Applied and Computational Mathematics Division

Summary of Activities for Fiscal Year 2021

NIST
National Institute of
Standards and Technology
U.S. Department of Commerce

Applied and Computational Mathematics Division

Summary of Activities for Fiscal Year 2021

Ronald F. Boisvert, Editor
Applied and Computational Mathematics Division
Information Technology Laboratory

April 2022

Abstract

This report summarizes recent technical work of the Applied and Computational Sciences Division of the Information Technology Laboratory at the National Institute of Standards and Technology (NIST). Part I (Overview) provides a high-level overview of the Division's activities, including highlights of technical accomplishments during the previous year. Part II (Features) provides further details on projects of particular note this year. This is followed in Part III (Project Summaries) by brief synopses of all technical projects active during the past year. Part IV (Activity Data) provides listings of publications, technical talks, and other professional activities in which Division staff members have participated. The reporting period covered by this document is October 2020 through December 2021.

For further information, contact Ronald F. Boisvert, 100 Bureau Drive, Mail Stop 8910, NIST, Gaithersburg, MD 20899-8910, phone 301-975-3812, email boisvert@nist.gov, or see the Division's Web site at https://www.nist.gov/itl/math/.

Keywords: applied mathematics; computational science and engineering; high-performance computing; mathematics of metrology; mathematics of biotechnology; materials modeling and simulation; mathematical knowledge management; mathematical modeling; network science; scientific visualization; quantum information science.

Cover Visualization: A snapshot of a simulation of the NIST monoclonal antibody (mAb) Standard Reference Material (SRM 8671) in suspension. The individual spheres are the amino acids that make-up each mAb. See page 82.

Section Visualizations: The "word cloud," which is found at the start of each Part of this document was created using Wordle, and the text of this document as input.

Disclaimer: Certain commercial entities, equipment, and materials are identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, and equipment are necessarily the best available for the purpose.

Contents

Part I

Overview

Introduction

Founded in 1901, the National Institute of Standards and Technology (NIST) is a non-regulatory federal agency within the U.S. Department of Commerce. Its mission is to promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life. The technical disciplines represented in the NIST Laboratories include physics, electrical engineering, nanotechnology, materials science, chemistry, biotechnology, manufacturing and construction engineering, fire research, information technology, mathematics, and statistics. The NIST Labs operate in two locations: Gaithersburg, MD, (headquarters—234 hectare/578-acre campus) and Boulder, CO (84 hectare/208-acre campus). NIST employs about 3 400 scientists, engineers, technicians, and support personnel. NIST also hosts about 3 500 associates from academia, industry, and other government agencies, who collaborate with NIST staff and access its facilities.

The Information Technology Laboratory (ITL) is one of six major organizational units that make up the NIST Labs. ITL's singular purpose is to cultivate trust in information technology and metrology. This is done through the development of measurements, tests, and guidance to support innovation in and deployment of information technology by industry and government, as well as through the application of advanced mathematics, statistics, and computer science to help ensure the quality of measurement science.

The Applied and Computational Mathematics Division (ACMD) is one of seven technical Divisions in ITL. At its core, ACMD's purpose is to nurture trust in metrology and scientific computing. To do so, ACMD provides leadership within NIST in the use of applied and computational mathematics to solve technical problems arising in measurement science and related applications. In that role staff members

- perform research in applied mathematics and computational science and engineering, including analytical and numerical methods, high-performance computing, and visualization,
- perform applied research in computer science and engineering for future computing and communications technologies,
- engage in peer-to-peer collaborations to apply mathematical and computational techniques and tools to NIST problems,
- develop and disseminate mathematical reference data, software, and related tools, and
- work with internal and external groups to develop standards, tests, reference implementations, and other measurement technologies for current and future scientific computing systems.

Division staff is organized into four groups:

- <u>Mathematical Analysis and Modeling Group</u> (*Timothy Burns, Leader*). Performs research and maintains expertise in applied mathematics, mathematical modeling, and numerical analysis for application to measurement science.
- <u>Mathematical Software Group</u> (*Michael Donahue, Leader*). Performs research and maintains expertise in the methodology and application of mathematical algorithms and software in support of computational science within NIST as well as in industry and academia.
- <u>High Performance Computing and Visualization Group</u> (*Judith Terrill, Leader*). Performs research and maintains expertise in the methodologies and tools of high-performance scientific computing and visualization for use in measurement science.
- <u>Computing and Communications Theory Group</u> (*Ronald Boisvert, Acting Leader; Oliver Slattery, Project Leader*). Performs research and maintains expertise in the fundamental mathematics, physics, computer science, and measurement science necessary to enable the development and analysis of current and future computing and communications systems.

The technical work of the Division is organized into seven thematic areas; these are described in the sidebar. Project descriptions in Part III of this document are organized according to these broad themes.

Division Thematic Areas

Broad Areas

Mathematics of Metrology. Mathematics plays an important role in measurement science. Mathematical models are needed to understand how to design effective measurement systems and to analyze the results they produce. Mathematical techniques are used to develop and analyze idealized models of physical phenomena to be measured, and mathematical algorithms are necessary to find optimal system parameters. Mathematical and statistical techniques are needed to transform measured data into useful information. We develop fundamental mathematical methods and tools necessary for NIST to remain a world-class metrology institute, and to apply these to measurement science problems.

