spectrum and find the light scattering or absorption resonance as a function of wavelength. Different structures, that were developed to form sensors for chemical and biomolecules, are shown in Figures 2 and 3.

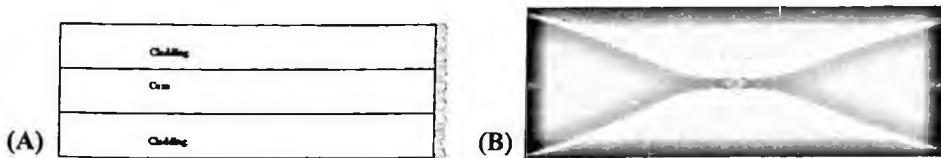


Fig. 2. (A) Deposition or formation of gold nanoparticles on the fiber tip, and (B) Tapered optical fiber employed for evanescent and surface plasmon sensing.

In Figure 2 A, use of end-face of the optical fiber as the sensing region, is shown. This is a simple method for LSP based sensing and geometry of the fiber tip can also be altered. However, the interaction volume is very small. Moreover, it requires the fiber to be interrogated in the reflection mode. Gold nanoparticles were deposited on the tip of the optical fiber as shown in Figure 2 A and light propagating in the optical fiber excited LSP in the gold nanoparticles. In order to evaluate the tip based fiber optic sensors in transmission mode, a platform to hold the sensor fiber and a collector fiber very close to each other, was developed. A V-groove having a width almost equal to the fiber diameter

was etched out in a silicon substrate by using standard photolithography and etching techniques. The sensor and collector fibers were placed in the V-groove, their tips at a certain distance $\sim 10 \mu\text{m}$ from each other, and held in that position using two magnets.

Another structure that was developed was a fiber taper, which is shown in Figure 2 B. Tapering the fiber allows the guided wave to escape the core of the fiber due to the fiber geometry and interact with the environment. This guided wave, after interaction with the environment, gets coupled back into the other end of the optical fiber. This allows in-line sensing of the medium around the fiber taper. As the light that escapes the taper is evanescent in nature, it can excite surface plasmons in a thin gold film on the tapered region and also in gold nanoparticles in that region. The fiber taper provides a larger interaction volume, but it significantly reduces the mechanical strength of the fiber since the diameter is typically reduced to less than $20 \mu\text{m}$. Hence employing this fiber sensor for large-area distributed sensing in tough environments is not possible.

In order to form chemical sensors that are sensitive yet robust, an in-line sensor that maintains the constant dimension of the optical fiber, but manipulates the light within the fiber to interact with the fiber/environment interface was designed¹⁰. These novel fiber structures in Figure 3 thus form an in-line sensing mechanism that is able to overcome the problem associated with fragile tapered fiber sensors.

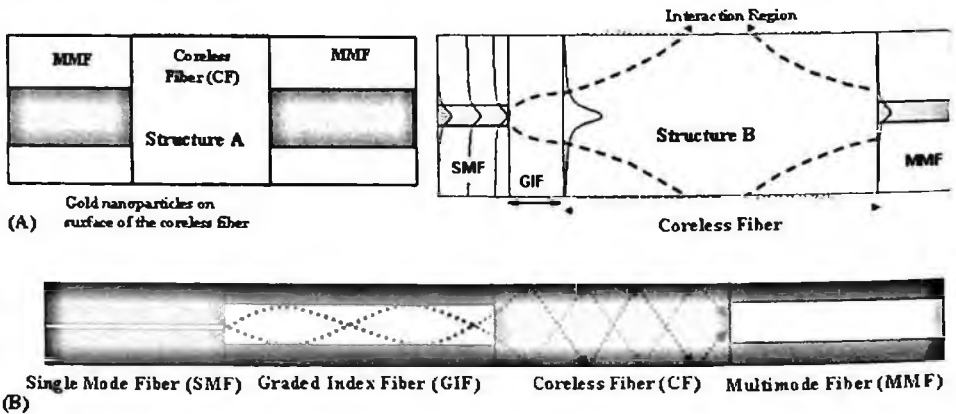


Fig. 3. Gold nanoparticles are formed on the surface of (A) A fused MMF-CF-MMF structure and a fused SMF-GIF- CF-MMF structure, (B) Ray Tracing schematic of Structure B, which is an in-line fiber structure consisting of single mode, graded index, coreless, and multimode fibers to expand the Gaussian beam to the surface of the coreless fiber.

Structure A is the simplest structure formed by fusion splicing a section of a coreless fiber to two pieces of continuous multimode fibers. In this sensor, different modes propagating in the multimode fiber expand to the surface of the coreless fiber in the central region and after reflection from the coreless fiber-environment boundary, get coupled back into the second multimode fiber fused to the coreless fiber. The evanescent waves at the coreless fiber-environment boundary excite plasmon resonance is the metallic nanoparticles formed on the surface of the coreless fiber section, as shown in Figure 3 A. This structure is not very effective for sensing purposes as there are a lot of

modes in a multimode fiber and this leads to a broadening of the plasmon resonance related dip in the transmission spectrum. A substantially more efficient structure was constructed by using multiple optical elements¹⁰ as shown in Figure 3 B. The optical elements that can be fused together to form a continuous fiber sensor, include:

- **Single Mode Fiber:** This fiber transmits light to the sensor, and also provides a well defined aperture for a Gaussian wave to expand from.
- **Multimode Fiber:** It can also be used transmit light to the structure, but resulting beam within the sensor is multi-moded and less controllable. However, it can be used efficiently as a collection fiber to transmit the signal to the detector, as shown in Figure 3 B.
- **Fiber Lenses:** Made of graded index optical fiber, by altering the length the graded index segment can be used as either a converging or diverging lens, and fused in-line with the sensor structure. Using combinations of lenses, a near collimated beam can be used to interrogate the nanoparticles placed on the surface. The nearly collimated beam can also be focused into the collection fiber.
- **Expansion Elements:** Made of coreless optical fiber, these serve a dual purpose. They can be used to expand the beam, but they also provide the coreless fiber/environment interface.

The mechanism of expanding the Gaussian beam propagating in a single mode fiber to the surface of the coreless fiber is illustrated by Figure 3 B. This figure also shows that the graded index fiber, if cut at the correct length can serve the purpose of either a converging lens or a diverging lens. In the case of structure B, a diverging lens was formed by cleaving the fiber at the appropriate length, as shown in Figure 3 B. The Gaussian beam expands in the coreless region till it reaches the coreless fiber/environment boundary. At the interface, the Gaussian beam gets reflected or folded multiple times due to total internal reflection. The reflected beams superpose to form the resultant beam that gains a near Gaussian profile after a certain distance of beam propagation. This Gaussian beam is then effectively collected by the multimode fiber due to its large numerical aperture. One could also employ a combination of a graded index fiber, cleaved at the correct length to form a converging lens, and a single mode fiber instead of the multimode fiber. The converging lens would allow light reflected back from the coreless fiber/environment interface to be effectively coupled back into the collector single mode fiber.

3. Experimental

The sensor shown in Figure 2 A was developed by stripping and cleaving F-MLD multimode fibers obtained from Newport Corporation with a 100 μm core and a 140 μm cladding diameter and then evaporating a very thin gold film (4-8 nm) on the fiber tips^{11, 12}. Electron-beam evaporation process was employed for carrying out the

gold film depositions. An electron beam, with a current varying between 65 and 85 mA, was directed towards a crucible containing gold slugs, and thickness of the film deposited on the fibers was measured by a quartz crystal monitor. Optical fibers were attached to the electron beam substrate holder by employing a fiber holder made from a thick aluminum foil. The evaporation of gold was carried out at a chamber pressure of 3×10^{-6} Torr. In order to form nanoparticles, the thin gold films deposited on fiber tips were annealed at different temperatures by employing a micro-coil heater. Upon annealing, the transmission spectrum showed a plasmon resonance related dip, indicating the development of gold nanoparticles on the tip of the fibers. The annealing of the gold films was also carried out using controlled, low-intensity plasma arcs. A Sumitomo Electric fusion splicer was used to plasma arc-anneal the gold-coated fiber tips and a schematic of the process is shown in Figure 4 A. It was observed that the plasmon resonance related dip in transmission spectra that appeared upon application of 1 plasma arc became sharper and moved to lower wavelengths when 10 plasma arcs were applied, as shown in Figure 4 B. This could be attributed to change in size and shape of the gold nanoparticles upon further plasma annealing of the nanoparticles. The fiber taper structure shown in Figure 2 B was developed by using the Sumitomo fusion splicer. A commercial software was employed to precisely control the rate of pull of the optical fiber as a plasma arc was applied to develop the fiber taper. The sensor structure described in Figure 2 A was evaluated for its sensing capability by inserting the fiber holder, with the sensor and collector fibers into several media having refractive indices higher than that of air such as water, methanol, and ethanol. Optical transmission, of white light through the fiber sensors, was measured by using a SpectraPro®-500 spectrograph having a Spectradrive® stepping motor scan controller.

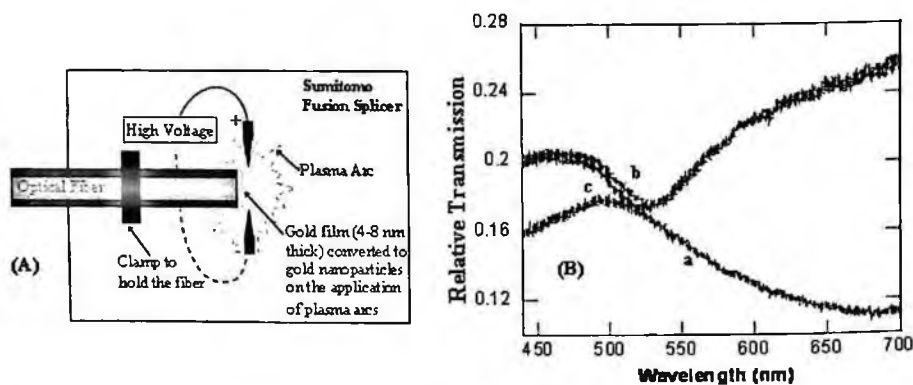


Figure 4. (A) Plasma arc annealing of 4-10 nm gold film on the optical fiber tip using a fusion splicer (B) Transmission spectra before plasma arc application (a), after 1 plasma arc (b) & after 10 plasma arcs (c).

In order to form the in-line structure A, a multimode fiber was fusion spliced to a cleaved end of a coreless fiber. The coreless fiber was then cleaved after a length of 35.3 mm using a diamond scribe and pull assembly developed by us. The cleaved end of the coreless fiber was fusion spliced to another multimode fiber, which served as the collector fiber. The length of the coreless fiber was determined by employing calculations that are described previously¹⁰. The composite fiber structure B, was developed by first fusing GIF1 to SMF1 using a fusion splicer. The first graded index fiber, i.e. GIF1, was cleaved at approximately 1622 μm length by using a microscope setup and a diamond scribe. Then a coreless fiber was fused and cleaved at an appropriate length. This length depended on the desired level of interaction of the propagating light with the environment and varied between 17.6 mm and 35.3 mm. A multimode fiber was then fused to the SMF1-GIF1-CLF arrangement. 8 nm gold films were then deposited on the surface of the coreless fiber segments of structures A and B and annealed at 330 °C for 10 minutes forming nanoscale gold islands and nanoparticles on the surface of the coreless fiber. The sensor structures A and B were evaluated for their sensing capability by inserting the in-line fiber sensors into several mediums having refractive indices higher than that of air such as water, methanol, and ethanol. Optical transmission of white light through the in-line fiber sensors was measured by using the spectrometer described above.

This work was also extended to examine the sensors to detect biological molecules. The Biotin/Streptavidin system was chosen for initial studies since it has a large free energy of association, and is stable under reasonable temperatures and pH ranges. Since it is also used as a test system, it permits sensing results to be compared with other sensor systems. The Biotin molecule works well with gold via thiol bonding and the Streptavidin molecule can also be designed to be selectively receptive to a variety of other molecules, toxins or proteins. Monitoring the binding response over time also allows reaction kinetics, and concentrations to be determined. In order to carry out measurements for sensing biomolecules, Biotin-HPDP ((N-(6-(Biotinamido)hexyl)-3'-(2'-pyridyldithio)-propionamide) was obtained from Pierce Biotechnology. It was attached to gold nanoparticles on the surface of the in-line sensor fibers and the fiber tips by first reducing the Biotin-HPDP using a reducing agent, Dithiothreitol, to form free thiol groups that are available for attachment to the gold nanoparticles. Streptavidin, obtained from the bacteria *Streptomyces Avidinii*, was obtained from Sigma Aldrich. In order to form the Biotin-HPDP solution, 1 mg of Biotin-HPDP was dissolved in 100 μl of DMSO to get a concentration of 10 mg/ml. A solution of Dithiothreitol was separately prepared and 100 μl of that solution was added to the Biotin-HPDP solution to get a net Dithiothreitol concentration of 50 mM. After around 4 minutes the disulfide bond in the Biotin-HPDP is reduced by the reducing agent and the solution turns yellow in color. Fiber tip sensors and the in-line fiber optic sensors were immersed in a 2 ml mixture of water and methanol (1:1) in a small boat-shaped fluid cell. The reduced Biotin-HPDP solution was then added to the water-methanol mixture. Right after the reduced Biotin solution was added to the solution containing the sensor, optical transmission as a function of time, of white light through the fiber-tip and in-line fiber sensors, was measured by using a

spectrometer. Biotin-HPDP was allowed to react with the gold nanoparticles for up to 2 hours and a transmission spectrum through the fiber sensors was taken every minute. After binding the thiol group of the reduced Biotin-HPDP with the gold nanoparticles, a wash buffer (a dilute PBS buffer solution with a pH of 7.2) was flowed into the fluid cell to wash away the Biotin-HPDP that is not bonded to the gold nanoparticles. Then, a solution of Streptavidin was made by dissolving 0.15 mg of Streptavidin in 150 μ l of 7.2 pH PBS buffer solution. This solution was then added to the fluid cell with the fiber sensors so as to surround the fiber sensors completely. Right after the Streptavidin solution was added to the fluid cell holding the fiber sensor, optical transmission of white light through the fiber-tip and in-line fiber sensors was measured by using a spectrometer. Streptavidin molecules were allowed to react with the Biotin molecules that were in turn bound to the gold nanoparticles via the thiol bond, for up to 4 hours and a transmission spectrum through the fiber sensors was taken every 5 minutes.

As a means of incorporating the fiber sensors into systems, incorporation of in-line fiber sensors into non-woven fabrics was explored. In order to incorporate an in-line sensor into a textile fabric, a spun-bonded base fabric was taken and an in-line fiber sensor was placed on top of it. Subsequently, non-woven fibers of polyurethane were electrospun on top of the optical fiber sensor to hold the sensor fiber in its position, thereby incorporating it in a non-woven fabric matrix. The electro-spinning was carried out at a voltage of 18 KV using a solution of polyurethane in the solvent DMP.

4. Results and Discussion

The spectral response of fiber tip based sensors are shown in Figure 5. It was observed that in air, the position of the plasmon resonance related dip was \sim 526 nm. When the fiber sensor was placed in methanol, having a refractive index of 1.33, the plasmon resonance related dip moved to 541 nm and when it was placed in acetone, having a refractive index of 1.36, the plasmon resonance related dip moved to 545 nm.

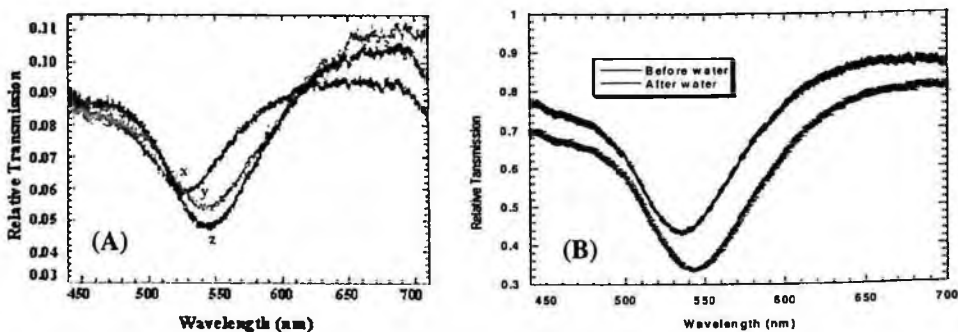


Figure 5. (A) Shift in transmission spectrum when a fiber tip containing gold nanoparticles, formed by thermal annealing, is placed in air with refractive index \sim 1 (x), in methanol with refractive index \sim 1.33 (y), and acetone with refractive index \sim 1.36 (z), and (B) Shift in transmission spectrum when a fiber tip containing gold nanoparticles is placed in water.

Optical transmission spectrum of the two nanoparticle coated in-line fiber structures A and B showed a plasmon resonance related dip indicating the formation of nanoparticles on the fiber surface. When the sensor fibers were dipped in water (a change of refractive index of ~ 0.33 from air) and media of a refractive index higher than air, there was a shift in the spectral position of the plasmon resonance related dip towards higher wavelengths. The spectral response of a sensor formed by using structure B is shown in Figure 6. When the structure B based sensor was inserted in water, a peak shift of 42 nanometers per refractive index unit was obtained. In the case of structure A, a peak shift between 80-120 nanometers per refractive index unit was obtained for different samples that were prepared under similar conditions. The shift in the resonance related dip for structure A was higher than that for structure B but the plasmon resonance related dip was very broad for structure A. Hence, the peak position could not be as accurately determined for the structure A sensor as compared with the structure B sensor.

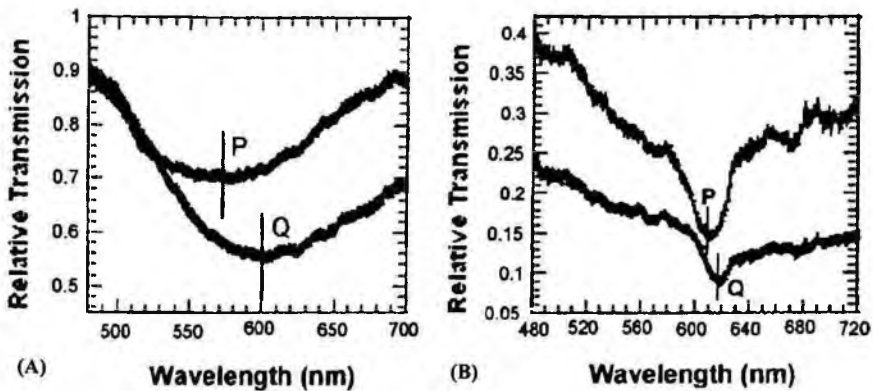


Figure 6. Spectral response of in-line fiber structures having gold nanoparticles on their surface showing the shift in plasmon resonance related dip when placed in water for (A) Structure A & (B) Structure B. Lines labeled P are in air and lines labeled Q are in water. Figure reprinted from Optics Letters¹⁰.