High Performance Computing and Visualization. Computational capability continues to advance rapidly, enabling modeling and simulation to be done with greatly increased fidelity. Doing so often requires computing resources well beyond what is available on the desktop. Developing software that makes effective use of such high-performance computing platforms remains very challenging, requiring expertise that application scientists rarely have. We maintain such expertise for application to NIST problems. Such computations, as well as modern experiments, typically produce large volumes of data, which cannot be readily comprehended. We are developing the infrastructure necessary for advanced interactive, quantitative visualization and analysis of scientific data, including the use of 3D immersive environments, and applying the resulting tools to NIST problems.

Current Focus Areas

Mathematics of Biotechnology. As proof-of-concept academic work in engineering biology meets the market realities of bringing lab science to product initiation, there are needs to compare biological products, measure whether desired outcomes are realized, and optimize biological systems for desired behaviors. NIST is working to deliver tools and standards to measure such biological technologies, outputs, and processes from healthcare to manufacturing and beyond. We support this effort with the development and deployment of innovative mathematical modeling and data analysis techniques and tools.

Materials Modeling. Mathematical modeling, computational simulation, and data analytics are key enablers of emerging manufacturing technologies. The Materials Genome Initiative (MGI), an interagency program with the goal of significantly reducing the time from discovery to commercial deployment of new materials using modeling, simulation, and informatics, is a case in point. To support the NIST role in the MGI, we develop and assess modeling and simulation techniques and tools, with emphasis on uncertainty quantification, and collaborate with other NIST Laboratories in their efforts to develop the measurement science infrastructure needed by the materials science and engineering community.

Quantum Information Science. An emerging discipline at the intersection of physics and computer science, quantum information science is likely to revolutionize 21st century science and technology in the same way that lasers, electronics, and computers did in the 20th century. By encoding information into quantum states of matter, one can, in theory, enable phenomenal increases in information storage and processing capability. At the same time, such computers would threaten the public-key infrastructure that secures all of electronic commerce. Although many of the necessary physical manipulations of quantum states have been demonstrated experimentally, scaling these up to enable fully capable quantum computers and networks remains a grand challenge. We engage in (a) theoretical studies to understand the power of quantum computing, (b) collaborative efforts with the multi-laboratory experimental quantum science program at NIST to characterize and benchmark specific physical realizations of quantum information processing, and (c) demonstration and assessment of technologies for quantum networking.

Foundations of Measurement Science for Information Systems. ITL assumes primary responsibility within NIST for the development of measurement science infrastructure and related standards for IT and its applications. ACMD develops the mathematical foundations for such work. This can be very challenging. For example, many large-scale information-centric systems can be characterized as an interconnection of many independently operating components (e.g., software systems, communication networks, the power grid, transportation systems, financial systems). Exactly how the structure of such large-scale interconnected systems and the local dynamics of its components leads to system-level behavior is only weakly understood. This inability to predict the systemic risk inherent in system design leaves us open to unrealized potential to improve systems or to avoid potentially devastating failures. A looming new example of importance to NIST is the Internet of Things. We are developing models to aid in the development of applications from individualized health IT devices to large-scale sensor networks.

Mathematical Knowledge Management. We work with researchers in academia and industry to develop technologies, tools, and standards for representation, exchange, and use of mathematical data. Of particular concern are semantic-based representations which can provide the basis for interoperability of mathematical information processing systems. We apply these representations to the development and dissemination of reference data for applied mathematics. The centerpiece of this effort is the Digital Library of Mathematical Functions, a freely available interactive and richly linked online resource, providing essential information on the properties of the special functions of applied mathematics, the foundation of mathematical modeling in all of science and engineering.

Highlights

In this section we identify some of the major accomplishments of the Division during the past year. We also provide news related to ACMD staff. For the second year in a row, the COVID-19 pandemic provided a wealth of challenges to overcome during 2021. The NIST campus remained closed to all Division staff except those whose work required laboratory facilities and equipment only accessible while on campus. The daily face-to-face interactions which are the lifeblood of a research organization moved online, which made them difficult at best. The spontaneity that feeds collaboration continued to be largely lost. While staffing levels grew during the period, it continued to be difficult to integrate new staff members into the life of the Division. We managed to cope with these challenges and developed new methods of interaction. In fact, routine virtual interactions enabled staff working in Gaithersburg to know their colleagues in Boulder better, and vice-versa. As I write this in early 2022, most of our staff remain in full-time teleworking mode. We look forward to a return to normalcy later this calendar year.

Figure 1. *For the second year in a row, the COVID-19 pandemic kept most ACMD staff working from home, which made collaborations and the assimilation of new staff members very challenging. The corridor outside the Division office illustrates how lonely working on campus was during this period.*

Recent Technical Highlights

ACMD has made significant technical progress on many fronts during the past year. Here we highlight a few notable technical accomplishments. Further details are provided in Part II (Features) and Part III (Project Summaries).

- Division staff have had some remarkable success in the problem of identifying an initial state which evolves under the 2D viscous Burgers' equation into a close approximation of noisy data given at a future point in time. This *data assimilation problem* is ill-posed and hence extremely difficult. Using a direct non-iterative method based on marching backward in time with a stabilized explicit finite-difference scheme, Division staff have demonstrated a large class of examples in which reasonable recovery is possible. As one might expect in such a challenging problem, there are also unsuccessful examples. See page 47.

- ACMD staff members made a ground-breaking advance in techniques for *identifying substances using DART-MS* (Direct Analysis in Real Time Mass Spectrometry). One important application is screening for controlled substances in forensics for seized drug evidence. The team's insight was to utilize the statistical nature of the peaks in mass spectra in order to determine similarity. See page 19.

- Capitalizing on their success last year in using machine learning to successfully predict the Kováts retention index of substances, an ITL/MML team led by ACMD turned their attention to the *prediction of normal boiling points* of chemical compounds. Their machine-learning-based model was also successful here. See page 37.

- By applying Chebyshev expansions adaptively, Division staff have enabled more efficient and precise *representation of pure fluid phase equilibria*, which will provide fundamental improvement in the modeling routines offered by the NIST Reference Fluid Thermodynamic and Transport Properties Database. See page 71.

- The need for efficient and accurate antibody measurements has been an ongoing need during the COVID-19 crisis. This year, several staff members derived a series of new results showing that *diagnostic classification associated with antibody assays* can be recast as an optimization problem. Solutions to this problem result in improved classification accuracy over traditional methods, as has

been demonstrated in collaboration with colleagues at the School of Public Health at Johns Hopkins University. See page 53.

- We have been extending the capabilities of our dense-suspension simulator, QDPD, to match the physical experiments being performed at NIST's NCNR on the *flow of monoclonal antibody proteins* in solution. An early application of this new simulation capability, in collaboration with researchers at the University of Illinois and the University of Delaware, has been published in the journal *Structural Dynamics*. See page 82.

- Pulmonary edema is a medical condition caused by accumulation of excess fluid in the human lungs. It can occur due to heart problems, pneumonia, exposure to certain toxins, as well as viruses such as COVID-19. We have investigated the feasibility of a simple wearable system that can be used at home to *detect or monitor excess fluid buildup in the lungs*. The proposed wearable system is composed of several small wearable antennas located on the chest and back area. Through simulation studies, we have demonstrated that the frequency and time responses of the communication channel between these antennas can potentially be exploited to predict the onset of pulmonary edema. See page 15.

- Working with staff of the National Renewable Energy Laboratory (NREL), ACMD mathematicians have improved the search for *optimal airfoil shapes*, for, e.g., wind turbine blades, using a new, much richer set of data-driven deformations than designers had previously considered. See page 39.

- ACMD's High Performance Computing and Visualization Group has operated an *immersive visualization environment* (IVE) for many years and have successfully used it to provide critical insights to a wide variety of NIST metrology projects. To take advantage of recent advances in visualization hardware, we are moving our IVE to ParaView, a fully open-source software environment. As part of this effort, we are partnering with Kitware to extend and enhance ParaView's capabilities for immersive visualization. We have also joined the Khronos Group, which is working to develop standards for extended reality and for rendering scientific data. See page 93.

- Modern engineered systems, such as information and communication systems and power systems, are actually comprised of many sub-systems. Due to this interdependence, it is possible for a local failure or infection of a system by malware to spread to other systems, potentially compromising the integrity of the overall system. With this in mind, we studied the problem of *measuring the cyber risks* to nodes (e.g., subsystems) and finding optimal cybersecurity investments for hardening vulnerable nodes in large systems, which will minimize the long-run average costs to a system operator. See page 117.

- One difficulty with scaling-up trapped-ion quantum computers is the management of lasers used to apply logic gates. To overcome this difficulty, the NIST Ion Storage Group has developed techniques to apply all logic gates using only microwave fields. ACMD researchers designed a strategy for *estimating the fidelity with which such microwave gates could entangle qubits*. The observed fidelity rivaled the best fidelities achieved with lasers. This advance was reported in *Nature*. See page 99.

- In December 2021 *Physics World,* a publication of the Institute of Physics, announced their top 10 physics breakthroughs for 2021. The number one breakthrough was a project with significant participation by ACMD. A team led by John Teufel and Shlomi Kotler of PML was cited for *entangling two macroscopic vibrating drumheads*, "thereby advancing our understanding of the divide between quantum and classical systems." Ezad Shojaee, Alex Kwiatkowski, Shawn Geller, Scott Glancy, and Manny Knill of ACMD developed the theory and wrote, analyzed, and tested the data analysis code that enabled this accomplishment. They were all co-authors on the paper reporting the results published in Science in May[1], where a news summary stated the following:

---

[1] https://www.science.org/doi/10.1126/science.abf2998

> *"Quantum entanglement occurs when two separate entities become strongly linked in a way that cannot be explained by classical physics; it is a powerful resource in quantum communication protocols and advanced technologies that aim to exploit the enhanced capabilities of quantum systems. To date, entanglement has generally been limited to microscopic quantum units such as pairs or multiples of single ions, atoms, photons, and so on. Kotler et al. and Mercier de Lépinay et al. demonstrate the ability to extend quantum entanglement to massive macroscopic systems (see the Perspective by Lau and Clerk). Entanglement of two mechanical oscillators on such a large length and mass scale is expected to find widespread use in both applications and fundamental physics to probe the boundary between the classical and quantum worlds."*

See page 99.