The fiber tip sensor, employed in the transmission mode, and the in-line fiber optic structures were also employed for affinity sensing of biomolecules such as Biotin and Streptavidin. The chemical bond between Biotin and Streptavidin is extremely strong with a large free energy of association¹³. When the reducing agent Dithiothreitol breaks the disulfide bond in Biotin-HPDP, the free thiol (-SH) group of Biotin-HPDP attaches to the gold molecule. As Biotin has a high affinity for Streptavidin, one can chemically conjugate different biological molecules such as antibodies to Streptavidin and then react the Streptavidin with the Biotin. An example of such an antibody that can be chemically conjugated to Streptavidin is Immunoglobulin IgG. Thus an affinity sensor, which can specifically sense the presence of antigens such as anti-IgG, can be developed. The Streptavidin molecule can be modified, via conjugation to other biomolecules, to make it selectively receptive to a variety of other molecules, toxins and proteins. Thus Biotin-Streptavidin based optical sensing systems can not only sense the concentration of

biological agents and contaminants, they can also study the chemical kinetics of the different reactions taking place in the environment being sensed. Shift in plasmon resonance related dip wavelength, upon the addition of Biotin to a fluid cell containing the in-line sensor based on structure A, is shown in Figure 7 A. In this experiment, the spectral readings were taken every minute after the addition of Biotin to the fluid cell. It can be seen that there was no shift in the plasmon resonance related dip position after 4 minutes. In another experiment, Biotin was first attached to a fiber sensor and subsequently, the response of Streptavidin addition to the Biotin coated fiber was evaluated. Value of the wavelength, at which the plasmon resonance-related dip occurs, was plotted vs. time for the in-line fiber optic biosensor and is shown in Figure 7 B. Development of biosensors on in-line optical fiber structures can enable continuous online monitoring of biological agents present in the environment around the sensors. These fiber sensors potentially can be conveniently inserted into water filtration systems by incorporating them into the filtration system to sense the presence of harmful agents or contaminants.

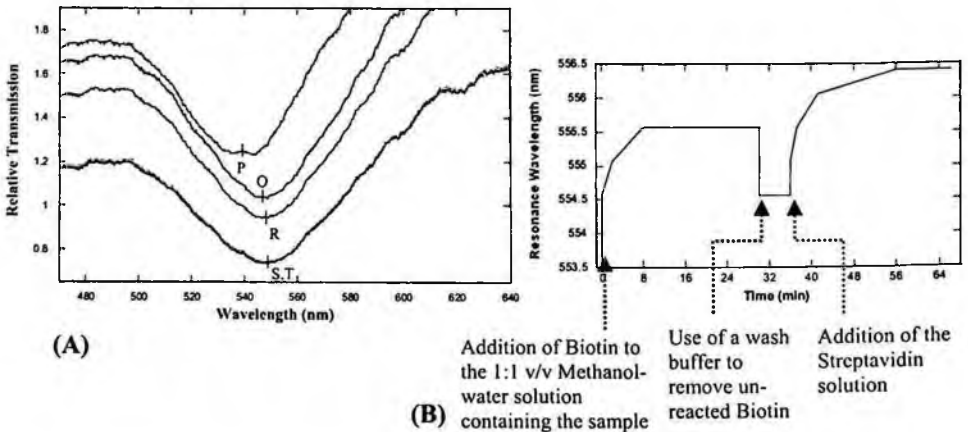


Figure 7. (A) Shift in plasmon resonance peak wavelength upon addition of Biotin; Peak wavelength before Biotin addition was at 539.48 nm (P), right after Biotin addition was at 546.36 nm (Q), 1 minute after Biotin addition was at 547.85 nm (R), 2 minutes after Biotin addition was at 548.8 nm (S), & 3 minutes after Biotin addition was at 548.88 nm (T), and (B) Response of Streptavidin addition to a Biotin coated fiber; The value of the wavelength at which plasmon resonance-related dip occurs was plotted vs. time for an in-line fiber optic biosensor based on structure A.

5. Conclusions

Sensors and other devices were developed on optical fibers by incorporating metallic and semiconducting nanoparticles on the surface of optical fibers. Several ways of forming sensitive and robust chemical sensors based on plasmon resonances of metallic islands and nanoparticles were demonstrated. Development of in-line optical fiber structures, involving single mode or multimode optical fibers fused to an arrangement of coreless

and graded index fibers, enabled us to expand the light propagating in the core of the optical fiber to reach the surface of a coreless fiber and effectively interact with nanoparticles on the fiber surface and the environment. Moreover, these optical fiber sensors and devices can be embedded into non-woven or woven textiles for large area distributed sensing. They can be also be inserted into water filtration systems to sense the presence of harmful agents or contaminants. These sensors were evaluated by monitoring their optical response to binding of Biotin/Streptavidin to the sensor surface and by placing the sensors into media of different refractive indices.

References

1. H. J. Simon, D. E. Mitchell, and J. G. Watson, Surface plasmons in silver films - a novel under-graduate experiment, *Am. J. Phys.*, **43** 630 (1975).
2. R. J. Green, R. A. Frazier, K. M. Shakesheff, M. C. Davies, C. J. Roberts and S. J. B. Tendler, Surface plasmon resonance analysis of dynamic biological interactions with biomaterials, *Biomaterials* **21**, 1823-1835 (2000).
3. J. Homola, S. S. Yee and G. Gauglitz, Surface plasmon resonance sensors: review, *Sens. Actuators B* **54**, 3 (1999).
4. J. Homola, On the sensitivity of surface plasmon resonance sensors with spectral interrogation, *Sens. Actuators B* **41**, 207 (1997).
5. A. Stratton, *Electromagnetic Theory*, New York: McGraw-Hill, 1941.
6. H. C. Van de Hulst Ed., *Light scattering by small particles*, John Wiley and Sons: New York 1957.
7. M. Born and E. Wolf, *Principles of Optics 7th ed.*, Cambridge University Press: New York 1999.
8. T. Okamoto "Near-Field Spectral Analysis of Metallic Beads" In *Near-Field Optics and Surface Plasmon Polaritons*, S. Kawata Ed., Springer-Verlag Berlin: Heidelberg 2001.
9. C. F. Bohren and D. R. Huffman Eds., *Absorption and Scattering of Light by Small Particles*, Wiley: New York 1983.
10. A. Dhawan and J. F. Muth, In-line fiber optic structures for environmental sensing applications, *Opt. Lett.* **31**, 1391-1393 (2006).
11. A. Dhawan and J. F. Muth, Optical nano-textile sensors based on the incorporation of semiconducting and metallic nanoparticles into optical fibers, Proceedings of the MRS Conference, Spring 2006.
12. A. Dhawan and J. F. Muth, Plasmon resonances of gold nanoparticles incorporated inside an optical fibre matrix, *Nanotechnology* **17**, 2504-2511 (2006).
13. Y. Mu, H. Zhang, X. Zhao, D. Song, Z. Wang, J. Sun, M. Li, and Qinhan Jin, An Optical Biosensor for Monitoring Antigen Recognition Based on Surface Plasmon Resonance Using Avidin-Biotin System, *Sensors*, **1**, 91-101 (2001)

PROGRESS IN HIGH EFFICIENCY UV LED RESEARCH FOR REAGENTLESS BIOAGENT DETECTION AND WATER PURIFICATION

MEREDITH L. REED, GREGORY A. GARRETT, ANAND V. SAMPATH, PAUL H. SHEN,
CHARLES J. COLLINS, MICHAEL WRABACK

*U.S. Army Research Laboratory, Sensors and Electron Devices Directorate,
Adelphi, MD
Meredith.Ford@arl.army.mil*

JIANPING ZHANG, XUHONG HU, JIANYU DENG, ALEX, LUNEV, YURIY BILENKO,
THOMAS KATONA, REMIS GASKA

*Sensor Electronic Technology, Inc., 1195 Atlas Road
Columbia, SC*

We present material and device characterization of 280 nm semiconductor ultraviolet light emitting diodes. These devices exhibit low series resistance, wavelength stability with increasing current, and have a half-life in excess of 570hrs, depending upon the injection current. Time-resolved photoluminescence studies of these materials prior to fabrication have been correlated with the device performance. We also discuss the potential for use in water purification.

Keywords: UV; LED; water purification.

1. Introduction

Several emerging applications for ultraviolet (UV) light emitting diodes (LEDs) include spectroscopy, chemical and biological analysis, water and food treatment, solid-state lighting, and short-range ultraviolet non-line of site communications. High power UV LEDs based on III-Nitrides are preferred for these applications because they are small, power efficient, contain no hazardous materials, and are highly resistant to shock. This is in contrast to current technology where low-pressure mercury vapor lamps are utilized. These mercury lamps have several disadvantages in that they require high voltages for operation, are toxic, and are limited to a peak emission wavelength of ~254nm.

Water sterilization by ultraviolet light is relatively straightforward and clean. Purification occurs when microbial agents present in water are exposed to UV light within the germicidal region, ~245-295nm as shown in figure 1. The most effective peak wavelength for killing or destroying the DNA of bacteria is at 265nm, which is the main driving force for replacing mercury lamps with UV LEDs that are designed to emit

specifically at that wavelength.¹ Recently, Sensor Electronic Technology Inc. (SET) has produced AlGaIn-based UV LED germicidal lamps emitting in the range from 260 – 280nm for water purification.^{2,3} Results of SET UV LED germicidal lamps emitting at 280nm reduce the level of e-coli spiked water flowing at 150ml/min and 300ml/min by 99.99% and 99.0%, respectively, where typical rates for an individual water treatment system requires 500ml/min.⁴ Although this demonstration is encouraging, LED optical output power must be increased for adoption in many water purification applications.

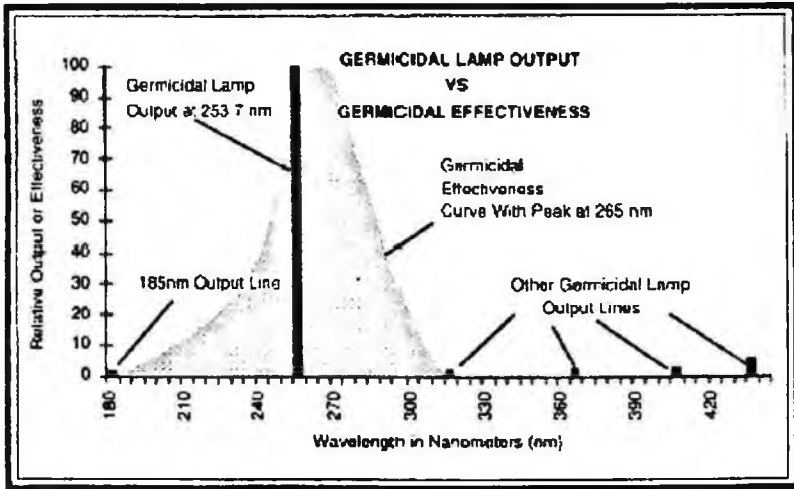


Fig. 1. UV spectrum showing germicidal lamp output versus the effectiveness of germicidal destruction at that wavelength.⁵

The shortest wavelength attainable for AlGaIn-based UV LEDs is ~200nm. However, several technical barriers prevent the realization of efficient LEDs at these short wavelengths. Currently UV LEDs are plagued by low output powers because as the bandgap of AlGaIn alloys increase these films become more insulating, suffer from lower internal quantum efficiency, and are subject to problems such as current crowding and self-heating. Despite these constraints, UV LEDs with wavelengths as short as 247nm are commercially available.⁶ In this paper, we evaluate the performance of 280 nm UV LEDs, with a view toward potential improvements that may lead to significantly better future performance. We also discuss the potential of nanoscale compositional inhomogeneities (NCI) to improve the quantum efficiency and overall performance of UV LEDs.

2. Experimental

Devices were grown, fabricated and packaged at Sensor Electronic Technology (SET), Inc. Figure 2 shows a schematic of a typical SET device structure grown by MOCVD.⁷ Electrical and optical testing was performed at ARL. Electroluminescence, lifetime measurements, L-I, I-V and spectral response measurements were taken using an Ocean Optics SD2000 spectrometer and integrating sphere with UV rated fiber optic cable in conjunction with a Keithley 2420 source meter. Time-resolved photoluminescence (TRPL) measurements were performed prior to device fabrication.

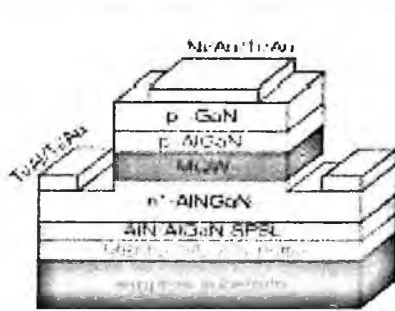


Figure 2. Schematic of typical UV LED structure.

3. Optical Testing

Time-resolved photoluminescence measurements (TRPL) performed by ARL utilize a femtosecond excitation pulse tuned to pass through the transparent substrate and n-type current injection layers to excite carriers directly in the MQW active region of the diode structure. The TRPL results can be correlated with the LED output powers and external quantum efficiencies directly. This type of characterization enables one to chart and understand materials growth improvement prior to full device fabrication, while allowing one to separate active region quality from that of current injection layers. Figure 3 shows TRPL data from the active region of a 280 nm device structure prior to fabrication. The active region of the 280 nm LED structure exhibits a complex TRPL decay that is approximated by a dominant 927 ps decay and a weaker 347 ps component. In addition, the decay time is almost independent of pump intensity for all measured values until the highest pump intensity (smallest neutral density filter) for which the decay becomes noticeably faster. This decrease in photoluminescence decay time signifies the onset of nonlinear radiative recombination, characteristic of higher quality samples with less defects contributing to nonradiative recombination. 280 nm LEDs fabricated from these wafer batches exhibited ~0.6 mW of output power at 20 mA CW.

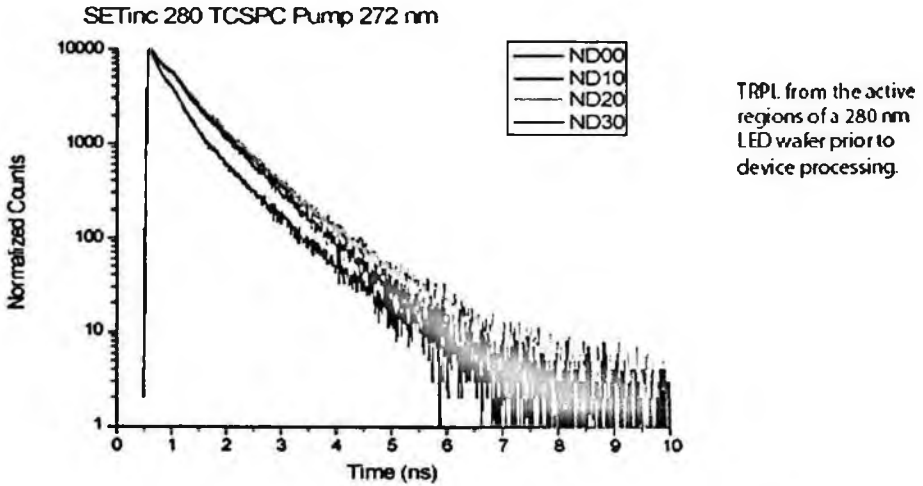


Figure 3. TRPL from the active regions of a 280 nm LED wafer prior to device processing.

4. Electro-Optical Testing

Figures 4 and 5 show typical I-V curves and electroluminescence spectra collected by ARL for 280nm UV LEDs. Analysis of the I-V curve for the 280nm UV LED gives a series resistance of only $\sim 9\Omega$, indicating excellent electrical properties of the n- and p-type layers of this structure. An optical power output of 0.6 mW was obtained at 20mA and an operating voltage of 5.39V. This gives a wall-plug efficiency of 0.56% for this

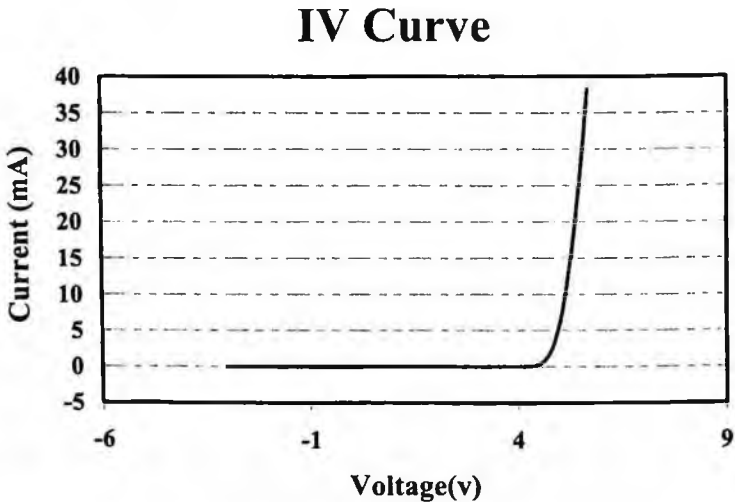


Figure 4. Typical IV characteristics of UV LED emitting at 280 nm.

device. CW electroluminescence for the same 280 nm device shows no wavelength drift over the range of injection currents tested, as shown in figure 5. This wavelength stability indicates that low thermal resistance packaging has largely mitigated the deleterious effects of device self-heating. The inset shows a log plot of the same spectra showing deep-level intensities only 1/1000 of the primary emission peak at 280 nm.

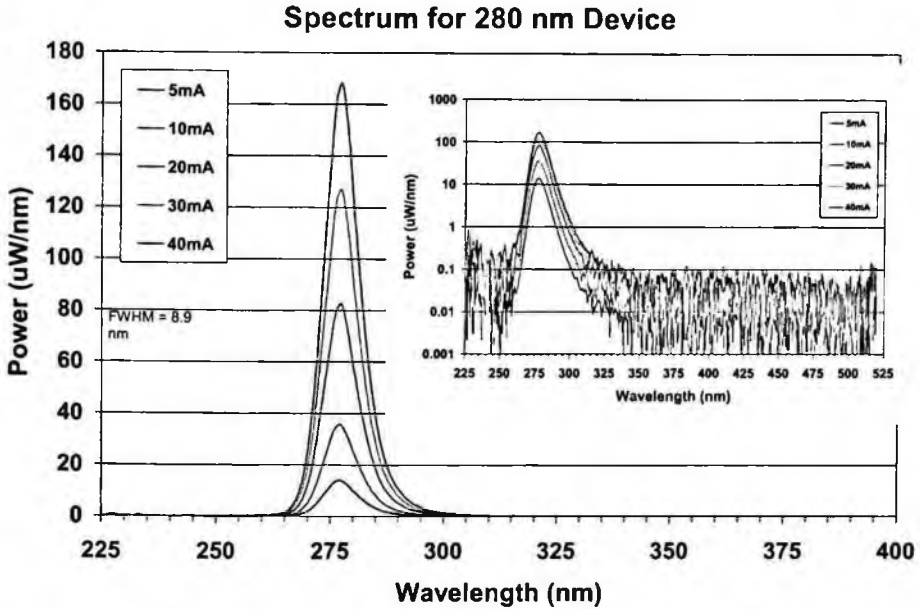


Figure 5. Electroluminescence spectrum of a typical 280 nm device at different injection currents. Inset shows log plot of the electroluminescence spectrum.

Lifetime tests on a typical 280 nm LED under continuous current injection are shown in figure 6. For a 25 mA drive current, output power is initially greater than 1 mW, but rapidly drops to 0.78 mW during the ~ 20 hr burn-in period. The LED power then follows a more linear decay with time, with a half-life of ~ 250 hrs. When the device current was increased to 50 mA to recover an optical power of ~ 1.2 mW, a second burn-in took place that left the device on a longer decay slope, with a half-life of ~ 570 hrs. These results suggest two degradation mechanisms. During burn-in, lower quality portions of the device cease to function leaving behind more stable regions of the device that decay more slowly. Under the increased electrical stress at 50mA a similar behavior is observed. Examination of digital images of the operating chip support the conclusion of partial junction area burn out by showing the appearance of dark regions with time. The origin of these localized areas of device burn-out are currently under investigation.