- Working with colleagues in PML and at the University of Maryland, an ACMD researcher examined the class of *quantum adiabatic optimization algorithms* through the lens of optimal control theory. In this algorithm, the system is initialized in the ground state of some initial Hamiltonian. If the Hamiltonian changes slowly enough, the quantum adiabatic theorem says that the system will stay in the ground state, allowing one to find the ground state of desired Hamiltonians that might encode some classical problem. In their study they were able to identify an optimal procedure for evolving such a Hamiltonian. See page 104.

- For several years, ACMD staff have made remarkable progress in the automation of quantum dot control for quantum computing applications. (See page 24.) Recently, they turned their attention to the processing of data from cold atom experiments, whose data comes from images. In particular, they are working with colleagues in PML on *detecting dark solitons in Bose–Einstein condensates* (BEC). They have developed an automated classification and positioning system for identifying such structures using deep convolutional neural networks to eliminate the need for human image examination. Furthermore, they published a first-of-its kind labeled dataset of dark solitons for further machine learning research. See page 84.

- This year our quantum communications experimental team began several new efforts in the area of *quantum networking*. For example, they (a) designed and built tools that will be deployed into quantum networks for characterization of quantum network components and quantum network links, (b) began to build and synchronize nodes on a campus-wide quantum network testbed in collaboration with colleagues from ITL, PML and CTL, and (c) developed a plan for a regional quantum network testbed (called the DC Quantum Network testbed or DC-QNet) in collaboration with other NIST colleagues and in coordination with a number of local federal agencies. See page 109.

- Finally, ACMD staff played a key role in the analysis of data for two important studies of *gender, equity, and inclusion at NIST*. See pages 134 and 135.

Technology Transfer and Community Engagement-

The volume of technical output of ACMD remains high. During the last 15 months, Division staff members were (co-)authors of 48 articles appearing in peer-reviewed journals, 28 papers in conference proceedings, and 13 published in other venues. Twenty-one additional papers were accepted for publication, while 38 others are undergoing review. Division staff gave 59 invited technical talks and presented 34 others in conferences and workshops. Staff members were co-inventors on two newly awarded patents. Six new invention disclosures were filed, four of which have resulted in patent applications.

ACMD continues to maintain an active website with a variety of information and services, most notably the Digital Library of Mathematical Functions, though legacy services that are no longer actively developed, like the Guide to Available Mathematical Software, the Matrix Market, and the SciMark Java benchmark still see significant use. During calendar year (CY) 2021, the division web server satisfied more than 6 million requests for pages during more than 793 000 user visits. Another indication of the successful

transfer of our technology is references to our software in refereed journal articles. For example, our software system for nano-magnetic modeling (OOMMF) was cited in 203 such papers published in CY 2020 alone.

Members of the Division are also active in professional circles. Staff members hold a total of 14 editorial positions in peer-reviewed journals. For example, Barry Schneider is an Associate Editor-in-Chief for IEEE's *Computing in Science and Engineering.* Staff members are also active in conference organization, serving on 25 organizing/steering/program committees.

Service within professional societies is also prevalent among our staff. For example, Bonita Saunders is a member of the Board of Trustees of the Society for Industrial and Applied Mathematics (SIAM). Staff members are also active in a variety of working groups. Ronald Boisvert and Andrew Dienstfrey serve as members of the International Federation for Information Processing (IFIP) Working Group 2.5 on Numerical Software, Donald Porter is a member of the Tcl Core Team, Bruce Miller is a member of W3C's Math Working Group, and Sandy Ressler is a member of the W3C Advisory Committee. Barry Schneider represents NIST on the High-End Computing (HEC) Interagency Working Group of the Federal Networking and Information Technology Research and Development (NITRD) Program. Further details can be found in Part IV of this report.

Staff News

Once again, this year ACMD experienced an unusually large number of staffing changes. Three permanent staff members retired. At the same time, we welcomed four new permanent staff members and four new postdocs. All of our new staff members were onboarded while the NIST campus was closed during the pandemic. This has made the assimilation of new staff into Division research efforts more difficult than usual. Our senior staff members have been diligent in working to give our new staff members and postdocs the best experience possible under the circumstances. We once again provided remote internship opportunities to a large number of students, including 16 graduate students, 10 undergraduate students, and 3 high school students. See Table 7 on page 139 for a list of our interns.

Further details on our staff changes and awards are provided below.

Arrivals

William Sherman joined ACMD as a Federal employee on July 7, 2021. He had been working under contract to NIST for the previous year. Bill received a B.S. in Computer Engineering (1986) and an M.S. in Computer Science (1989) both from the University of Illinois. An expert on virtual reality (VR), Bill is the author of four books on the topic. He has wide experience in scientific visualization having worked as a researcher at the National Center for Supercomputer Applications (15 years), the Desert Research Institute (5 years), and Indiana University (11 years). At NIST, Bill will work on the development and application of the NIST CAVE, as well as in standardization of VR software.

Elsie (Meliza) Lane joined ACMD/Boulder as an Administrative Assistant on August 16, 2021. Meliza spent 25 years working for the Colorado Dept. of Transportation and Colorado Parks and Wildlife, and 16 months at the Dept. of the Interior Bureau of Reclamation before joining NIST. She will be providing admin services to our Boulder staff as well as providing backup for admin work in Gaithersburg.

Lisa Ritter began an NRC Postdoctoral Associateship appointment with ACMD on September 13, 2021. She received a Ph.D. in Mathematics from the University of Hawaii at Manoa last spring where she did a thesis on superintegrable systems, their algebraic structures, and how to use them to find solutions to Painlevé equations. At NIST she will be working with her ACMD advisor Howard Cohl in the search for new solutions to the Painlevé equations. She will also contribute to the Digital Repository of Mathematical Formulae effort.
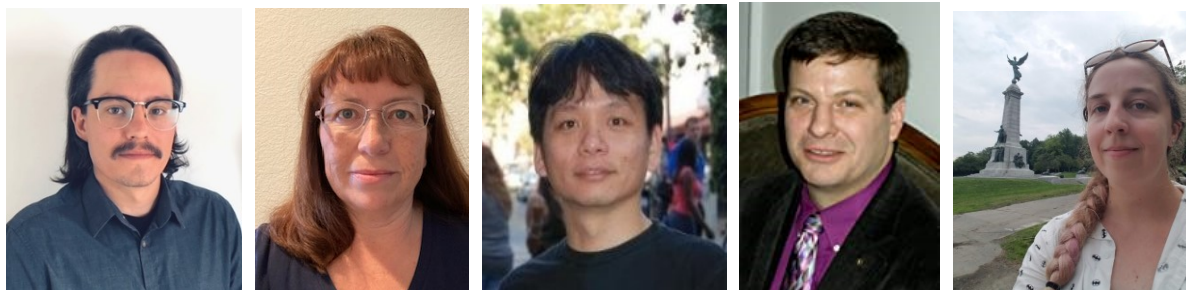
Figure 2. *ACMD welcomed five new Federal staff members during 2021. From left to right: Zachary Grey, Elsie (Meliza) Lane, Simon Su, William Sherman, and Lisa Ritter.*

In November 2021 **Zachary Grey** made the transition from NRC Postdoctoral Associate to a full-time permanent ACMD staff member in Boulder. Zach received a Ph.D. in Computational and Applied Mathematics from the Colorado School of Mines in November 2019. His research interests focus on the use of differential geometry to solve real-world problems, including those requiring dimension reduction and uncertainty quantification. At NIST he has worked with CTL staff on optimizing unlicensed band spectrum sharing with subspace-based pareto tracing. He also established a collaboration with the National Renewable Energy Laboratory on optimizing the shape of wind turbine blades.

Simon Su joined ACMD as a permanent staff member on November 8, 2021. He comes to us from the Army Research Laboratory where he served as an expert in large-scale scientific data analytics and visualization, including immersive virtual reality. He will continue work on such topics in ACMD's High Performance Computing and Visualization Group. Simon holds a Ph.D. in Computer Science from the University of Houston (2001). He also has experience working at a variety of academic labs, such as the Desert Research Institute, the University of Louisiana, and Princeton University, as well as in industrial labs, such as Ball Aerospace, and Lockheed Martin.

Three new postdocs were welcomed via the NIST Guest Researcher Program. **Prajakta Bedekar** graduated from the University of Houston with a Ph.D. in Mathematics in 2020. A modeler of biological systems, she has joint appointments at NIST and at the Johns Hopkins University (JHU). At NIST she is working with Paul Patrone on applications of mathematical modeling in cytometry. **Rayanne Luke** received a Ph.D. in Mathematics from the University of Delaware in 2021 where she specialized in mathematical applications in biology and medicine. She is also a joint NIST-JHU postdoc who is working with Paul Patrone on data analysis in biometrology. **Sesha Challa** (Pavan Kumar) received a Ph.D. in Physics from Pusan National University (India) in 2021. At NIST he is working with Paulina Kuo on optical frequency conversion and electro-optic modulation.

Departures

Steven Satterfield, who served as technical manager of ACMD's immersive scientific visualization lab (the CAVE), retired in December 2020. Steve's work in visualization has enabled advances in a wide variety of NIST measurement science projects, from determining the underlying mechanism of shake gels, to measuring the critical dimensions of engineered polymer scaffolds used in tissue engineering, to revealing the internal interactions of particles in flowing cement. The latter work was honored with the 2020 NIST Judson French Award. The visualization tools he helped design and create have made ACMD's CAVE an agile virtual measurement environment in which new interactive visualizations, and the insights they provide, can be spun up quickly. Steve continues to be affiliated with ACMD as a guest researcher.