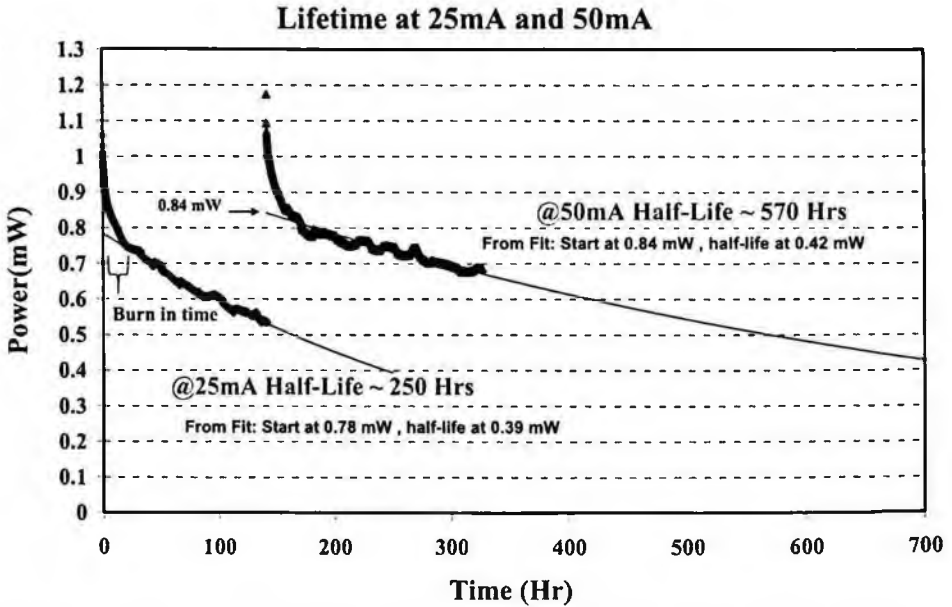


Figure 6. 280 nm LED lifetime testing. The initial curve is for 25 mA drive current. The second curve is for 50 mA drive current.

5. Future Directions: NCI AlGa_N active regions

ARL has developed a very high internal quantum efficiency (>30%) MBE-grown active region based on nanoscale compositional inhomogeneities (NCI) in AlGa_N.⁸ Incorporation of this highly efficient optically active layer can lead to leap-ahead device performance for UV LEDs. ARL has already demonstrated mW class devices incorporating this active region. Figure 7 shows that the pump intensity dependent TRPL lifetimes in these active regions are comparable to those in the LED sample, with a dominant lifetime > 700 ps at low intensity and a fast initial decay at the highest intensity. ARL will investigate growth of these novel NCI active regions on optimized n-type base layers, and evaluate their applicability for future high efficiency LEDs.

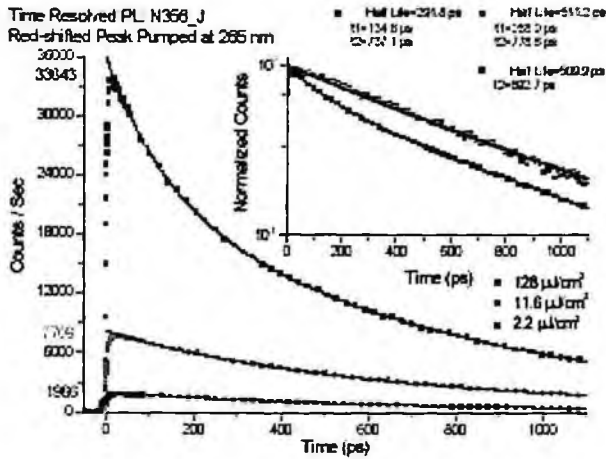


Figure 7. TRPL from an ARL high internal quantum efficiency NCI AlGaIn LED active region. Faster initial decay at highest intensity provides evidence of density dependent decrease of radiative lifetime.

6. Conclusions

Considerable improvements have been achieved in the field of UV LEDs by reaching shorter wavelength emission while at the same time increasing power output. However, further improvements in power output and device lifetime are necessary to enable widespread market acceptance in a myriad of potential applications. Materials and device testing and characterization will foster further improvements in device performance. Additionally, materials development at ARL has resulted in the creation of a novel NCI UV emitting layer that exhibits enhanced internal quantum efficiency in excess of 30%. Combination of these NCI active layers with high quality, low defect density current injection layers like those employed in SET 280 nm devices may lead to substantial improvements in deep-UV light emitting diodes.

References

1. Y. Bilenko, et al., Jap. J. Appl. Phys. 44, L98 (2005).
2. *SET launches UV LED germicidal lamp*, Compoundsemiconductor.net, February 21, 2006.
3. R. Gaska and J. Zhang, Proc. SPIE 6037, 603706 (2005).
4. *Brighter LEDs improve water purification rates*, Compoundsemiconductor.net, November 14, 2005.
5. Lupal, Myron, *UV Offers Reliable Disinfection*, http://www.cwqa.com/html/UV_Article.htm
6. <http://www.s-et.com/>
7. R. S. Qhalid Fareed, J. P. Zhang, R. Gaska, G. Tamulaitis, J. Mickevicius, R. Aleksiejunas, M. S. Shur, M. A. Khan, *Migration enhanced MOCVD (MEMOCVD™) buffers for increased carrier lifetime in GaN and AlGaIn epilayers on sapphire and SiC substrate*, physica status solidi (c) Volume 2, Issue 7, 2095-2098 (2005).
8. G. A. Garrett, A.V. Sampath, C. J. Collins, E. D. Readinger, W. L. Sarney, H. Shen, M. Wraback, V. Soukhovcev, A. Usikov, V. Dmitriev, Physica Status Solidi (c) Volume 3, Issue 6, 2125-2128 (2006).

TOWARDS SENSING SINGLE OR A FEW BIO-MOLECULAR ARCHITECTURES: DESIGN OF FUNCTIONAL SURFACES

PEIJI ZHAO

*Department of Electrical and Computer Engineering, North Carolina State University Raleigh, NC 27695, USA
pzhao@eos.ncsu.edu*

DWIGHT WOOLARD

*The US Army Research Office, Army Research Lab. RTP NC 27709-2211, USA
dwight.woolard@us.army.mil*

JORGE M. SEMINARIO

*Department of Chemical Engineering, Texas A&M University, College Station, TX 77843 USA
seminario@tamu.edu*

ROBERT TREW

*Department of Electrical and Computer Engineering, North Carolina State University Raleigh, NC 27695, USA
trew@ncsu.edu*

There is considerable interest in electrical sensing of biomolecular binding since it has the potential to be label free, to work easily in aqueous environments native to the biomolecules, and to be integrated with small, fast, and inexpensive microelectronics as detection instrumentation. Although electrochemical methods have been used successfully in detections of DNA molecules with Ag labels at very high sensitivity (\sim p ml), detection of DNA molecules in terms of label free techniques has a lower sensitivity (\sim μ ml). Here, the surface attachment chemistry is critical towards the detection of ultra-low concentration of biomolecules. In this article, based on density functional theory, we have calculated and analyzed the electrical characteristics of the contact between aromatic molecules and silicon (100) – 2×1 surfaces. Design principles for silicon based electrodes of electrochemically biomolecular sensing instruments for label-free sensing of single or a few biomolecular molecules have also been discussed.

Keywords: aromatic molecules, silicon surface, contact attachment, Schottky contact, Ohmic contact, electrical characteristics

1. Introduction

Biomolecular sensing systems have been developed dramatically in the last years. Two basic concepts, direct sequencing [1] and DNA hybridization [2], are used to detect and identify polynucleotides. The latter concept is more widely used, given its simplicity and

the inherent exceptional specificity of complementary DNA strands to hybridize [3]. By detecting hybridization of unknown target DNA molecules to known single-stranded oligo- or polynucleotides, DNA sequences are readily identified. Among the DNA molecule detection methods that are widely used in practices, the dominant approach to detect hybridization is based on fluorescent labeling of the target denatured DNA. Fluorescent labeling has given rise to extremely efficient diagnostic systems, especially since the emergence of gene chips and nanoscale particle labeling technique. However, fluorescent labeling technique has some intrinsic drawbacks: (1) Labeling not only constitutes an additional costly and time-consuming preparatory step, it can introduce modifications in DNA strand conformation that lower the precision of gene detection; (2). label degradation, labeling yield fluctuations, fluorescence efficiency site dependence and fluorophore quenching inhibit the quantifiability of genetic diagnostics, which would be very important for various analysis, especially in cancer research; (3). Labels may interfere with the antibody-antigen binding event. Therefore, a strong interest in developing alternative label-free detection schemes exists.

So far, there are two types of label-free DNA molecule detection techniques widely studied in laboratories. In the past decades, optical-based label-free detection of DNA molecules gains great development. The sensitivity of the technique has reached the order of pico-mole level. In contrast, although electrochemical method has obtained more and more attentions in the area of detection of DNA molecules because of the effectiveness and costless of the method, the sensitivity of DNA sensors with silicon electrodes is in the order of μ to nano-mole [4], which is low for the propose of detection of ultra-low concentration of DNA molecules, although sensitivity of electrochemistry based sensors has been greatly improved with the incorporation of microelectronic techniques.

Based on measuring the change of charges on DNA monolayer on the surface of electrode, electrochemical method can be used in monitoring DNA hybridization processes [5], [6]. This is or detection of bio-molecules in terms of electrochmetrical methods. Hence, enhancement of the sensitivity of the electrodes of the sensors is a key for the improvement of the sensitivity of the sensors. Here, the sensitivity of the electrodes of the sensors is not only associated with the surface attachment chemistry but also related to the electric characteristics of the contact between the probe molecule and the surface of the electrodes.

The absorption of organic molecules on semiconductor surface has been drawing attention in the past years for the possibility of combining the wide range of functionality of organic molecules with existing semiconductor-based infrastructures [7, 8]. It is also crucial to future molecular devices since the attachment of the molecules onto the surface of electrodes also play important roles in determining the functionality of the devices [9].

The attachment of π -conjugated molecular systems on the surface of substrate has become a research interests in recent a few years [10~12]. Theoretical and experimental

researches show that π -conjugated molecular systems can be absorbed onto the surface of silicon (100) surfaces via direct Si – C linkage [13], Si–S linkage [14], Si–N linkage [15]. Previous research of the attachment of π -conjugated system to the surface of silicon mainly focus on the binding geometrical structures and the stability of the structures. In this paper, we discuss the electrical behaviors of the systems. Specifically, we are interested in the answers of the following questions: 1). which is the most electrically effective binding structures? 2). how to realize the Ohm contact or the Schottky contact in aromatic molecule – silicon systems. The answers to these two questions are important not only to the hybridized bio-molecule-semiconductor architectures for biomolecular sensing, but also to the research of transport through molecular devices because a correct boundary condition to the model of molecular devices. Furthermore, the electrical characteristics of the interface between the surface of the substrate and the self-assembled molecular monolayer is crucial to the detection of ultra-low concentration of bio-molecules.

In this paper, in terms of density functional theory, we will study the attachment of π -conjugated system on silicon (100) – 2×1 surface. Since the fabrication of a π -conjugated molecule system on silicon surface is the first step in fabricating electric driven bio-molecular detection devices, we discuss the electrical characteristics of the attachment of aromatic molecules on silicon surfaces. In the second section, we discuss the theoretical method and computational method used in this paper. The numerical calculation results and the analyses of the results will be given in the third section of the paper. The conclusion is provided in the last section.

2. Theoretical and Computational Methods

Silicon (100) surfaces play a decisive role in many technologically important device applications. The structure of the surface has been widely studied, both experimentally and theoretically [16~18]. The knowledge of the surface provides a good reference for the study of the interactions of the surface and the complex molecules. Thus, in this research, we investigate the issue of attachment of aromatic molecules to the dimerized silicon (100) surface. This is an importance in device fabrication. In theoretical research, slab models and cluster models have been extensively used to describe Si(100) 2×1 surfaces [16~18]. Among these models, the cluster model has been extensively used in theoretical research of the surface since it is simple and amenable to model the fundamental features of the surface. The smallest cluster model for the 2×1 -reconstructed Si(100) surface is the Si_9H_{12} one-dimer cluster, which includes two surface Si atoms representing the dimer structure, four second layer silicon atoms, two third layer silicon atoms, and one fourth layer silicon atom. Although this cluster can be used to simulate the bulk feature of silicon, the one dimer cluster does not describe the surface reaction adequately [17]. In order to adequately describe the surface at low computational cost, we adopt the three dimer cluster model $Si_{21}H_{20}$. As the energy gain per dimer for the three-dimer cluster is between 0.15-0.20eV and in agreement with the slab results to

better than 0.05eV, it is possible to infer the characteristics of the Si(100)-2×1 surface from the three-dimer cluster $Si_{21}H_{20}$ [18]. In order to analyze the previously stated two questions, that is, 1). which is the most electrically effective binding structures? 2). how to realize the Ohm contact or the Schottky contact in aromatic molecule, we numerically calculated the Mulliken population of the system using PBELYP density functional with the 6-31G** basis set as coded in the program GAMESS. Geometry optimization has been performed on (100) – 2×1 surface model of silicon.

3. Calculation Results and Analyses

For π -conjugated molecular systems to be used as conductors or semiconductors, they must retain π -conjugation, even after absorption onto the surface of silicon. Among the π -conjugated systems, benzene molecule, the simplest π -conjugated system, can be absorbed onto the Si (100)- 2×1 surface by Si – C linkage. However, the formation of new Si-C bonds leads to the loss of π -conjugation of the system [19]. The broken of the π -conjugation greatly reduces the electron transport capabilities of the molecule which is the result of the localization of the electrons in the molecule. Thus, the key issue of attaching a π -conjugated system to the surface of silicon is to find suitable attaching mechanism to ensure the π -conjugation after the attachment of the molecules. Experimentally, this task can be finished using the following methods: 1). Using a sulfur atom as a tether for selective attachment of aromatic molecules to silicon (100) surface [14]; 2). Formation of a benzoimine-like conjugated structure through the absorption of benzonitrile on silicon (100) surface [20]; 3). Using Ge-covered silicon (100) surface as a tempting template and highly selective binding of styrene through cycloaddition reaction between vinyl group of styrene and an Ge = Ge dimer [21]; 4). as well as direct Si – C linkage between a polycyclic aromatic molecule and silicon (100) surface [22]. Among these techniques of binding aromatic molecules to silicon (100) surfaces, the charge distribution in the Si – C linkage and the Si – S linkage determine the electrical features of the contacts.

3.1 Electrical Characteristics of Si – S linkage

Considering the attachment of benzenethiol on silicon (100) - 2×1 surface. Fig. 1 illustrates the binding structures of benzenethiol and silicon (100) - 2×1 surface [14]. It also shows the Milliken population of each atom. For an isolated sulfur atom, it has sixteen electrons. As shown by the figure, the sulfur atom has 16.30 electrons in the bonded structures. Since more electrons are located on the site of sulfur atom, the sulfur atom establishes a potential barrier between the benzene molecule and the silicon surface. In the meantime, there is depletion of electrons on the points of silicon surface and the benzene molecule that contacts the sulfur atom. The formation of the potential barrier on the sulfur atom can be understood in terms of the electronegativities of the atoms. The electronegativities of the isolated sulfur atom, carbon atom, and silicon atom are 2.58, 2.55, and 1.9, respectively. Thus, the sulfur atom attracts electrons from the silicon atom contacting to it in the binding process. On the other hand, an isolated sulfur atom and a

carbon atom have almost the same electronegativity. They may form a non-polar covalent bond. However, when the sulfur atom is bonded to the surface of silicon, the scale of the electronegativity of sulfur increases as discussed before. Thus, the sulfur atom also attracts electrons from the carbon atom connecting to it and forming a polarized covalent bond, as shown by Fig.1. Because of the potential barrier formed between the benzene molecule and the surface of silicon, this type of contact is a Schottky – like contact.

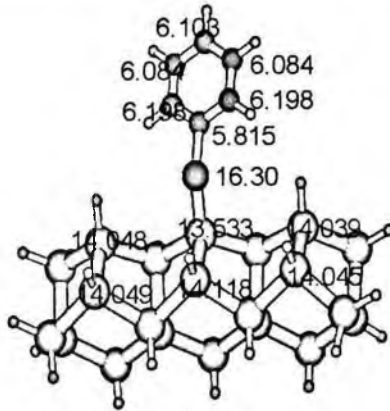


Fig. 1 Geometrical structure of the attachment of benzenethiol on the surface of silicon (100) – 2×1 surface and Milliken population of the atoms.

3.2 Electrical Characteristics of Si – C linkages

In this subsection, we consider the attachment of benzonitrile molecule onto silicon (100) – 2×1 surface. The geometrical structure of the attachment shown by Fig. 2 has been experimentally verified [20]. Comparing to Si – S linkage, this linking structure has two

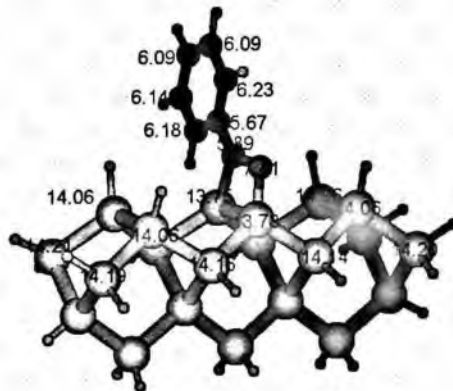


Fig. 2 Geometrical structure of the attachment of benzonitrile on the surface of silicon (100) – 2×1 surface and Milliken population of the atoms.

advantages: 1). the carbon, nitrogen, and silicon atoms form a circularly structure. This structure is stable and can retain the π -conjugation of the molecule even after the molecule is absorbed onto the surface of silicon, thereby making it possible to retain the π -conjugation of the system. This is an important technique requirement to molecular based information processing systems; 2). There is a pathway of electrons between the molecule and the silicon surface. As we know the carbon atom, silicon atom, and nitrogen atom have six electrons, fourteen electrons, and seven electrons, respectively. As shown by Fig.2, there is electron population deficiency of the carbon atoms and silicon atoms near the attachment point. Furthermore, the Milliken populations of the atoms illustrated by Fig. 2 show that there is no potential barrier between the benzene molecule and the silicon surface. This deficiency of electron population of atoms provides a potential channel for the transfer of electrons from the molecule to the surface of silicon or from the surface of silicon to the molecule. This conduction channel is created by the difference of the electronegativities of the atoms. Here, the nitrogen atom is a key factor for its electronegativity (3.04) is much larger than those of the carbon atom and the silicon atom. Thus, the nitrogen atom attracts electrons from carbon atoms and silicon atoms. It is this attraction of electrons from carbon atoms and silicon atoms that leads to the creation of the electronic conduction channel through the surface of silicon to the molecule. Since there is no potential barrier between the molecule and the silicon surface, this type of contact is Ohmic-like contact.

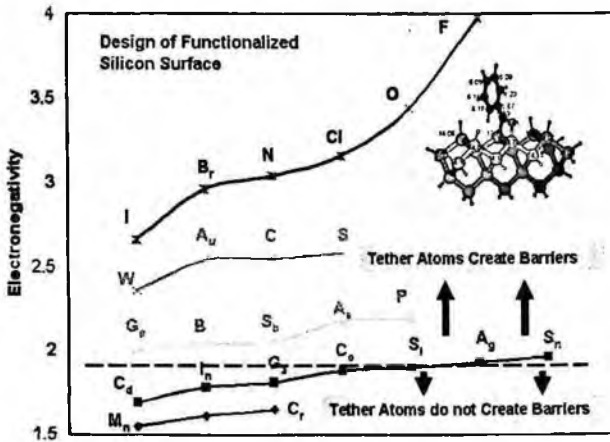


Fig. 3 Design of Functionalized silicon surfaces. The dish line in the figure is the reference line for silicon surfaces.

4. Conclusion

In this research, we have, in terms of density functional theory, studied the electrical characteristics of attachment of aromatic molecules onto silicon (100) – 2×1 surfaces. Our numerical calculation results show that, in the case of Si – S linkage, the sulfur atom,

the tether atom, establishes a potential barrier between aromatic molecules and silicon surfaces. This type of contacts is a Schottky-like contact. For Si – C linkage, the aromatic molecule may attach to the surface of silicon via a cyano group selectively binding to Si (100) through a typical 1, 2 – dipolar cycloaddition of cyano group with a silicon dimer. Here the carbon in the cyano group does not establish a potential barrier between the benzene molecule and the surface of Si (100). An electron motion channel is formed while the benzonitrile is absorbed onto the surface of Si (100) for the electronegativity of the nitrogen atom is higher than those of carbon atoms and the silicon atoms. These results illustrate basic rules for design the electrical characteristics of the contact between aromatic molecules and the silicon surfaces as shown by figure 3. The dish line in the figure is the reference line for silicon surface. The tether atoms above this line have greater electronegativities than silicon atoms. They will create barriers between the linker molecule and the surface, which form a Schottky-like contact. Below this line, the electronegativity of the tether atoms is a little bit less than that of silicon atoms. These atoms do not create barriers between the linker molecule and the surface, which form an Ohmic-like contact.

References

1. Guy A. Caldwell and Jeffrey M. Becker, Rapid PCR Sequencing of Plasmid DNA Directly from Colonies of *Saccharomyces cerevisiae*, *Promega Notes Magazine Number 44*, Nov. 1993, p.06; <http://www.people.virginia.edu/~dnaseq/dnaseq3.html>.
2. http://www.modernbio.com/dna_hybridization_analysis.htm; Jane A. Ferguson, T. Christian Boles, Christopher P. Adams, and David R. Walt; A fiber-optic DNA biosensor microarray for the analysis of gene expression, *Nature Biotechnology* 14, 1681 - 1684 (1996); Palecek E, Fojta M, Tomschik M, Wang J., Electrochemical biosensors for DNA hybridization and DNA damage, *Biosens Bioelectron.* 15;13(6):621-8 (1998).
3. Ultrathin Electrochemical Chemo- and Biosensors Technology and Performance Series: Springer Series on Chemical Sensors and Biosensors, Vol. 2 Mirsky, Vladimir M. (Ed.) 2004; ELECTROCHEMISTRY OF NUCLEIC ACIDS AND PROTEINS : *Towards Electrochemical Sensors for Genomics and Proteomics*, Edited By E. Palecek, F. Scheller, J. Wang .
4. B. Lillis, E. Hurley, S. Galvin, A. Mathewson, H. A. Berney, A novel, high surface area, capacitance based silicon sensor for DNA hybridization detection, The 16th European Conference on solid-state transducers, Prague, Czech Republic, Sept. 15-18, 2002; Moreno-Hagelsieb, L.; Lobert, P.E.; Pampin, R.; Remacle, J.; Flandre, D.; DNA detection based on capacitive Al₂O₃/Al microelectrodes Micro Electro Mechanical Systems, 2004. *17th IEEE International Conference on. (MEMS)*, 2004 Page(s):308 - 311
5. Thomas Ming-Hung Lee, Lu-Lu Li, and I-Ming Hsing, Enhanced Electrochemical Detection of DNA Hybridization Based on Electrode-Surface Modification, *Langmuir*, 19 (10), 4338 - 4343, 2003.
6. Wei Cai, John R. Peck, Daniel W. van der Weide, Robert J. Hamersa, Direct electrical detection of hybridization at DNA-modified silicon surfaces, *Biosensors and Bioelectronics*, 19(9), 1013(2004).
7. Sonja B. Braun-Sand and Olaf Wiest, Theoretical Studies of Mixed-Valence Transition Metal Complexes for Molecular Computing. *J. Phys. Chem. A* 107, 285(2003).
8. J. Twamley, Quantum-Cellular-Automata Quantum Computing with Endohedral Fullerenes, *Phys. Rev. A* 67, 052318(2003).

9. Sonja B. Braun-Sand and Olaf Wiest, Biasing Mixed-Valence Transition Complexes in Search of Bistable Complexes for Molecular Computing, *J. Phys. Chem. B* 107, 9624(2003).
10. M. Preuss, W. G. Schmidt, K. Seino, and F. Bechstedt, Methylchloride adsorbed on Si (100): an ab initio study, *Applied Surface Science* 234, 155(2004).
11. Feng Tao, Xhong Hai Wang, Yee Hing Lai, and Guo Qin Xu, Attachment of Styrene and Phenylacetylene on Si (111) - 7×7 : The influence of Substitution Groups on the Reaction Mechanism and Formation of π -conjugated Skeletons, *J. Am. Chem. Soc.*, 125, 6687(2002).
12. M. V. Lee, D. Guo, M. R. Lindford, and H. Zuihof, Molecular Modeling of Alkyl Monolayers on the Silicon (100) - 2×1 Surface, *Langmuir*, 20 9108(2004).
13. F. Nunzi, et al, Density Functional Study on the Interaction of a Polycyclic Aromatic Molecule and the Silicon (001) Surface, *J. Phys. Chem. B* 108, 10881(2004); F. Tao, et al, Formation of a Benzoimine-like Conjugated Structure through the Adsorption of Benzonitrile on Si(100), *J. Phys. Chem. B* 106, 3557(2002).
14. S. K. Coulter, M. P. Scheart, and R. J. Hamers, Sulfur Atoms as Tethers for Selective Attachment of Aromatic Molecules to Silicon (001) Surfaces, *J. Phys. Chem. B* 105, 3079(2001).
15. F. Tao, et al, Dative and Di- Binding States of Pyridine on Si(100) and Their Thermal Stability, *J. Phys. Chem. B* 107, 6384(2003); W. F. Bergerson, et al, Assembly of Organic Molecules on Silicon Surfaces via the Si-N Linkage, *JACS* 121, 454(1999)
16. Jürgen Fritsch and Pasquale Pavone, Ab initio calculation of the structure, electronic states, and the phonon dispersion of the Si(100) surface, *Surface Science*, V344, 159(1995).
17. James Shoemaker, Larry Burggraf, and Mark S. Gordon, An ab initio cluster study of the structure of the Si(100) surface, *J. Chem. Phys.*, 112, 2994(2000).
18. S. B. Healy, C. Filippi, P. Kratzer, E. Penev, and M. Scheffler, Role of Electronic Correlation on the Si(100) Reconstruction: A Quantum Monte Carlo Study, *Phys. Rev. Lett.* 87, 016105(2001).
19. W. A. Hofer J. Fisher, G. P. Lopinski, and R. A. Wolkow, Absorption of Benzene on Si (100)- 2×1 : Absorption energies and STM image analysis by *ab initio* Methods, *Phys. Rev. B* 63, 085314(2001).
20. F. Tao, Z. H. Wang, and G. Q. Xu, Formation of a Benzoimine-like Conjugated Structure through the Absorption of Benzonitrile on Si (100), *J. Phys. Chem.* 106, 3557(2002).
21. Y. P. Zhang, L. Yang, Y. H. Lai, G. Q. Xu, and X. S. Wang, Self-assembly of one-dimensional molecular nanostructures on the Ge-covered Si (100) surface, *Appl. Phys. Lett.* 84, 401(2004).
22. M. P. Schwartz, R. J. Halter, R. J. McMahon, and R. J. Hamers, Formation of an Atomically Abrupt Interface between a Polycyclic Aromatic Molecule and the Silicon (001) Surface via Direct Si - C Linkage, *J. Phys. Chem. B* 107, 224(2003).

COUMARIN DYE AS A FLUORESCENCE SENSOR FOR METHANOL VAPOR

NATHAN STEVENS

DANIEL L. AKINS*

*Center for Analysis of Structures and Interfaces (CASI)
Department of Chemistry, The City College of New York, Convent Avenue and 138th Street
New York, New York 10031, USA
akins@sci.ccny.cuny.edu*

The sol-gel method has been employed in the fabrication of mesoporous composite films consisting of a nonionic surfactant, Pluronic P123, as the organic component, and silica as the inorganic component. The hybrid nature of these films resulted in their having an internal structure consisting of nanometer size self-assembled organic mesostructures surrounded by a silica framework. These films served as the host matrix for the laser dye coumarin 481 (C481) and an energy transfer complex formed between C481 and J-aggregated meso-tetra(4-sulfonatophenyl)porphyrin (TSPP). Upon exposure to methanol vapor, a rapid and reversible decrease in fluorescence intensity occurs for films containing C481 alone as well as containing both C481 and TSPP. Steady-state and time-resolved spectroscopic studies suggest that the decrease in fluorescence intensity is primarily due to an excited state interaction between methanol and C481; while, additionally, morphological changes within the film appear to play a role for films containing both C481 and TSPP.

Keywords: Coumarin dye; Methanol; Porphyrin; J-aggregate; Sol-gel; Composite film; Optical sensor

1. Introduction

Over the years, thin-film technology has been widely employed in the development of artificial olfactory systems for the detection of a wide range of volatile organic compounds (VOC).¹⁻⁴ The success of these films is in large part due to the ability of VOC to rapidly diffuse through and interact with the intercalated sensing agent or agents, inducing a measurable response. The ability to readily deposit such films on a variety of substrates is another factor that has contributed to their successful adoption. Depending on the composition of the film, the sensing response may be optical or electrical in nature. An electrical response is most often measured as a change in the electrical current on exposure to VOC vapors,^{5,6} while an optical response is often a change in absorption intensity at particular wavelengths.^{7,8} Moreover, the flexibility inherent to thin-film technology has led to continued research in the development of films with desirable sensing characteristics such as fast response, good analyte discrimination, high sensitivity, and reversibility.

We report here on the changes in fluorescence intensity of a coumarin 481 and a coumarin/TSPP (C481/TSPP) complex that are intercalated into mesoporous inorganic/organic composite films on exposure to methanol vapor. Films were prepared from sols that utilized silica, the inorganic component, and the nonionic surfactant Pluronic P123, *i.e.*, poly(ethylene oxide)-poly(propylene oxide)-poly(ethylene oxide) (PEO-PPO-PEO) triblock copolymer (structure shown in Fig. 1), serving as the organic component. Steady-state fluorescence measurements indicate that the methanol vapor is able to rapidly diffuse into the film and result in a reduction in the fluorescence intensity of the sensing agent. Moreover, a combination of time-resolved and steady-state fluorescence measurements have aided in understanding the change in fluorescence intensity, when C481 and TSPP are present, as primarily attributable to an excited state interaction between methanol and an energy transfer complex. The data also suggests that a change in film internal morphology enhances the response to methanol vapor.

2. Experimental

The synthesis of the silica/organic sol follows a formulation previously reported to prepare a pure silica sol.⁹ In the present study, 40.0 ml of tetraethylorthosilicate (TEOS), 106.6 ml ethanol, 11.4 ml distilled water, and 0.2 ml of HCl (2.6 M) were placed into a 250 ml Erlenmeyer flask and vigorously stirred for 1 hour. Ten grams of Pluronic P123 (BASF) triblock copolymer paste was then added to the resultant partially hydrolyzed sol and the mixture was stirred for an additional 4 hours until the sol was visually clear. To aliquots of this sol, C481 (alone as well as in combination with TSPP) was added, with sonication used to promote dissolution. The concentration (in mg/ml of sol) of C481 used was 1.5 mg/ml, while that of TSPP was 0.04 mg/ml.

Films were fabricated by drop-casting the dye/sol mixture onto glass capillary tubes or glass coverslips. The chemical structures of C481, TSPP and Pluronic P123 are shown in Fig. 1.

UV-Vis absorption measurements of the coated films and solution samples were acquired using a Perkin-Elmer Lambda 18 spectrometer while fluorescence spectra were obtained using an Ocean Optics fiber optic spectrometer (Model HR4000). Time-resolved fluorescence spectra were obtained using a streak camera system, described in detail elsewhere.¹⁰ Film thickness was measured using a Mikropack NanoCalc-2000 spectroscopic reflectometer. Surface topology and roughness were measured using a Thermomicroscope ExplorerTM atomic force microscope (AFM). The pore structure of the composite was determined from adsorption of nitrogen at its boiling point: an ASAP 2010 analyzer (Micromeritics, Inc., Norcross, GA) was used for the measurements of

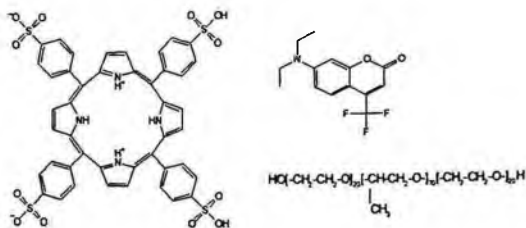


Fig. 1. Chemical structures of C481, TSPP, and Pluronic P123.

adsorption isotherms. The pore structures of three composite samples were measured: one was formed from the pure sol; a second contained ~ 0.2 wt% antimony doped tin oxide nanoparticles, Nanophase Technologies (NanoTek[®]), with an average particle size quoted as 20 nm; and the third contained ~ 0.2 wt% of an organic dye.

For methanol sensing, a simple vapor generation and fluorescence detection system was utilized. A diagram of this system is presented in Fig. 2. The generation of methanol vapor was accomplished by flowing a carrier gas, in this case dry air, through a 20 ml

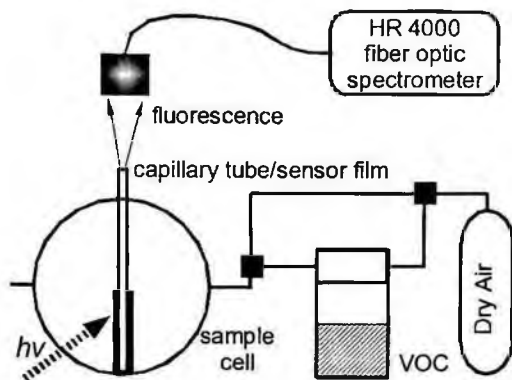


Fig. 2. Schematic diagram showing the key vapor generation and fluorescence detection components of the sensor evaluation system. VOC refers to a volatile organic chemical.

capillary tube acted as a waveguide, directing the fluorescence from the intercalated dye to the collection lens of the Ocean Optics spectrometer. The concentration of methanol vapor was obtained analytically by measuring the decrease in methanol volume in the sample vial over a known period of time while maintaining a constant flow rate of air. Although simple in design, this setup allowed the rapid evaluation of numerous sets of samples.

3. Results and Discussion

In deciding on the composition of the thin-film a number of factors needed to be taken into account. First and foremost, was the capability of the film's matrix to effectively solubilize the high concentration ($\sim 10^{-3}$ M) of C481 and TSPP used. The film also needed to be porous to permit the analyte vapor to rapidly diffuse through and interact with the dye. Moreover, high quality films needed to be fabricated with relative ease, using the simplest of deposition techniques possible. As a result of these requirements, the materials that we found most suited for fabrication of the thin-film was a sol-gel derived inorganic/organic composite. This type of composite has the solubilizing properties of a polymer while retaining the porosity of a sol-gel derived xerogel.¹¹

Although several organic molecules were tried as the organic component for the composite film, the one that resulted in the highest quality film was the nonionic surfactant Pluronic P123.¹² One advantage of using this surfactant is that films can be readily fabricated using the simple drop-casting method. AFM topographic images reveal that the resultant films, in our setup, are formed with average thickness of $6 \pm 0.5 \mu\text{m}$ and have an average RMS surface roughness of only 2.4 nm over a $100 \times 100 \mu\text{m}$ area; the latter determination indicates that there are few surface defects, which are due, undoubtedly, to the films' inorganic/organic hybrid nature, with Pluronic P123 functioning as a drying control agent.¹³

Given the concentration of Pluronic P123 used, the quality of the films should also be dependent on the specific interactions that take place between the copolymer and developing silica network. Several studies have shown that Pluronic P123 molecules self-assemble into rod-shape nanometer size mesostructures that serve as templates around which the silica network forms.¹⁴ Within the resulting films, these supramolecular assemblies occupy a significant portion of the overall volume, leading to distinct hydrophobic domains. It is these hydrophobic domains that serve as a favorable microenvironment in which polar organic dyes can reside. Support for the existence of these mesostructures is provided by the observation that when no dye is added to the sol, the resulting composite is nonporous. Upon dye addition, however, the composite becomes mesoporous, with an average pore size distribution of $2.7 \pm 0.1 \text{ nm}$.

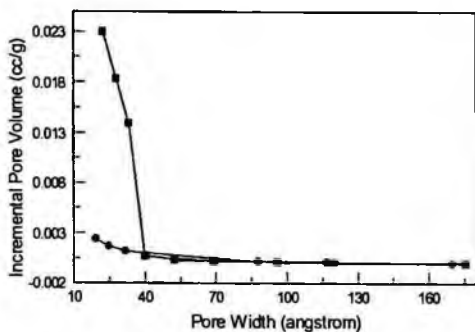


Fig. 3. Plot of pore volume versus pore diameter of the composite materials containing an organic dye (filled squares) or antimony doped tin oxide nanoparticles (filled circles) at equivalent weight percent.

As a reference system, antimony doped tin oxide nanoparticles were added to the sol at the same weight percent as was used for the dye. In this case, the resulting composite was found to be of significantly reduced porosity, supporting the supposition that the hydrophobic interactions between the dye and the Pluronic P123 drive the formation of the rod-shape mesostructures. The porosity difference between the dye and nanoparticle containing composites can be shown in Fig. 3.

As mentioned earlier, the primary aim of this study was to use changes in an intercalated dye's fluorescence intensity as a means to detect the presence of methanol vapor; for the present study the highly luminescence laser dye C481 is chosen. Prior work has already shown that the fluorescence intensity of a closely related dye, coumarin 460, was dependent on the methanol concentration in solution.¹⁵

Additionally, for the present study, as an extension of our efforts, TSPP was co-intercalated along with C481 in order to form a Coulombic-type energy transfer complex.

For this latter case, individual TSPP molecules are expected to self-assemble into J-aggregate structures.¹⁶ Although these J-aggregates strongly absorb in the spectral region where C481 fluoresces, they have very low radiative quantum yields. As a result of these properties a long-range energy transfer process from C481 (donor) to TSPP J-aggregates (acceptor) should result. Furthermore, due to the distance dependence of such an energy transfer process,¹⁷ the fluorescence from the complex should be sensitive to changes in the film's internal structure. If the film expands when exposed to methanol vapor then these two molecular species will move further apart, likely leading to a decrease in the energy transfer efficiency and conversely an increase in the fluorescence intensity; if the film contracts, then the opposite would occur. Having such a complex should allow changes in the film's internal morphology, in the presence of methanol vapor, to be actively probed.

The absorption and fluorescence spectra of the C481 and TSPP are shown in Fig. 4. The absorption maximum of C481 is similar to that in solution.¹⁸ The absorbance spectrum of TSPP, on the other hand, shows clear evidence of J-aggregate formation: sharp narrow band centered at 490 nm.¹⁶ Evidence for the formation of the energy transfer complex can be seen when the fluorescence spectra of the C481 films with and without TSPP are compared. When no TSPP is present, the fluorescence spectrum of C481 is characterized by a single broad band, but when TSPP is present there is a sharp dip in the spectrum at the wavelength corresponding to the J-aggregate (490 nm). Moreover, the shape of the fluorescence spectrum is strongly dependent on the collection geometry. As is indicated in Fig.4, the spectrum acquired using the right-angle (RA) configuration differs from that observed for the spectrum acquired in the front-face (FF) configuration. This is most likely due to the well-known inner filter effect.¹⁷ In the right-angle configuration, the photons that are detected have traveled a much greater distance through the film than those detected in the front-face configuration. As a result of this, the emitted photons from the C481 molecules are absorbed and scattered by the TSPP J-aggregates to a greater degree, leading to significant changes in the fluorescence band shape. In a sense, the TSPP J-aggregates act as miniature spectral filters for the emitted C481 photons.