Geoffrey McFadden retired from NIST in January 2021 after some 40 years of Federal service. During his tenure at NIST Jeff made fundamental contributions of lasting importance to the mathematics of materials science. Understanding the mechanisms behind pattern formation in materials, which occurs, for example, during the solidification of metal alloys and the growth of crystals is crucial for the development of effective

Figure 3. *Geoffrey McFadden retired from NIST in January 2021.*

techniques for the manufacturing of new materials. Jeff's significant accomplishments in this area include the understanding of three-dimensional cellular interface morphologies, thermo-solutal convection during alloy solidification, the effect of electrical currents on solidification processes, modification of dendritic growth due to fluid flow, the effect of shear flows and step motion on crystal growth, anisotropies of interface energies, kinetics in crystal growth, and order-disorder transformations. He has also made fundamental contributions to phase-field and diffuse-interface models for solidification which have become a powerful predictive tool for materials design. Jeff's work was highly collaborative, as evidenced by the 98 co-authors of his 146 research papers, both highly unusual numbers for a mathematician. Jeff has also been an extraordinary mentor of young researchers evidenced by his direct training of 12 postdocs. Jeff received numerous awards, including the Department of Commerce Gold Medal, the Arthur Flemming Award, and the Washington Academy of Sciences Award for work of Merit and Distinction in Mathematics and Computer Science. He was named a Fellow of the American Physical Society, the Society for Industrial and Applied Mathematics and the Washington Academy of Sciences. He was made a NIST Fellow in 2004. Jeff continues his association with ACMD as a guest researcher and a NIST Fellow Emeritus.

Marjorie McClain retired from ACMD in September 2021 after more than 42 years of government service at NIST. Marjorie was originally hired in the Operations Research Division of the former NBS Center for Applied Mathematics. Over the years she was involved in a wide variety of projects including an IRS post-of-duty location modeling system, specification languages for linear programs, algorithms and software for the calculation of envelope curves, surface approximation methods, including non-oscillatory splines for modeling urban environments, and calculation of uncertainties in coordinate measuring machine measurements. In more recent years, she worked closely with ACMD staff member Bill Mitchell on a survey and experimental comparison of *hp*-adaptive strategies for the solution of challenging elliptic partial differential equations. Marjorie played a largely behind-the-scenes, but critical role in the Digital Library of Mathematical Functions (DLMF) project performing such tasks as ensuring the uniform use of standardized macros in the LaTeX input and linking the extensive DLMF bibliography with online sources. She also worked with Howard Cohl on the development of the follow-on digital repository of mathematical formula.

Figure 4. *Marjorie McClain retired from NIST in September 2020.*

Recognition

ACMD staff members were recognized with a variety of awards this year, including the following.

Anthony Kearsley received the <u>Washington Academy of Science (WAS) Award for Excellence in Research in Applied Mathematics</u>, an honor that was conferred at the annual WAS awards banquet held on September 9, 2021. The citation reads "For outstanding contributions in the application of mathematics with impact in a broad range of fields, including physics, chemistry, biotechnology, and engineering." He was also named a Fellow of the WAS.

Six ACMD staff members were honored with four different 2020 DOC and NIST Awards which were commemorated in January 2021:

- **Bonita Saunders**: <u>NIST Diversity, Inclusion and EEO Award</u>. First presented in 1977, this award is granted for exceptionally significant accomplishments and contributions to Equal Employment Opportunity/Diversity goals. The award consists of an engraved plaque and a $5000 honorarium. Bonita

was honored for exemplary service as a role model, mentor, and tutor in support of STEM careers by women and minorities.

- **William George** and **Steve Satterfield** : <u>NIST Judson C. French Award</u>. First presented in 2000, this award is granted for significant improvement in products delivered directly to industry, including new or improved NIST calibration services, Standard Reference Materials (SRMs), and Standard Reference Databases. The award consists of an engraved plaque and a $5000 honorarium. Bill and Steve were honored along with Nicos Martys of the NIST Engineering Lab and Blaza Toman of the ITL Statistical Engi-

Figure 5. *Colleagues and mentees of Anthony Kearsley celebrate his recognition by the Washington Academy of Sciences. From left to right: Prajakta Bedekar, Ronald Boisvert, Arun Moorthy, Anthony Kearsley, Ryan Evans, Robert DeJaco, Chen Qu, Danielle Brager, and Matthew Roberts.*

neering Division for development of a concrete rheology standard reference material (SRM), the first SRM certified through computer simulation.

- **Scott Glancy** and **Emanuel Knill:** <u>DOC Gold Medal</u> for their work on the team that accomplished the world's first demonstration of deterministic quantum gate teleportation between physically separated trapped ions.

- **Barry Schneider**: <u>DOC Bronze Medal</u> for outstanding initiative as part of the team that developed and deployed Enki, a new computational capability in support of NIST initiatives in big data and AI.