A histogram showing the variation in the fluorescence intensity when a composite film containing C481 is exposed to air then methanol vapor over several cycles is presented in Fig. 5. The histogram for films containing the C481/TSPP energy transfer

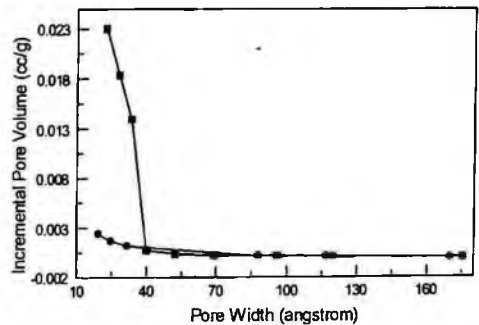


Fig. 4. Normalized absorption (ab) and fluorescence (fl) spectra of films containing C481 and TSPP. Fluorescence spectra are obtained at both the right angle (RA) and front face (FF) collection geometry using an excitation wavelength of 400 nm.

complex is shown in Fig. 6. Each cycle consists of dry air flow, at a rate of 500 ml/min for 30 seconds, followed by dry air/methanol vapor (concentration of ~ 150 ppm) for 30 seconds at the same flow rate. A 30 second interval is chosen since no further spectral changes are observed to take place at this point. The fluorescence intensity is obtained by spectral integration in the region from 490 to 540 nm. Several measurements were made on each type of sample to ensure the reproducibility of the results.

When films containing C481 is exposed to methanol vapor, the histogram that results (Fig. 5) indicates a clear response in the fluorescence intensity. After the initial exposure, the intensity decreases by $\sim 40\%$ and recovers to $\sim 94\%$ when dry air is then introduced

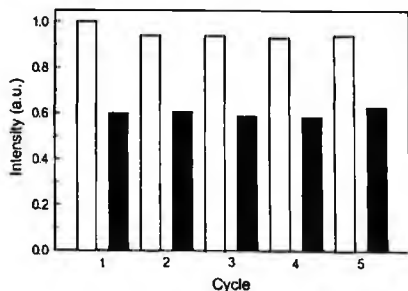


Fig. 5. Histogram of the response to methanol vapor (black bars) at a concentration of 150 ppm for a film containing C481.

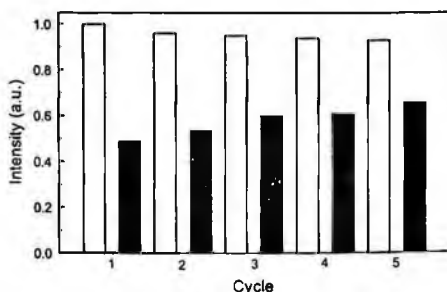


Fig. 6. Histogram of the response to methanol vapor (black bars) at a concentration of 150 ppm for a film containing the C481/TSPP energy transfer complex.

into the sample cell. This trend is more or less consistent over the remaining cycles. A noticeable aspect of this data is that the fluorescence intensity never fully recovers. One likely reason for this is that a small fraction of the methanol molecules become trapped within the film and cannot be completely removed. Such a transient effect has been observed for other methanol sensing films.⁵ To test this assertion, dry air was flowed through the sample for an extended period of time (30 min), and resulted in an almost complete recovery of the fluorescence intensity (98%). The lack of complete recovery may indicate an irreversible morphological change occurring within the film or, possibly, a chemical change in the C481 molecules.

As previously reported by Dadge *et al.*,¹⁵ the observed decrease in fluorescence intensity is likely due to an excited state interaction between methanol molecules and C481 molecules in the S_1 state. Upon photoexcitation, C481 forms radiative intramolecular charge transfer (ICT) state in which electron density from the nitrogen is transferred to the oxygen atom.¹⁸ This results in the nitrogen having a partial positive charge and the oxygen a partially negative charge. Interaction with methanol molecules would help to stabilize this highly polar excited state, while likely resulting in additional nonradiative decay pathways to the ground state, and thus, a decrease in the fluorescence intensity. Direct evidence for the formation of this ICT excited state is provided by time-resolved fluorescence data. In methanol solution, C481 exhibits a single component lifetime (0.34 ns), but within the film a multicomponent lifetime is found with time

constants of 5.29 ns (66 %), 1.75 ns (15 %), and 4.78 ns (19 %) (Data not shown). The 1.75 ns component is a rise component and indicates the formation of the radiative ICT state from the initial excited state species as time progresses.

For films containing C481 and TSPP, the response to methanol vapor is a fluorescence intensity decrease whose magnitude, unlike when C481 alone is the sensor, is dependent on the number of exposures to methanol (*i.e.*, the sensing cycles). As can be seen (Fig. 6), on initial exposure to methanol vapor the fluorescence intensity decreases to ~50 % of the original, versus ~60 % for films containing only C481. However the enhanced magnitude of this response is not retained and comes into parity with the C481 films over several sensing cycles. Furthermore, the decrease in intensity at cycle 5 for these films is not as great as that for the C481 films, indicating the loss of the initial enhancement. Moreover, we found that the fluorescence intensity does not recover to the same extent as for C481 films, despite an extended period of dry air flow.

The results described above for the C481/TSPP J-aggregate film can be explained as due to a morphological change occurring within the film's structure. The initial enhancement in methanol detection can be explained in terms of Förster energy transfer theory.¹⁷ According to this theory, the overall energy transfer efficiency from a donor (*i.e.*, C481) to an acceptor (*i.e.*, TSPP J-aggregate) is strongly dependent on the average distance between the two. Since a small variation in the average distance results in a substantial change in the energy transfer efficiency, any morphological changes within the composite film that causes the average distance between the C481 and TSPP J-aggregate to vary, would result in a change in the fluorescence intensity. If the distance between the donor and acceptor decreases, then the fluorescence intensity will undergo a greater decrease in fluorescence intensity compared to the single fluorophore system, it appears that a morphological change occurs within the film that brings the C481 and TSPP J-aggregates into closer proximity. The loss of this enhanced response over several sensing cycles indicates that this morphological change is not fully reversible and helps explain why the fluorescence intensity cannot be fully recovered to the initial value. In other words, over several cycles the average distance between C481 and TSPP J-aggregates permanently decreases, thereby increasing the energy transfer efficiency, which translates into a reduction in the overall fluorescence intensity of C481.

4. Conclusion

Mesoporous films made by coating a sol gel consisting of silica/Pluronic P123 nanocomposite were prepared. The organic domains within the self-assembled mesostructures within the film act as favorable microenvironments for intercalation of the highly luminescent laser dye C481. The incorporated dye is shown to be able to indicate the presence of methanol vapor at a concentration of ~ 150 ppm as a result of changes in the fluorescence intensity. The co-intercalation with C481 into the films of TSPP J-aggregates led to the formation of an energy transfer complex that exhibits an enhanced response to methanol vapor, but only for the first few sensing cycles, at which point the response returns to a constant value close to that for the system in which only

C481 is present. This enhancement is associated with a morphological change within the film that leads to the C481 (donor) and TSPP J-aggregates (acceptor) coming into closer proximity, and thereby, as a result of energy transfer to the aggregate, a decrease in the fluorescence intensity of C481. Future work will focus on evaluating the sensing response of this system and those combining other dyes that would sense a range of volatile organic compounds. The film's composition also will be modified in an attempt to enhance sensing response of an energy transfer complex and to increase the number of enhanced sensing cycles.

5. Acknowledgments

DLA thanks the NSF and DoD-ARO for support of this work, in part, through the following awards: (1) the NSF-IGERT program under grant DGE-9972892; (2) the NSF-MRSEC program under grant DMR-0213574; the (3) NSF-NSEC program under grant CHE-0117752; and (4) DoD-ARO under Cooperative Agreement DAAD19-01-1-0759 and grant W911NF-04-1-0029.

References

1. J. Spadavecchia, G. Ciccarella, P. Siciliano, S. Capone, R. Rella, Spin-coated thin films of metal porphyrin-phthalocyanine blend for an optochemical sensor of alcohol vapors, *Sens. Actuators, B, Chem* **100**, 88-93 (2004).
2. M. Vollprecht, F. Dieterle, S. Busche, G. Gauglitz, K.-J. Eichhorn, B. Voit, Quantification of quaternary mixtures of low alcohols in water: temporal-resolved measurements with microporous and hyperbranched polymer sensors for reduction of sensor number, *Anal. Chem.* **77**, 5542-5550 (2005).
3. R. C. Bailey, J. T. Hupp, Large-scale resonance amplification of optical sensing of volatile compounds with chemoresponsive visible-region diffraction gratings, *J. Am. Chem. Soc.* **124**, 6767-6774 (2002).
4. B. Kieser, F. Dieterle, G. Gauglitz, Discrimination of methanol and ethanol vapors by the use of a single optical sensor with a microporous sensitive layer, *Anal. Chem.* **74**, 4781-4787 (2002).
5. M. F. Mabrook, C. Pearson, M. C. Petty, Inkjet-printed polypyrrole thin films for vapour sensing, in press.
6. L. Jiang, H.-K. Jun, Y.-S. Hoh, J.-O. Lim, Sensing characteristics of polypyrrole-poly(vinyl alcohol) methanol sensors prepared by in situ vapor state polymerization, *Sens. Actuators, B, Chem* **105**, 132-137 (2005).
7. J. H. Krech, S. L. Rose-Pehrsson, Detection of volatile organic compounds in the vapor phase using solvatochromic dye-doped polymers, *Anal. Chim. Acta* **341**, 53-62 (1997).
8. S. T. Dubas, C. Iamsamai, P. Potiyaraj, Optical alcohol sensor based on dye-Chitosan polyelectrolyte multilayers, *Sens. Actuators, B, Chem* **113**, 370-375 (2006).
9. Y. Huang, K. Chou, Studies on the spin coating process of silica films, *Ceram. Int.* **29**, 485-493 (2003).
10. N. Stevens, Nonlinear Optical Switching Properties of Dye-doped Inorganic/Organic Composite Films, Ph.D. thesis, The Graduate Center, New York, pp. 28-29 (2006).
11. M.M Collinson, Recent trends in analytical applications of organically modified silicate materials, *Trends Anal. Chem.* **21**, 30-38 (2002).

12. P. Alexandridis, T. A. Hatton, Poly(ethylene oxide)-poly(propylene oxide)-poly(ethylene oxide) block copolymer surfactants in aqueous solutions and at interfaces: thermodynamics, structure, dynamics, and modeling, *Colloids Surf., A* **96**, 1-46 (1995).
13. U. De Rossi, S. Daehne, R. Reisfeld, Photophysical properties of cyanine dyes in sol-gel matrices, *Chem. Phys. Lett.* **251**, 259-267 (1996).
14. A. Bearzotti, J. M. Bertolo, P. Innocenzi, P. Falcaro, E Travers, Humidity sensors based on mesoporous silica thin films synthesized by block copolymers, *J. Eur. Ceram. Soc.* **24**, 1969-1972 (2004).
15. J. W. Dudge, V. N. Krishnamurthy, R. C. Aiyer, Nitrogen laser induced fluorescence in laser dyes for sensing of organic compounds, *Sens. Actuators, B, Chem* **113**, 805-808 (2006).
16. D. L. Akins, S. Özçelik, H. Zhu, C. Guo, Fluorescence decay kinetics and structure of aggregated tetrakis(p-sulfonatophenyl)porphyrin, *J. Phys. Chem.* **100**, 14390-14396 (1996).
17. J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, 2nd. ed. (Kluwer Academic, New York, 1999).
18. S. Nad, M. Kumbhakar, H. Pal, Photophysical properties of coumarin-152 and coumarin-481 dyes: unusual behavior in nonpolar and in higher polarity solvents, *J. Phys. Chem. A* **107**, 4808-4816 (2003).

NANOSCALE IMAGING TECHNOLOGY FOR THZ-FREQUENCY TRANSMISSION MICROSCOPY

DWIGHT WOOLARD and PEIJI ZHAO

*Department of Electrical and Computer Engineering, NC State University, Raleigh, NC 27695, USA
dwight.woolard@us.army.mil*

CHRISTOPHER RUTHERGLEN, ZHEN YU and PETER BURKE

*Department of Electrical and Computer Engineering, University of California, Irvine, CA 22697, USA
pburke@uci.edu*

STEVEN BRUECK and ANDREAS STINTZ

*Center for High Technology Materials, University New Mexico, Albuquerque, NM 87106, USA
brueck@chtm.unm.edu*

A novel nanoscale-engineering methodology is presented that has potential for the first-time development of a microscope-system capable of collecting terahertz (THz) frequency spectroscopic signatures from microscopic biological (bio) structures. This unique THz transmission microscopy approach is motivated by prior studies on bio-materials and bio-agents (e.g., DNA, RNA and bacterial spores) that have produced spectral features within the THz frequency regime (i.e., ~ 300 GHz to 1000 GHz) that appear to be representative of the internal structure and characteristics of the constituent bio-molecules. The suggested THz transmission microscopy is a fundamentally new technological approach that seeks to avoid the limitations that exist in traditional experiments (i.e., that must average over large numbers of microscopic molecules) by prescribing a viable technique whereby the THz frequency signatures may be collected from individual bio-molecules and/or microscopic biological constructs. Specifically, it is possible to envision the development of a "nanoscale imaging array" that possesses the characteristics necessary (e.g., sub-wavelength resolution) for successfully performing "THz-frequency microscopy."

Keywords: THz microscopy; nanoscale engineering, bio-molecules, DNA, RNA.

1. Introduction

This paper introduces a novel nanoscale-engineering methodology for the first-time development of a microscope-system capable of collecting terahertz (THz) frequency spectroscopic signatures from microscopic biological (bio) structures. Here, the main challenge is the realization of a very sensitive spectroscopic-sensing technique that can be applied successfully to measure (and map) the THz-absorption images produced when long-wavelength radiation is applied to bio-agents that possess distinct structural features at nanoscale dimensions and below. This unique THz transmission microscopy is motivated primarily by previous studies on biological materials and agents that have produced spectral features within the THz frequency regime (i.e., ~ 300 GHz to 1000 GHz) that appear to be representative of the internal structure and characteristics of the biological samples – e.g., DNA, RNA and bacterial spores. However, these same studies

(that utilized conventional apparatus and required difficult-to-control bulk samples for averaging over large apertures) have shown the serious need for new techniques that can extract spectral/spatial information from individual microstructures without the masking/variation effects associated with bulk macroscopic samples.¹ Therefore, THz microscopy is the obvious solution for the effective probing of small bio-agents and the collection of *reproducible THz signatures* that have real value for detection applications.

The nonexistence of a THz microscope is a simple manifestation of the difficulties (i.e., diffraction limit) involved in performing deep sub-wavelength (DS) imaging using conventional approaches. Specifically, if the traditional point of view is used, the resolution size is limited by the spot-size that can be produced by a Gaussian beam. However, this is not the only avenue to producing an image. In fact, DS imaging at optical and THz frequencies is an active area of research where a number of approaches (or tricks)² are under investigation for significantly reducing the measurement aperture. For example, both interferometric and near-field (nano-probe) techniques can realize nano-resolution, and it is possible to utilize a pump (short-wavelength) plus probe (long-wavelength) method to affect the same result. However, all these methods effectively utilize a “microscopic” source (i.e., to realize a small aperture) and then collect the measured-energy with a “macroscopic” detector, and this severely limits sensitivity in spectroscopic (i.e., narrow bandwidth & small intensity) imaging applications. An obvious alternative approach is to employ a “*microscopic*” detector or imaging array (i.e., that would set small apertures) and a “*macroscopic*” source (i.e., that would provide for lots of power and increased sensitivity). Of course this is very challenging from a technology perspective, but is exactly the one proposed here. Specifically, a novel quantum-cell (Q-cell) sensor concept (with nanoscale pixel resolution) is introduced that allows for sensing the intensity of THz photon-flux via their polarizing effect on single-electrons which is combined with a proven RF single-electron-transistor (SET) based frequency-domain multiplexing approach to statistically readout the polarization-state of these electrons. As will be shown, when such Q-cell detection and RF sampling concepts are combined with advanced materials growth, fabrication & testing techniques, THz-frequency transmission microscopy becomes feasible such that spectral/spatial information can be acquired at the nanoscale. Hence, this proposed nanoscale imaging technology has potential for the study and characterization of microscopic bio-materials/agents, and therefore has important relevance for defense and security applications against bio-threats.

In particular, the specific technology objective of this work is the identification of a very sensitive instrument that can be used to successfully measure spectroscopic absorption from bio-structures that have nanoscale spatial dimensions. Specifically, since the objects of interest have dimensions on the order of 100's of nanometers or less (which is much less than the spectral wavelength), this is equivalent to realizing THz-frequency deep sub-wavelength spectroscopic imaging (THz-DSSI). Therefore, the goal is the design and demonstration of a *THz microscope* which is exceedingly difficult from a conventional point of view (i.e., since one is well below the diffraction limit). However, such a THz microscope offers significant potential payoffs for the scientific characterization of bio-materials/agents and for the development of advanced sensor methodologies. Indeed, this technology development has been motivated because THz-frequency spectroscopy is potentially useful for detecting and identifying bio-agents, but important signature-

phenomenology related problems must be solved before the technique can be made amenable to practical application. As will be discussed in the sections that follow, novel nanoelectronic-based technologies and methodologies offer promise for enabling THz-signature based detection and identification of bio-materials and bio-agents in the future.

2. Scientific Background

During the last ten years, spectroscopic measurements conducted on bio-materials/agents have produced spectral features within the THz frequency regime (i.e., ~ 300 GHz to 1000 GHz) that appear to be representative of the internal structure and characteristics the specific bio-systems under consideration – e.g., DNA, RNA and bacterial spores.¹ However, the THz spectroscopic approach has been very problematic because the spectral features observed from bulk solid-state samples tends to be very weak (i.e., ~ 1-5% local variation in spectral absorption) and of limited number within the band (i.e., < 50-100 spectral features).² Also and very noteworthy to the THz sensing problem are recent results that suggest the underlying dynamics (i.e., vibrational modes) that produce the THz-frequency absorption characteristics might be seriously underestimated by measurements performed on large samples made up of microscopic bio-molecules or bioparticles (i.e., either ordered or random in nature). In particular, recent theoretical and experimental results suggest that the absorptions characteristics of isolated bio-systems may possess optically active vibrational (or phonon) modes that produce spectral absorption features that are much stronger and/or sharper (i.e., and therefore less overlapping between individual spectral signatures) than those obtained from measurements made on bulk samples.^{2, 3} Here it is important to note that bulk samples have been required in almost all the reported THz spectral studies to date because a sufficiently large and uniform aperture must be utilized in order to successfully measure the very weak phenomenon within the microscopic bio-system.⁴ Needless to say, one obviously needs extremely high sensitivity in performing such measurements and when large collections of either randomly placed or highly ordered bio-molecules are considered it becomes extremely difficult to execute repeatable spectroscopic experiments because it is difficult to prepare identical systems and to subsequently control them during testing. Clearly, a fundamental approach for avoiding the previously discussed limitations is to prescribe a new technique whereby the THz signatures can be collected from individual bio-molecules, and this motivates the search for a viable THz microscopy technology as is presented here.

From the usual perspective, it might at first seem somewhat foolish to image with long wavelength radiation due to the diffraction limit – i.e., if one wants more resolution for contrast imaging then a shorter wavelength is always better. However, the specific goal here is to collect spectral signatures within the THz regime and to do this for very small target objects. This being accepted, it should not be surprising that there are existing approaches for addressing this problem. Indeed, one could utilize interferometric imaging to significantly reduce resolution as this has been very effective in nanoscale fabrication, and alternatively, there are many groups actively utilizing near-field nano-probe based techniques to reduce the available aperture. Apparently, both of these approaches can significantly benefit from pump-probe type operations, so it is certainly reasonable to attack the THz microscope problem using these techniques. Although, there are noteworthy challenges associated with the very small effective apertures. Furthermore, it

can be expected that there will be definite sensitivity limits associated with these methods when spectroscopic operation is desired. As noted earlier, both of the previous approaches effectively utilize a “microscopic” source (i.e., to realize a small aperture) and then collect the measurement energy with a “macroscopic” detector. A somewhat more obvious approach that could lead to successful microscopy would be to employ a “microscopic” detector or imaging array (i.e., to set a small aperture) and a “macroscopic” source (i.e., to provide more power and increased sensitivity). Of course, this is very challenging from a technology perspective and is usually dismissed in any discussion on the subject at hand. Indeed, if one seeks to achieve THz microscopy using a microscopic detection methodology then nanotechnology innovations of the type to be discussed here will be critical.