In late 2021 NIST announced that three ACMD staff members would be recipients of the 2021 <u>DOC Silver Medal</u>. The second highest honorary award granted by the Secretary of Commerce, this recognizes exceptional performance characterized by noteworthy or superlative contributions which have a direct and lasting impact within the Department. These staff members are:

- **Thomas Gerrits** of ACMD, along with five colleagues in PML and one in CTL were recognized for exceptional scientific achievement culminating in the first demonstration of an array of superconducting nanowire single-photon detectors (SNSPDs) with more than one thousand pixels. The NIST array represents a 15X improvement in size and pixel count and over one hundred million times better noise properties than conventional arrays. Such arrays of SNSPDs are required for the most demanding applications of imaging at ultralow light levels for astronomy, deep space communications, and medical imaging. The group's technology is already being adopted by industry.

- **Ryan Evans** and **Anthony Kearsley** of ACMD and a team of seven research chemists in MML were recognized for advancing the state-of-the-art in methods for determining the higher order protein structure of biotherapeutics. The group led two technically challenging international comparison studies, establishing the reproducibility and limitations of novel precision measurement methods. Additionally, they developed new data acquisition and analysis tools and new hardware that have hastened broad industry adoption and provide a foundation for measurements that ensure quality and accelerate development of life saving biotherapeutics and biosimilars.

A research paper whose lead author is **Lucas Brady** of ACMD (and a postdoctoral researcher at the NIST/UMD Joint Center for Quantum Information and Computer Science, QuICS) was selected as winner of the 2021 <u>USRA Q2B Applied NISQ Computing Paper Award</u> for 2021. The paper, entitled "Behavior of Quantum Algorithms," was co-authored with Lucas Kocia (Sandia Labs, and a former ACMD NRC postdoc), Przemyslaw Bienias, Aniruddha Bapat, and Yaroslav Kharkov (UMD), and Alexey Gorshkov (NIST PML). The Universities Space Research Association (USRA) announced the award on December 9,
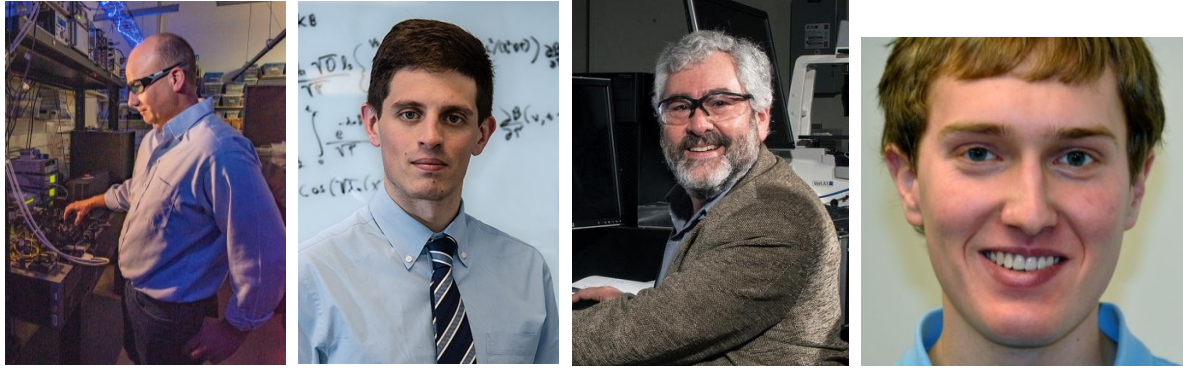
Figure 6. *Recent award winners from ACMD: Left to right: Thomas Gerrits, Ryan Evans, Anthony Kearsley, and Lucas Brady*

2021, at the Q2B conference in Santa Clara, CA. The award is presented to a paper that represents an outstanding advance in the field of applied noisy intermediate scale quantum (NISQ) computing. The paper of Brady et al. unifies the analysis of three major approaches to quantum optimization: adiabatic quantum computing, quantum annealing, and the quantum approximate optimization algorithm (QAOA), resulting in important new insights into their behavior. They also present a new algorithmic protocol that does a better job of approximating the optimal one obtained from optimal control theory. The paper has been posted to the arXiv here: https://arxiv.org/pdf/2107.01218.pdf.

   Daniel Lozier, retired staff member from ACMD (and now a guest researcher) was inducted into NIST's "wall of fame" in October 2020. The NIST Portrait Gallery of Distinguished Scientists, Engineers and Administrators honors NBS/NIST alumni for outstanding career contributions to the work of NBS/NIST. Portraits and biographies of those selected are displayed in the corridor adjacent to the NIST cafeteria at Gaithersburg, and in the Digital Portrait Gallery at NIST Gaithersburg and NIST Boulder sites. Dan was honored for his leadership of the NIST Digital Library of Mathematical Functions project as well as for his research contributions to the field of special functions.