3. THz Microscope Concept & Supporting Device Technology

Recent research by our collaborative group have focused on the preliminary design and analysis steps that will provide the basic guidelines for the future fabrication and demonstration of novel nanoelectronic devices and circuits that can be used to rapidly measure incident THz radiation with nanoscale spatial resolution (or pixel sizes) and with very high sensitivity (i.e., approaching the quantum limit). The specific focus to date has been on the *detector elements* and *readout circuitry* that can be used to perform THz-DSSI on microscopic bio-agents (e.g., bio-cells) with nanoscale structural features. This THz Transmission Microscope concept utilizes quantum-cell detectors (QCDs) that exhibit a polarization response to select THz radiation, along with RF single-electron-transistor circuitry that can rapidly measure these polarization changes (to infer incident energy flux). A hypothetical implementation of the full-array microscope architecture is illustrated in Fig. 1(a) where QCDs (i.e., the green cubes) are designed to respond to narrow-band THz radiation through changes to the spatial-state polarizations of trapped electrons (see “right-“ & left-pointing” arrows). In this example, the scheme has been made compatible to Quantum Cellular Automata (QCA) communication for facilitating the future integration of the array. Architectures such as this could be used for transmission-mode microscopy of the type illustrated in Fig. 1(b). Here, nano-sized detectors (10-100 nm’s) allow for measuring the presence of nano-features (e.g., virus) that might absorb higher levels of the THz radiation— i.e., as indicated by the dark red spot. Clearly, this is a very feasible approach to THz-DSSI if one can define QCD’s that *respond statistically* to the number of collected THz photons (incident intensity) and if

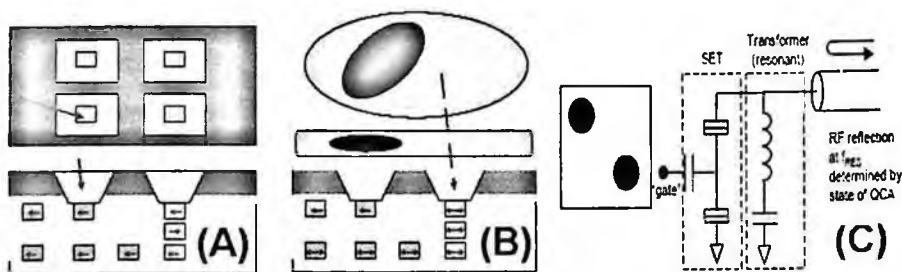


Fig. 1. Detector elements and readout circuitry. (a) Hypothetical implementation of the full-array microscope architecture; (b) THz radiation induced electron polarization and the scheme of communication for future integration of the arrays; (c) RF measurement of the polarization of the trapped electron in the detection cell.

the associated polarization processes are amenable to electrical measurement (i.e., *long time constants*) by known techniques for the RF measurement of the QCA state such as illustrated in Fig. 1(c). A summary of the technology elements and proposed measurement scheme will now be given to illustrate the basic approach.

3.1 Quantum-Cell Detector (QCD) Concept

The basic approach presented here for achieving THz microscopy leverages single-electron transition-processes and their dependence to long and short wavelength radiation. Hence, the first required element is a QCD that responds to single-photon excitation with a change of electron polarization. In a practical implementation, both vertically-defined heterostructure-semiconductor Q-dots and laterally-defined nano-probe induced Q-dots are realistic options as both have merits for realizing the needed QCD component. However, the fundamental principles of the detection methodology are perhaps best understood from the QCD which is illustrated in Fig. 2 which consists of a vertically-grown multi-layered Q-dot set into a shielded configuration to define the nanoscale aperture. Here, as show in Fig. 3, the sensor consists of a three-well Q-dot system where the first-well serves as a single-electron source (SES) that allows for controlled injection of a single-electron into a double-well pair that makes up the THz Radiation Detection Cell (THz RDC). This type of architecture is important for defining the proper quantum-states and the associated photonic processes that are necessary for enabling a statistically-based measurement of the incident radiation intensity that will actually allow for the THz microscopy function as is explained in the next subsection.

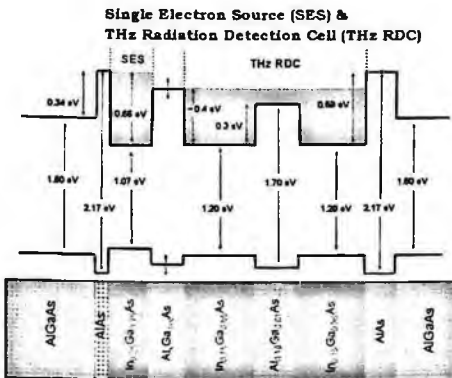


Fig. 2. Vertically-grown multi-layered Q-Dot Detector with nanoscale aperture.

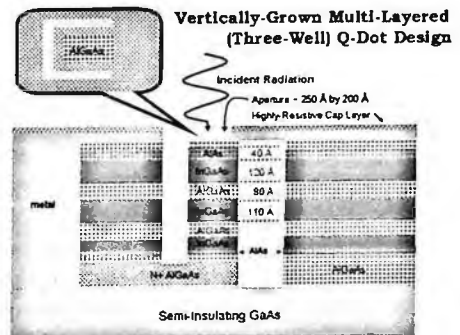


Fig. 3. Semiconductor heterostructure and Band structure for Q-Dot Detector

3.2 QCD Measurement Approach

The QCD sensor utilizes a triple-well Q-dot system to enable monitoring of the induced polarizations of a single electron due to exposure to THz radiation. Here, as shown in Fig. 4, the SES allows for optical-pulsed injection of a single electron into the THz-RDC, which has been asymmetrically engineered to possess THz energy-separations between the left ground-state and the right excited-state. Here, the microscopy measurement of the transmitted THz power through the nano-sized aperture is now amenable to statistical RF sampling of the RDC polarization using a single-electron-transistor SET based approach

that will be discussed later. Indeed, RF sampling of the resulting polarization process (i.e., via a SET impedance measurement) over many cycles of the combined THz-IR excitation enables a statistical determination of the transition rate (and therefore the THz field intensity) by the method as illustrated in Fig. 5. Here, the statistical rate of removal of the injected electron from the RDC is dictated by the THz induced transition rate τ_E and since this is fundamental quantum mechanical process it can be assessed by sampling over many transitions to numerically determine the value of the incident field intensity.

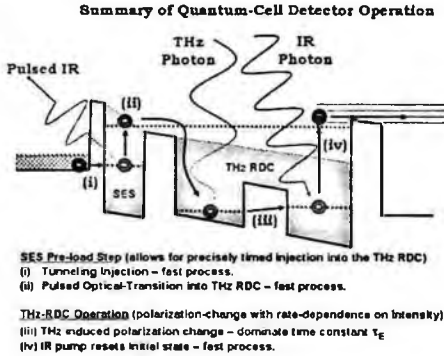


Fig. 4. Description of the electronic and photonic dynamics required for QCD operation.

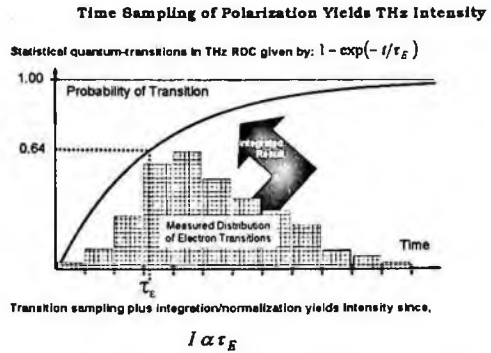


Fig. 5. Statistical determination of transition rate and associated intensity of incident electric field.

3.3 QCD Feasibility Analysis

The approach of statistically counting the quantum transitions (which in a sense is counting the THz photons) to infer the value of the incident field intensity, I , at THz frequencies sets specific requirements on the electron coherence time and the energy-state separation tuning ability associated with the QCD device. Specifically, the THz RDC must be physically compatible with state-of-the-art pulsed RF measurement techniques as the expected sampling rates (e.g., $\sim \mu$ secs) will dictate the ultimate sensitivity of the approach. This fact was illustrated in Fig. 5 and since one needs to numerically integrate the results of the measured transition-distribution, a time increment equal to μ secs must be small in comparison to the value of τ_E . Fortunately the THz QCD offers much latitude for engineering both the THz detection-frequency, f_{THz} , and the transition-rate τ_E , as will now be demonstrated. The technique presented for collecting a nanoscale-size pixel of information at THz frequencies relies on the technological capability for estimating the transition-rate of the excited electron which must be performed statistically. In order to accomplish this, one must design the QCD such that τ_E is long compared to the sampling rate. Since the aperture is sub-wavelength at THz frequencies, there will be a significant photonic mismatch (e.g., low quantum efficiency) and this is actually an advantage because this will lead to a longer value of τ_E . However, it is still important to insure that the QCD offers flexibility for designing τ_E around the requirements for frequency-tuning to allow for some spectral contrasting of the image. Specifically, one would like to be able to specify long values of τ_E independent of the value of the incident THz intensity. Preliminary studies of the QCD operation show that this is the case, as will now be discussed. Since it is possible to specify the SES electron injection and the IR-induced elimination of the electron after the THz excitation as very

fast processes (see Fig. 4), an analysis of the THz RDC alone is sufficient to estimate the latitude in designing for long τ_E times. As summarized in Fig. 6, a Schrodinger-equation based analysis of the RDC was applied to determine the ground and excited state of the trapped electron, and these results along with the associated spatial distribution for a zero-biased structure that yields $f_{THz} = 600$ GHz are illustrated. This same approach was then used to assess the both the ability to voltage-tune f_{THz} and to prescribe the dependency of τ_E on radiation intensity according to the RDC design and these results are summarized in Fig. 7. Hence, the collective simulation results given in Figs. 6 and 7 demonstrate that the QCD approach is viable for deriving the intensity (and power) of the transmitted THz beam over a nanoscale aperture, if a technological approach can be implemented (see next sub-section) for sampling the single-electron transitions at μsec sampling rates. It is also noteworthy to point out that this ability to design for very long transition-times (or equivalently the ability to suppress the scattering rates to $\sim 10^5 \text{ s}^{-1}$) has important ramifications for designing quantum gates for quantum computing applications as has been discussed in a companion publication.⁵

Schrödinger Based Simulations of THz RDC Operation

Numerical simulations for electron wavefunction over 300-900 GHz Range

$$H_0 \Psi = E \Psi \quad H_0 = -\frac{\hbar^2}{2m} \nabla^2 + V(\vec{r})$$

$$\Psi(\vec{r}) = \frac{2}{\sqrt{L_x L_y}} \sin\left(\frac{2\pi x}{L_x}\right) \sin\left(\frac{2\pi y}{L_y}\right) \phi(z)$$

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dz^2} + V(z)\right) \phi(z) = E_z \phi(z)$$

$$E_z = E - \frac{\hbar^2}{2m L_x^2} - \frac{\hbar^2}{2m L_y^2}$$

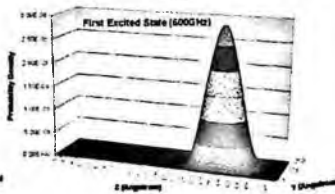
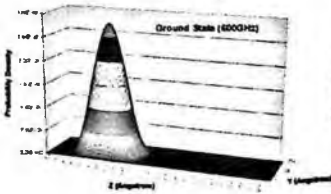
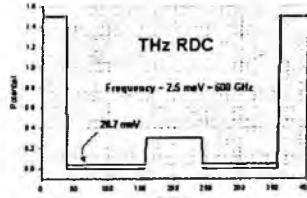


Fig. 6. Schrodinger-function based determination of energy states in THz RDC.

Analysis of Voltage Tuning & Transition-Time (τ_E)

THz RDC operation dominated by the THz-induced τ_E which is derivable as

$$\omega_A = \frac{2\pi}{\hbar} \left| \langle \Psi_f | H_{op}' | \Psi_i \rangle \right|^2 \delta(E_f - E_i - \hbar\omega)$$

$$\langle \Psi_f | H_{op}' | \Psi_i \rangle = -\frac{eA_0}{2\hbar} (E_f - E_i) \delta < \Psi_f | \tau | \Psi_i \rangle \quad A_0^2 = \frac{2I}{c \epsilon_0 \epsilon_r \omega^2}$$

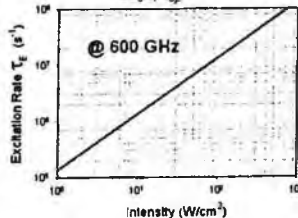
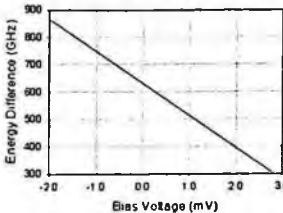


Fig. 7. Analysis of energy state dependence on voltage and excitation rate dependence on intensity.

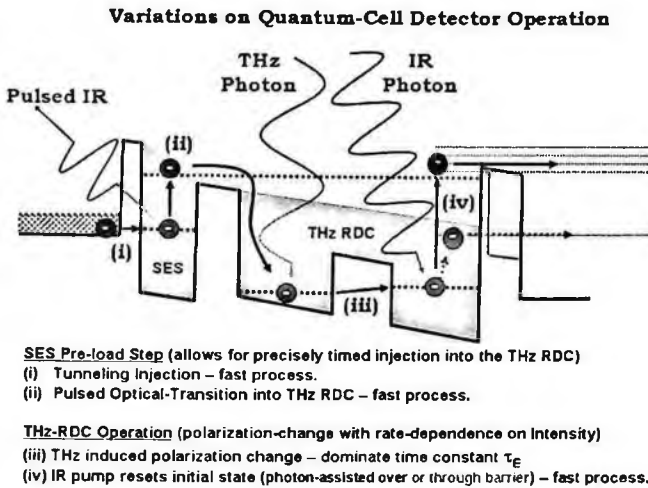


Fig. 9. Two variations in IR pump-assisted QCD design.

the right-most barrier (see modified barrier-structure in green and associated process given by the blue-dashed arrows).

The ability to obtain single-photon counting using methods such as avalanche photon detectors and negative electron affinity photocathode photomultipliers has thus far been limited to the visible and infrared regions. The vertical QCD which utilizes a triple-well quantum (Q) dot system of the type illustrated in Fig. 9 offers a novel approach to sense THz radiation. Here, the detector is first primed into active-mode by tunnel injection into the top Q-dot (QD1) of the SES followed by an IR pulse that puts the electron into the middle Q-dot (QD2) of the THz-RDC. This electron will remain in the QD2 until a THz photon induces the electron's transition to QD3. Finally, an IR photon ejects the electron from QD3 thus resetting the detector. Since the electron injection into the QCD system and ejection from the detector are quick transitions, only the middle Q-dot (QD2) will be occupied for a significant period of time. Consequently, to successfully read-out the state of our triple Q-dot system one must be able to differentiate between the two possible states: (1) if THz photons are present, the electron will quickly be ejected from the entire QCD system and no electrons will be present in QD2; or (2) if no THz photon is present, QD2 will remain occupied by an electron.

The approach proposed here for distinguishing between these two states is to utilize a carbon nanotube (CNT) RF single-electron transistor (RF-SET) that is capacitively coupled to the quantum-dot based detector as shown in Fig. 10. An RF-SET is advantageous for this application because it is an ultra-sensitive electrometer capable of differentiating the presence of single charges and it readily allows for multiplexing where by multiple RF-SETs can be readout simultaneously with each identifiable by its characteristic resonant frequency.⁷ Researchers have demonstrated RF-SET charge of sensitivity of order $10^{-3} e/\sqrt{\text{Hz}}$. A recent advance at Yale⁸ has allowed this to be achieved at rates with sub-microsecond temporal resolution with the so-called RF-SET,

and this has been used by Rimberg to measure the charge state of a quantum dot in real time.⁶ Therefore, based on this preliminary work by other groups, it seems very feasible to determine the charge state of a quantum dot using a single electron transistor. Furthermore, utilizing a carbon nanotube as our RF-SET affords us the advantage of having a higher operating temperature as well as being easier to fabricate compared to its GaAs counterpart. Due to fabrication constraints, the QDC and the RF-SET can not be fabricated directly adjacent to one another. To reconcile this problem, a sense electrode is introduced (see Fig. 11) to act as the intermediary by capacitively coupling both the triple quantum-dot system as well as the CNT RF-SET to one another. In other words, when the RDC is in the primed-state (i.e. an electron is residing in QD2) a voltage will be induced on the sense electrode. Since this electrode also doubles as the gate for the

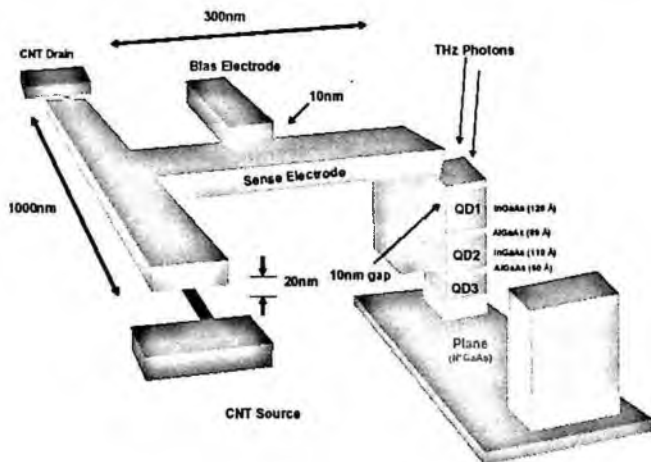


Fig. 10. The proposed structure of the SET-based readout circuit for THz detector. The quantum detector cell (QDC), which is the triple Q-dot system, is capacitively coupled through the sense electrode to a carbon based nanotubes RF-SET. A bias-electrode is used to control the bias potential on the sense electrode.

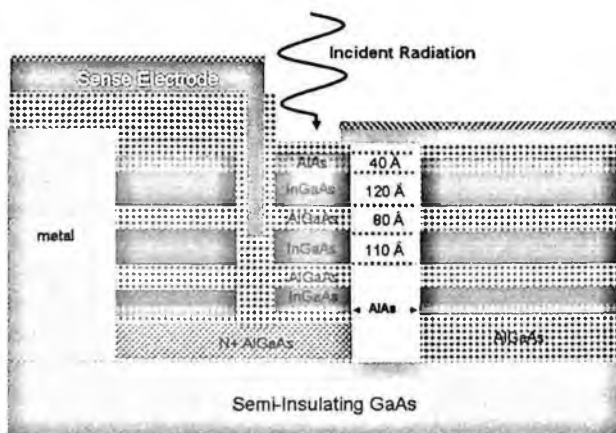


Fig. 10. A side view of the sense electrode for RF-SET integrated with the QDC device.

RF-SET, the induced voltage change will be translated over to the carbon nanotube RF-SET in the form of a change in the nanotube's impedance. A bias electrode is present to properly adjust the sense electrode's potential so to maximize the change in RF-SET impedance between the two states, $Z_{CNT-RFSET}$.

3.5 SET-Readout Circuitry Simulations

To differentiate between the two possible QDC states, i.e. whether QD2 is occupied or not, a reliable value for the capacitive coupling between the quantum dots and the sense electrode is needed. To that end, simulations for the capacitance of the proposed structure shown in Fig. 10 were performed using the software program FastModel. For simplicity purposes the bias electrode was not included in these simulations because it will have a negligible affect on the results. The relation between the charge and voltage vectors can be written out as a linear set of equations as

$$\begin{aligned} Q_{1;elec} &= \alpha_{11}V_{1;elec} + \alpha_{12}V_{2;dot1} + \alpha_{13}V_{3;dot2} + \alpha_{14}V_{4;dot3} + \alpha_{15}V_{5;CNT} \\ Q_{2;dot1} &= \alpha_{21}V_{1;elec} + \alpha_{22}V_{2;dot1} + \alpha_{23}V_{3;dot2} + \alpha_{24}V_{4;dot3} + \alpha_{25}V_{5;CNT} \\ &\dots \\ &\dots \end{aligned} \quad (1)$$

where α_{ij} represents the coupling capacitance. The simulated results are the capacitance matrix

$$\alpha = \begin{pmatrix} 97.34 & -3.762 & -1.934 & -1.757 & -39.53 \\ -3.762 & 8.135 & -2.930 & -0.2794 & -0.1004 \\ -1.934 & -2.930 & 9.022 & -3.176 & -0.0890 \\ -1.757 & -0.2794 & -3.107 & 7.358 & -0.1771 \\ -39.53 & -0.1004 & -0.0890 & -0.1771 & 51.45 \end{pmatrix} \begin{matrix} elec(aF) \\ dot1 \\ dot2 \\ dot3 \\ CNT \end{matrix}$$

To convert the capacitance matrix values α into physically meaningful capacitance values, represented by the traditional variable C , consider the circuit shown in Fig. 11. For simplicity, only the interconnecting capacitances associated with the top quantum dot (QD1) have been shown.

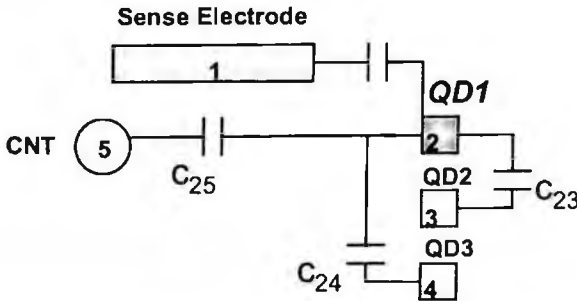


Fig. 11. An example showing all the capacitances of the top quantum dot (QD1) interconnecting with the other structures of the QDC device and SET readout circuit elements.

Here we can represent $Q_{2;dor1}$ as

$$Q_{2;dor1} = C_{21}(V_{2;dor1} - V_{1;elec}) + C_{22}(V_{2;dor1}) + C_{23}(V_{2;dor1} - V_{3;dor2}) + C_{24}(V_{2;dor1} - V_{4;dor3}) + C_{25}(V_{2;dor1} - V_{5;CNT}) \quad (2)$$

$$Q_{2;dor1} = V_{1;elec}(-C_{21}) + V_{2;dor1}(C_{21} + C_{22} + C_{23} + C_{24} + C_{25}) + V_{3;dor2}(-C_{23}) + V_{4;dor3}(-C_{24}) + V_{5;CNT}(-C_{25})$$

and by matching coefficients of $Q_{2;dor1}$ in (1) with (2) one obtains

$$\begin{aligned} \alpha_{21} &= -C_{21} \\ \alpha_{22} &= C_{21} + C_{22} + C_{23} + C_{24} + C_{25} \\ \alpha_{23} &= -C_{23} \\ \alpha_{24} &= -C_{24} \\ \alpha_{25} &= -C_{25} \end{aligned}$$

This same procedure can be carried out for all the remaining capacitance values C to obtain the coupling capacitances as shown in Fig. 12.

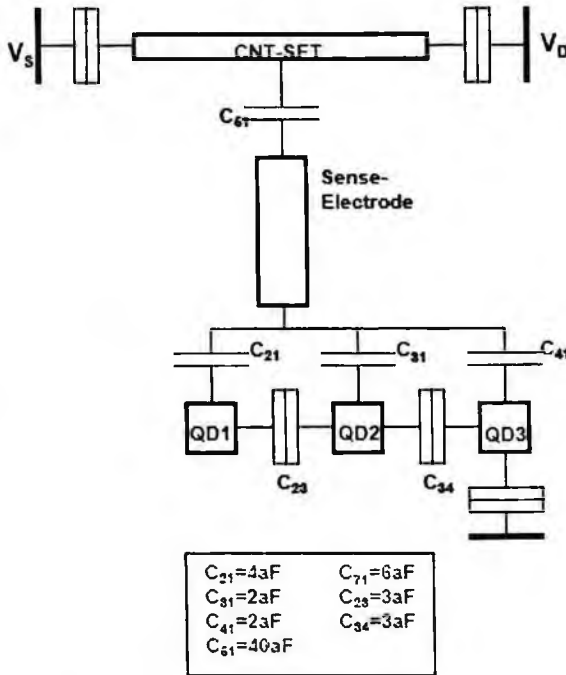


Fig. 12. A schematic of the proposed read-out circuit for the THz detector showing only the capacitances contacting the sense electrode. The capacitance values given were obtained from the simulation study results.

Having established the capacitive coupling of the read-out structure attention is now given to determining the voltage induced on the sense electrode due to a charge present in the quantum dots. This can be accomplished by inverting the above capacitance matrix for α to obtain the set of linear equations for V as follows

$$\begin{aligned} V_{1,elec} &= \alpha_{11}^{-1} Q_{1,elec} + \alpha_{12}^{-1} Q_{2,dot1} + \alpha_{13}^{-1} Q_{3,dot2} + \alpha_{14}^{-1} Q_{4,dot3} + \alpha_{15}^{-1} Q_{5,CNT} \\ V_{2,dot1} &= \alpha_{21}^{-1} Q_{1,elec} + \alpha_{22}^{-1} Q_{2,dot1} + \alpha_{23}^{-1} Q_{3,dot2} + \alpha_{24}^{-1} Q_{4,dot3} + \alpha_{25}^{-1} Q_{5,CNT} \\ &\dots \\ &\dots \end{aligned}$$

where the coefficients represented in matrix form are,

$$\alpha^{-1} = \begin{pmatrix} 1.62E16 & 1.18E16 & 1.06E16 & 9.17e15 & 1.25E16 \\ 1.18E16 & 1.53E17 & 6.51E16 & 3.70E16 & 9.59E15 \\ 1.06E16 & 6.52E16 & 1.60E17 & 7.43E16 & 8.82E15 \\ 9.10E15 & 3.64E16 & 7.27E16 & 1.71E17 & 7.77E15 \\ 1.01E16 & 8.62E15 & 8.46E15 & 7.84E15 & 2.91E16 \end{pmatrix} \begin{matrix} elec(aF^{-1}) \\ dot1 \\ dot2 \\ dot3 \\ CNT \end{matrix}$$

From these values one can readily obtain the potential induced on the sense electrode for our two cases, i.e. (1) when an QDC is in the primed state with an electron resides in QD2, or (2) when a THz photon excites the QD2 electron causing it to be ejected from the QDC. The corresponding voltage that would exist on the sense electrode if QD2 had a charge $Q_{3,dot2} = e$ with all other structures having $Q = 0$ is,

$$\begin{aligned} V_{1,elec}(Q_{2,dot1} = n \cdot e) &= \alpha_{11}^{-1} Q_{1,elec} + \alpha_{12}^{-1} Q_{2,dot1} + \alpha_{13}^{-1} Q_{3,dot2} + \alpha_{14}^{-1} Q_{4,dot3} + \alpha_{15}^{-1} Q_{5,CNT} \\ &= 0 + 0 + (1.06E16)(n \cdot e) + 0 + 0 \\ &= 1700 \mu V \cdot n \end{aligned}$$

For case (2), when the electron is ejected from the QDC the potential on the sense electrode is 0 V. This yields a $\Delta V = 1.7$ mV between the two cases. At this point, it is possible to determine how much the impedance of the SET will change as a result of a change in the gate voltage of $\Delta V = 1.7$ mV. The capacitance between the CNT-SET and sense electrode has been approximated to be $C_{51} = 40$ aF which means the distance between conductance peaks of the SET would be $\Delta V_g = e/C_{51} = 4mV$. Consequently, it is possible to derive the useful approximation,

$$\begin{aligned} \Delta G &= \frac{\partial G}{\partial V_g} \Delta V_g \\ &\approx \frac{G_{max}}{2mV} \cdot 1700 \mu V \\ &\approx 0.85 G_{max} \end{aligned}$$

If one assumes $G_{\max} = 10 \mu\text{S}$ (corresponding to $R = 100 \text{ k}\Omega$) and if the RF-SET is biased such that one of the double quantum-dot states gives a SET-conductance of $G_1 = G_{\max}$ then the other state would be $G_2 = G_{\max} - \Delta G = 1.5 \mu\text{S}$. In terms of the impedance seen by the matching circuit this would be comparable to distinguishing $100 \text{ k}\Omega$ from a $670 \text{ k}\Omega$ resistor. This is entirely feasible for achieving a successful readout with the available high signal-to-noise as will be discussed next in the impedance-matching section.

3.6 SET-Readout Tank Circuit Design

Fundamental to the RF-SET readout circuitry is the accompanying LC tank circuit which will transform the relatively high impedance of the carbon nanotubes RF-SET down to the characteristic impedance of 50Ω . In addition, when multiple THz detectors are integrated together, multiplexing of their signals can be accomplished by choosing unique LC tank-circuit resonant frequencies for each THz detector cell. The tank circuit shown in Fig. 13(a) consists of the parasitic capacitance to ground and an inductor ($C_{\text{parasitic}}$ is typically on the order of a fraction of a pF). Although the actual parasitic capacitance will vary between samples, for illustrative purposes we will assume $C_{\text{parasitic}} = 0.18 \text{ pF}$ and a $L_{\text{tank}} = 710 \text{ nH}$ chip inductor with a resistance at resonance of $R_{L\text{-Tank}} = 10 \Omega$. For this illustrative case, the S_{11} results have been simulated using the software ADS and as the RF-SET resistance changes from $R_{\text{RF-SET}} = 100 \text{ k}\Omega$ to $130 \text{ k}\Omega$ the S_{11} will change by over

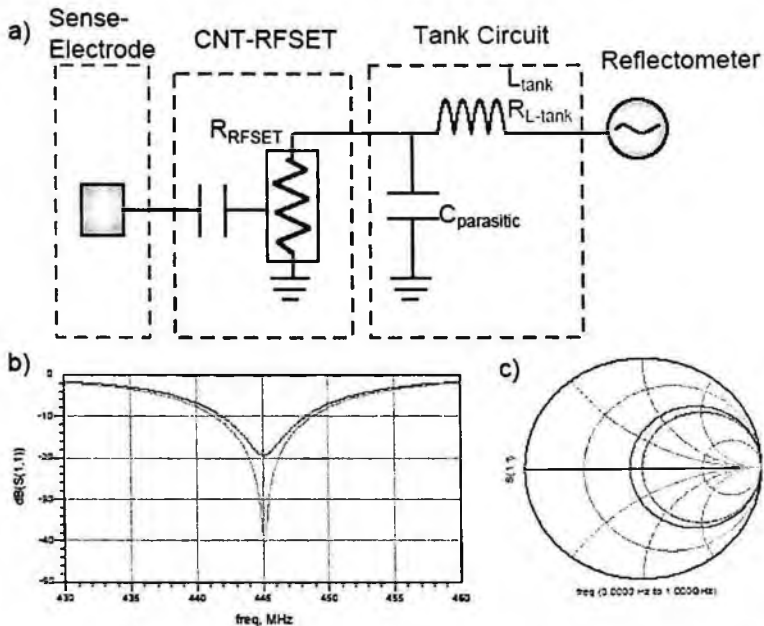


Fig. 13. a) A schematic of the LC-tank circuit used to transform the high impedance of the RF-SET down to the characteristic impedance of $Z = 50 \Omega$. b) Using the simulation software ADS, S_{11} is plotted for the case when $R_{\text{RF-SET}} = 100 \text{ k}\Omega$ (brown) and when it is $130 \text{ k}\Omega$ (green) along with the corresponding Smith chart. (c). The simulated component values of the tank circuit were $C_{\text{parasitic}} = 0.18 \text{ pF}$, $L_{\text{tank}} = 710 \text{ nH}$, and the resistance of the inductor at its resonant frequency $R_{L\text{-Tank}} = 10 \Omega$.

20 dB, as shown in Fig. 13(b). In practice, though one should expect a lower ΔS_{11} value due to the difficulty of precisely matching the RF CNT impedance to 50Ω due in part to 1) the constraint of limited bandwidth at which the cryogenic RF amplifiers maintain optimal bandwidth, 2) the limited availability of specific valued chip-inductors, and 3) variation in the parasitic capacitance between samples. Preliminary works has also been conducted to build and test an RF tank circuit with $L_{\text{tank}} = 10 \text{ nH}$ and $C_{\text{parasitic}} \sim 7 \text{ pF}$ and to demonstrate it at room temperature on chip resistors of various values, from 100Ω to $100 \text{ k}\Omega$. The results are given in Fig. 14. This demonstration circuit shows the ability to match to a $100 \text{ k}\Omega$ load and confirms the feasibility of the SET-based readout approach.

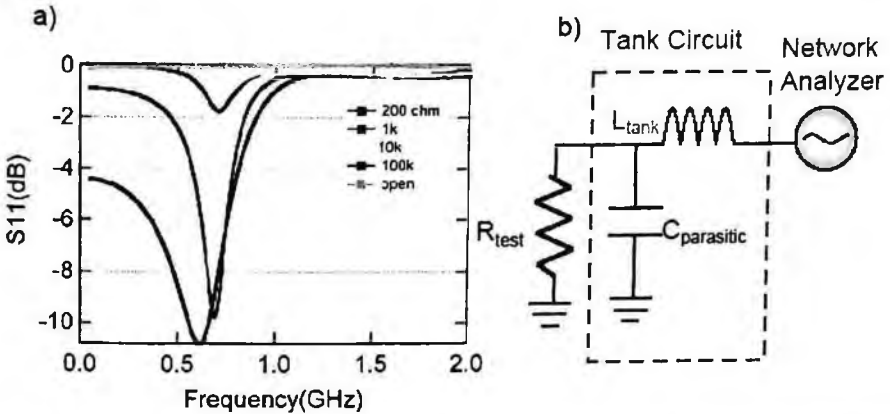


Fig. 14. A RF tank circuit (as illustrated in (b)) with $L_{\text{tank}} = 10 \text{ nH}$ and $C_{\text{parasitic}} \sim 7 \text{ pF}$ was tested to demonstrate how an impedance matching circuit would transform various resistance values (as given in (a)).

3.7 An Alternative All-Nanotube Integrated Design Concept

While the proposed concept design for the THz detector is feasible, it is also possible to suggest alternative designs that might offer advantages. As a potentially viable alternative to designing a vertically-defined heterostructure-semiconductor Q-dot system (e.g., InGaAs/AlGaAs triple Q-dot detector with optical injection), one could conceivably obtain the same result using a laterally-defined single-carbon-nanotube with multiple tops-gates as shown in Fig. 15. Compared to traditionally designed quantum dots (lateral gating of a 2DEG, vertical pillars), nanotube dots are by far the smallest and will allow for the simplest route to nanoscale spatial resolution. In addition, lateral gating of 2DEGs requires many electrodes to define each dot whereas a nanotube needs significantly fewer electrodes. Since the nanotube diameter is so small, transverse energy levels are “frozen out”, and the material is inherently one dimensional, simplifying the quantum excitation spectrum considerably. In addition, the nanotube small size causes the excitation energy scales in nanotube dots to fall naturally in the THz regime, whereas in lateral dots and the vertical design, they can be typically 10-100 GHz. Finally, a very important advantage of using nanotubes to define both the detection Q-dots and the readout circuitry is that one can easily *integrate* the entire technology onto one chip.

CNT Double Quantum-Dot

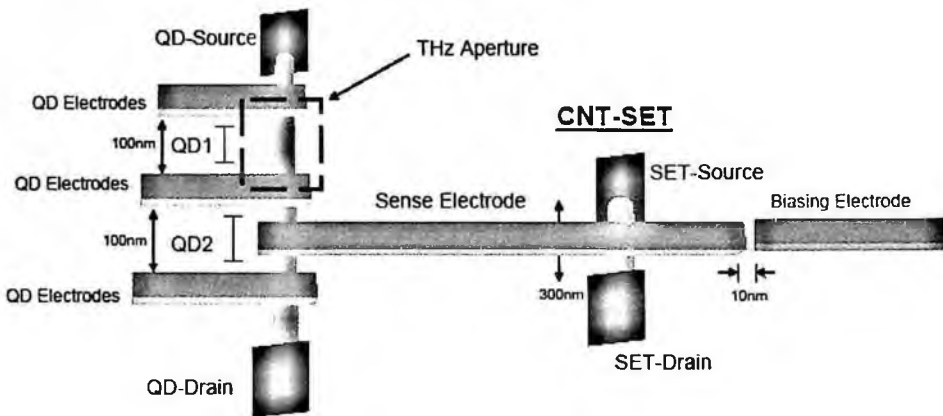


Fig. 15. This alternative structure for the THz detector and read-out circuit functions in the same manner as the one defined in Fig. 10 but utilizes QDs that are formed by nanotubes and potential electrodes. This structure offers fabrication advantages because only nanotubes and metallic gates would be involved.

4. Conclusions and Future Goals

This paper has presented a novel nanoscale-engineering methodology along with descriptive nanotechnology devices needed for the first-time development of a microscope-system capable of collecting terahertz (THz) frequency spectroscopic signatures from microscopic biological (bio) structures. The proposed THz microscope will represent a first of its kind (i.e., ability to perform THz spectral interrogation of microscopic targets) and it is proposed to be built upon completely new nanoelectronic concepts. The specific technological advantages include: (i) novel QCD design that allows for sampling the influence of THz-frequency photons on single-electron polarization which allows one to infer the intensity of the THz radiation over sub-wavelength apertures; and (ii) revolutionary CNT RF-SET technology that allows for ultrafast sampling of single-electron motion and the rapid readout of the information associated with a nanoscale THz sensing pixel; and when these technologies are implemented into a THz microscopy system they should allow for a unique and unprecedented long-wavelength spectral imaging capability.

The envisioned technology targets the goal of achieving the type of microscopy illustrated in the scenario of Fig. 16. Specifically, a scanning pixel array (or minimally a single pixel probe) is desired that will be effective for shadow-imaging biological targets with pixel resolution that is small in comparison to the structural features of the object. Here, the target object is assumed to be illuminated from below by THz radiation that is variable both in intensity and signal frequency. As hypothetically illustrated, expected bio-targets might be single cells or viruses of micron-size and below, so it will be necessary to define sensor elements capable of measuring the transmission of THz-frequency photons through materials/agents that possess geometrical structure on the nanoscale. Therefore, the aperture of the sensor must be on the order of 10-100's of nanometers while still being able to sense across a reasonable THz bandwidth with very

high detection sensitivity. This type of THz microscopy capability is attractive to enable the contrasting of structural features within the target that might exhibit a strong dependence on spectral absorption as illustrated in the top of Fig 16. For example, the circular green sub-unit within the cell (see optical visual image on the far left) could be transparent at THz frequency F_1 , but demonstrate a condition of opacity at THz frequency F_2 , and it possesses an inner light-green figure-eight feature that might absorb strongly only at a third THz frequency F_3 . While the exact scenario depicted in Fig. 16 is hypothetical, THz spectral signatures of the type assumed have been demonstrated in nearly perfect bulk-samples of bio-molecules and bio-cellular materials¹⁻⁴. Hence, a THz microscope of the type described would have profound implications for scientific inquiry and such a sensing capability would be valuable for practical applications related to medical diagnostics and bio-threat agent defense and security. Future work will focus on the further investigation and refinement of the technology presented here, and the pursuit of initial technology demonstrations will be conducted through a teamed multidisciplinary effort.

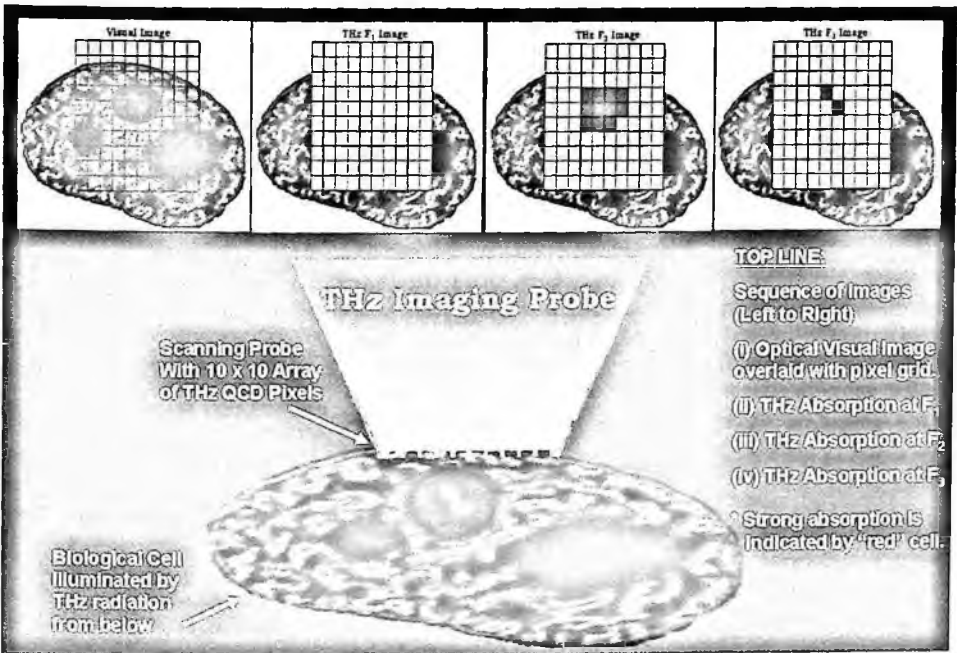


Fig. 16. Illustration of the operation scenario of the suggested THz-frequency transmission microscope.

Acknowledgment – This work was conducted under the support of the U.S. Defense Threat Reduction Agency (DTRA) as a part of the program led by Dr. Ngai Wong.

References

1. Terahertz Sensing Technology, Volume 2: Emerging Scientific Applications & Novel Device Concepts, Editors D. Woolard, W. Loerop & M. Shur (World Scientific, Singapore, 2003).
2. Dwight Woolard, et. al., "Terahertz Frequency Sensing and Imaging: A Time for Reckoning Future Applications?" *Proceedings of the IEEE*, **93**, pp. 1722-1743 (2005).
3. E. R. Brown, et. al., "THz-Regime Attenuation Signatures in Bacillus subtilis and a Model based on Surface Polarization Effects," accepted to *IEEE Sensors Journal* (2006).
4. T. Globus, M. Bykhovskaia, D. Woolard and Boris Gelmont, "Submillimeter-wave Absorption Spectra of Artificial RNA Molecules," *Journal of Physics D*, **36**, 1314-1322 (2003).
5. Peiji Zhao and Dwight Woolard, "Suppression of Electron-Phonon Scattering in Double-Quantum-Dot Based Quantum Gates," *Appl. Phys. Lett.*, **90**, 093507 (2007).
6. W. Lu, Z. Q. Ji, L. Pfeiffer, K. W. West and A. J. Rimberg, "Real-Time Detection of Electron Tunnelling in a Quantum Dot", *Nature*, **423**, 422-425, (2003); H. Hashiba, et. al., "Isolated Quantum Dot in Application to Terahertz Photon Counting," *Physical Review B*, **73**, 081310-1 (2006).
7. T. R. Stevenson, F. A. Pellerano, C. M. Stahle, K. Aidala and R. J. Schoelkopf, "Multiplexing of Radio-Frequency Single-Electron Transistors", *Applied Physics Letters*, **80**, 3012-3014, (2002).
8. R. J. Schoelkopf, P. Wahlgren, A. A. Kozhevnikov, P. Delsing and D. E. Prober, "The Radio-Frequency Single-Electron Transistor (Rf-SET): A Fast and Ultrasensitive Electrometer", *Science*, **280**, 1238-1242, (1998).

PREDICTING THE PATH OF ELECTRONIC TRANSPORT THROUGH A MOLECULAR DEVICE VIA A MOUNTAIN-PASS ALGORITHM

GREG RECINE

*Dept of Physics and Engineering Physics, Stevens Institute of Technology
Hoboken, NJ 07030, USA*

DWIGHT L. WOOLARD

*Army Research Office, OTP
Raleigh, NC 27695, USA*

The so-called "mountain-pass" theorem allows for the finding of a critical point on the path between two points on a multidimensional contour where the maximal elevation is minimal. By implementing the "elastic string algorithm", it is possible to not only find the critical point, but to compute the mountain-pass itself on a finite-dimensional contour. For a given molecule that sits between two probes making up a nanostructure device, we propose that the mountain-pass will be a likely path of electron transport through the molecule, the contour being the electronic potential of the molecule. This potential along this path will be used as the input potential for SETraNS, a 1D Wigner-Poisson electron transport solver in order to explore the current-bias characteristics of such the molecule in such a device. In order to calculate the mountain pass, the elastic string algorithm is used to set up a constrained non-linear optimization problem which is, in turn, solved via a Monte Carlo method. We will compute the mountain-pass for a well-known test contour in order to show the validity of this approach. The procedure developed here is to be combined with conformational analysis via the molecular modeling program AMBER and the quantum transport program SETraNS in order to predict molecular function. When achieved, this combined procedure will allow for the better design and implementation of nanoscale molecular devices for application such as sensing, switching and data processing.

1. Introduction

For a given molecular device, there should exist a preferred low-energy pathway that electrons will take in transversing the molecule from one fixed point to another. When considering this problem as a basic transport problem, the pathway can be represented by a potential function felt by the charge carrier that exists between the emitter and collector. This problem is analogous to finding the easiest path that passes though mountainous terrain (fig 1). Finding this so-called "Mountain-Pass"

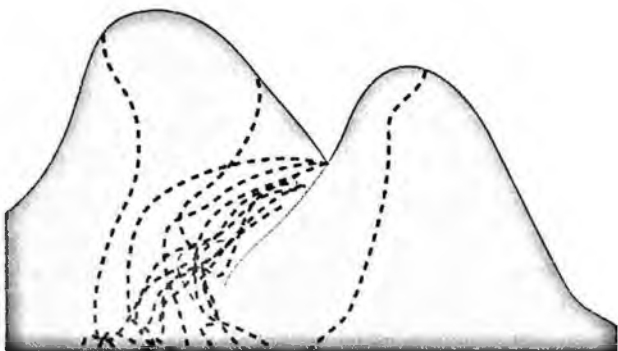


Fig. 1. Cartoon mountain pass symbolizing the path of least resistance

is a mathematical problem that has been greatly examined over the past 30 years since the introduction of the mountain pass theorem in 1973^{1,2}.

The mountain-pass theorem describes how to find the critical point x^* in a path between two points, x_a and x_b of a given n -dimensional function, f . The critical point being that point which must lie on the best pathway between two points, x_a and x_b . Moré and Munson³ describes the mountain-pass theorem in it's most basic form, stating that the mountain-pass theorem shows that if $x_a \in \mathbb{R}^n \mapsto \mathbb{R}$, and f has no critical points at infinity, then the critical point, x^* , in the path between x_a and $x_b \in \mathbb{R}^n$ with $f(x_b) \leq f(x_a)$, is such that the value of f at x^* is given by

$$\gamma = \inf_{p \in \Gamma} \{ \max \{ f[p(t)] : t \in [0, 1] \} \}, \quad (1)$$

Γ being the set of all paths that connect x_a to x_b . In two dimensions, this can be simply described as finding the maximum elevation on every path between x_a and x_b , then the minimum of this set of maximums will be the critical point. One example of an application of the mountain-pass theorem is in computational chemistry. Finding the critical point can show the lowest energy required for a molecule to make a transition between two stable states.

2. Computing Mountain Passes

In order to efficiently compute a mountain pass, the "elastic string algorithm" of Moré and Munson³ is employed. Briefly stated, the elastic string algorithm takes the possible paths between two points, x_a and x_b and approximates each one with a piecewise linear path with m breakpoints. Equation 1 now takes the form

$$\gamma = \min_{p \in \Gamma_\pi} \{ \max \{ f[p(t_k)] : 1 \leq k \leq m \} \} \quad (2)$$

where $\pi = \{t_0, \dots, t_{m+1}\}$ with $t_0 = 0$ and $t_{m+1} = 1$. The piecewise linear paths that connect x_a to x_b are defined as:

$$PL[0, 1] = \{p \in C[0, 1] : p(0) = x_a, p(1) = x_b; p \text{ linear on } (t_k, t_{k+1}), 0 \leq k \leq m\}. \quad (3)$$

Bounding each of the paths by a given length, L , defines

$$\Gamma_\pi = \left\{ p \in PL[0, 1] : \int_0^1 \|p'(t)\| dt \leq L \right\}, \quad (4)$$

where the breakpoints $x_k = p(t_k)$ are now:

$$p(t) = \frac{t - t_k}{t_{k+1} - t_k} x_{k+1} + \frac{t_{k+1} - t}{t_{k+1} - t_k} x_k; t \in [t_k, t_{k+1}]; \|p'(t)\| = \frac{x_{k+1} - x_k}{t_{k+1} - t_k}. \quad (5)$$

It can be shown that the length L will bound p if and only if the following is true:

$$\|x_{k+1} - x_k\| \leq h_k; 0 \leq k \leq m; \text{ where } \sum_{k=0}^m h_k = L. \quad (6)$$

The version of the mountain-pass theorem that is given in equation 2 is an optimization problem which can be solved by finding the breakpoints via another optimization problem, formulated by

$$\min \{ \nu(x) : \|x_{k+1} - x_k\| \leq h_k; 0 \leq k \leq m \} \quad (7)$$

where $\nu(x) = \max \{ f(x_1), \dots, f(x_m) \}$.

3. Computational Method

As a test of the computational methods, we used a six-hump ‘‘camel-back’’ function as the terrain on which we will find a mountain pass. The function is used to test many global optimization algorithm and has well known mountain passes. The function takes the form

$$f(\xi_1, \xi_2) = \left(4 - 2.1\xi_1^2 + \frac{1}{3}\xi_1^4 \right) \xi_1^2 + \xi_1 \xi_2 + 4(\xi_2^2 - 1)\xi_2^2, \quad (8)$$

and is shown in figure 2. We will define our starting and end points as $x_a = (-1.5, -0.6)$ and $x_b = (0.0, 0.8)$ as to match known results³.

Our end goal is to use an electrostatic potential of a molecule calculated via a molecular modeling codes, such as Amber⁴. The nature of this problem allows us to simplify the computational process. First, for a given molecule, the approximate path is already known and will be used as a starting guess for the algorithm that is very close to the solution. Also, only a rough mountain pass will be required since the potential will be a slowly varying, well behaved function. Another valuable insight is that, for the most part, the electron will be confined to transport in a plane, simplifying the problem even more. A Monte Carlo method was chosen as quick, simple way to explore the nearby solution space, in accordance with the constraints imposed by the elastic string algorithm and the simplifications just mentioned.

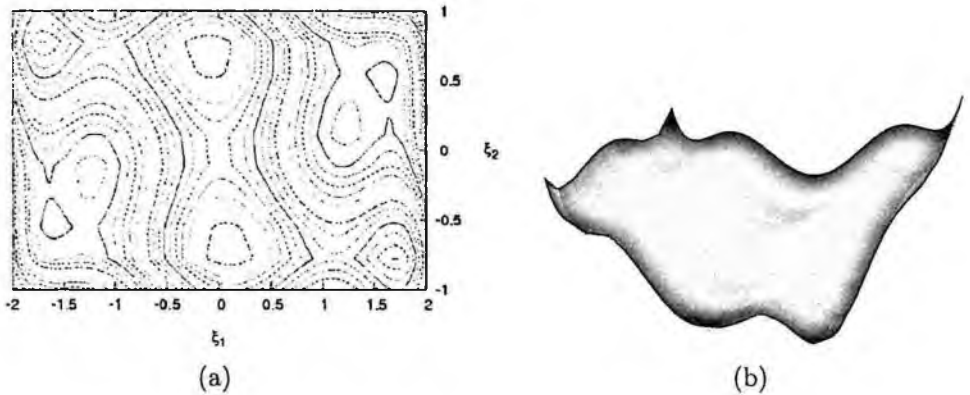


Fig. 2. Six-hump camel back function (a) contour map, (b) 3D topology

An interesting property of the elastic string algorithm is that even by choosing the number of breakpoints fairly small, say at $m = 20$, a very good approximation for the mountain pass is obtained. The maximum length of each segment, h_K , are all set to the same value, h , which is chosen to be

$$h = \frac{2}{m+1} \|x_b - x_a\|. \quad (9)$$

For a reasonably finite grid of data, we can compute most of (and in some cases, exactly all) the breakpoints of all the possible paths from x_a to x_b that satisfy the constraint 9. This is accomplished by guessing an initial mountain pass (a straight line between x_a to x_b , for example), then randomly modifying the break-point locations while satisfying the constraints in order to build up a series of pathways. It is then trivial to find the minimum maximum break-point value of all paths, the critical point x^* . We then know that this critical point is one that must lie on the mountain pass. We then use this critical point as our new endpoint and find the critical point between x_a and x^* . By repeating this process, we obtain the end result of an optimized set of breakpoints which the mountain pass must lie on. These m points represent our piecewise-linear low-energy mountain pass between the original endpoints x_a and x_b .

4. Results

The calculated mountain passes and the known solution for the test case are shown in figure 3. Eight different runs are shown, and the mountain passes they show are all quite similar. Figure 4 shows the calculated potential along the mountain passes and the potential along the known solution. Once again, they match up rather well. The result for one run was obtained in approximately 5 minutes on a 3GHz Xeon MP CPU. The procedure has yet to be optimized or properly parallelized, so faster

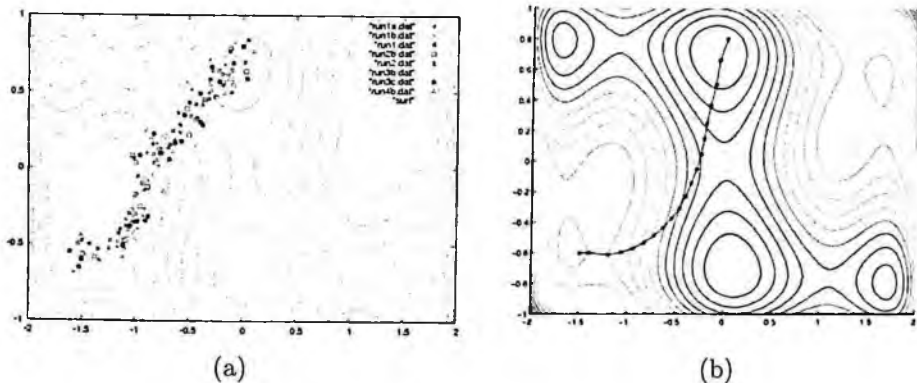


Fig. 3. Computed Mountain Pass: (a) Brute Force Method, (b) Known Solution

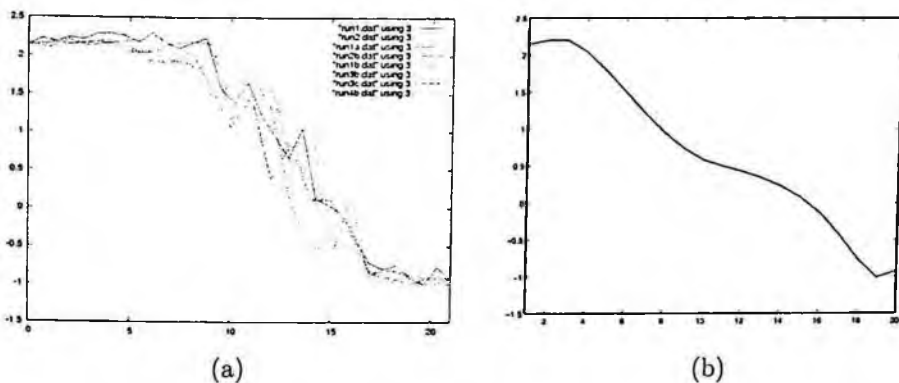


Fig. 4. Computed Mountain Pass Potential Profile: (a) Brute Force Method, (b) Known Solution

run times are expected. The procedure can be further refined by choosing more breakpoints, and tightening the constraints. This will lead to a longer run time, however. However, the mountain passes calculated by this method are close enough to the correct solution that the breakpoints can now be minimized further by using a Newton's method. It is import to note here that other, more general, implementations of the elastic string algorithm³ do just this: first generate an approximate solution, then use Newton's method to obtain a more accurate solution.

5. Conclusions

We have shown a fast, simple method that implements the elastic string algorithm to solve the Mountain Pass problem for predicting the path of a charge carrier through a non-trivial potential contour. Due to the nature of the problem, we were able to take advantage of some simplifications (good initial guess, slowly varying

contour, confined transport) in order to use a Monte Carlo technique to explore the solution space in accordance with the constraints imposed by the elastic string algorithm. The results obtained are rough, but are suitable for the problem at hand. The results are also close enough to the real solution to be used as a starting point for further minimization via Newton's method, if needed.

Acknowledgments

The authors would like to thank Prof. C. Tim Kelley for his fruitful discussions regarding the mathematics of the Mountain-Pass algorithm, and Prof. H.L. Cui for providing computational resources.

References

1. A. Ambrosetti and P. H. Rabinowitz, "Dual variational methods in critical point theory and applications," *J. Funct. Anal.*, vol. 14, pp. 349–381, 1973.
2. Y. S. Choi and P. J. McKenna, "A mountain pass method for the numerical solution of semilinear elliptic problems," *Nonlinear Analysis, Theory, Methods & Applications*, vol. 20, no. 4, pp. 417–437, 1993.
3. J. J. M. T. S. Munson, "Computing mountain passes," Argonne National Lab Preprint, Mathematics and Computer Science Division, ANL, Argonne IL, Tech. Rep. ANL/MCS-P957-0502, May 2002.
4. D. Case, T. Darden, I. T.E. Cheatham, C. Simmerling, J. Wang, R. Duke, R. Luo, K. Merz, B. Wang, D. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J. Caldwell, W. Ross, and P. Kollman, "Amber8," University of California, San Francisco., 2004.

SPECTRAL SENSING RESEARCH FOR WATER MONITORING APPLICATIONS AND FRONTIER SCIENCE AND TECHNOLOGY FOR CHEMICAL, BIOLOGICAL AND RADIOLOGICAL DEFENSE

This book provides unique perspectives on both state-of-the-art hyperspectral techniques for the early-warning monitoring of water supplies against chemical, biological and radiological (CBR) contamination effects as well as the emerging spectroscopic science and technology base that will be used to support an array of CBR defense and security applications in the future. The technical content in this book lends itself to the non-traditional requirements for point and stand-off detection that have evolved out of the US joint services programs over many years. In particular, the scientific and technological work presented seeks to enable hyperspectral-based sensing and monitoring that is real-time; in-line; low in cost and labor; and easy to support, maintain and use in military- and security-relevant scenarios.

World Scientific

www.worldscientific.com

6949 hc

ISBN-13 978-981-283-323-5
ISBN-10 981-283-323-4



9 789812 833235