Part II

Features

Computational Modeling and Design of a Wearable Wireless Monitoring System for the Detection of Pulmonary Edema

Pulmonary edema is a medical condition caused by accumulation of excess fluid in the human lungs. The fluid accumulation could occur because of heart problems, pneumonia, exposure to certain toxins, as well as viruses such as COVID-19. Pulmonary edema is a common condition in elderly experienced by 1 in 15 people aged 75 to 84 and just over 1 in 7 people aged 85 years and above. Detection and monitoring of pulmonary edema require access to medical facilities that are equipped with appropriate imaging technologies. Frequent access could be difficult for patients living in rural areas or during pandemics when resources are not readily available. The current pandemic climate has significantly increased the importance of Internet-of-Things (IoT) technologies that can enable remote or self-health monitoring. In this research, we have investigated the feasibility of a simple wearable system that can be used at home to detect or monitor excess fluid buildup in the lungs. To study this, a methodology has been developed to computationally emulate human lungs with various levels of fluid in the alveoli. The proposed wearable system is composed of several small wearable antennas located on the chest and back area. The antennas will operate at the MedRadio frequency band and will be optimized for signal penetration through the body. The frequency and time responses of the communication channel between these antennas for the lung models with varying levels of fluid have been measured and analyzed. The results show a correlation between the channel response and the level of fluids inside the lungs. This correlation can potentially be exploited by a simple wearable system to predict the onset of pulmonary edema.

Kamran Sayrafian

There are numerous tiny air sacs (known as alveoli) in the human lungs where oxygen is exchanged with carbon dioxide during the breathing process. The addition of fluid in these air sacs will make breathing difficult. If this condition is not treated in time, shortness of breath sets in, leading to acute respiratory distress syndrome (ARDS) which is a form of lung failure. X-ray, CT scan, and Electrical Impedance Tomography are typical technologies that are used to detect and measure fluid buildup in the lungs; however, physical presence and access of the patient to medical facilities having necessary equipment are required for accurate diagnosis. Frequent access to such facilities might not be easily feasible or convenient for patients living in remote areas or people who need to be continuously monitored. In addition,
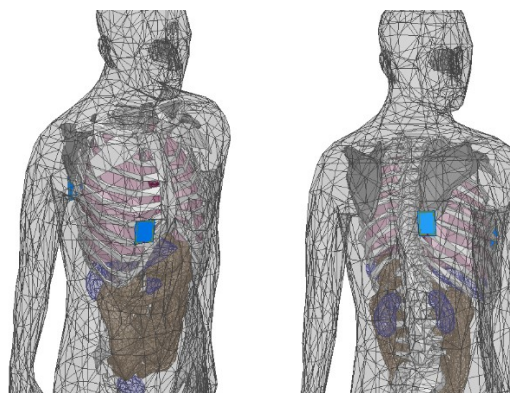
Figure 7. *Computational human body model with the antenna placements indicated.*

during pandemics, resources at the medical facilities are typically diverted to other critical tasks. Therefore, it is desirable to cut down the number of unnecessary visits by the patients during those times. Recent advances in IoT technology as well as proliferation and consumer adoption of wearable devices have created a fertile environment for new applications that enable health monitoring functions to take place at the patient's home. To the best of our knowledge, there are currently no wearable or mobile devices in the market that can be used to detect fluid buildup in the human lungs.

In this project, we first developed a methodology to computationally emulate human lungs with various levels of fluid in the alveoli. The lung models will be part of a full 3D computational human body model that has been used for a variety of research related to the application of wearable, implant, and ingestible electronics [1]. A sequence of virtual experiments has also been developed to evaluate the performance of a wearable system designed to monitor pulmonary edema. The proposed system is composed of several small wearable antennas located on the chest and back side of the body as shown in Figure 7. The antennas are designed to operate at the MedRadio (Medical Device Radio-communications Service) frequency band (401 MHz to 406 MHz) and optimized for signal penetration through the body. The MedRadio spectrum can be used for diagnostic and therapeutic purposes in implanted and wearable wireless devices.

There are 480 million alveoli in the human lungs on average. The mean size of a single alveolus is estimated to be around $4.2 \times 10^6 \, \mu m^3$ [2]. Computational modeling of such small structures within the lung tissue is nearly impossible. However, such microscopic level of modeling might not be necessary for the purpose of our study. The net effect of fluid buildup in the alveoli is the
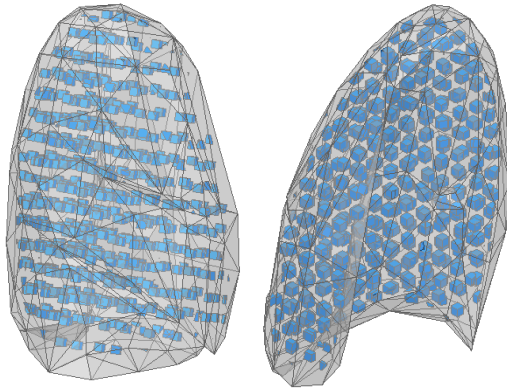
Figure 8. *Computational model of the lungs with a lattice of macro-alveolus.*

change in the dielectric properties of the lung's tissue (i.e., the medium where the electromagnetic waves travel through). However, as knowledge of these dielectric properties for various levels of fluid is not available, we propose a macroscopic computational model in which a cluster of alveoli are considered as one macro-alveolus. Using this approach, the following methodology is used to construct a lung model incorporating the effect of fluid buildup. A periodic lattice of small cubes is created inside the volume of the lungs. Each cube (referred to as Lattice Element (LE)) represents the overall volume of fluid that has been accumulated in a macro-alveolus. To remove the periodicity of the lattice, an algorithm has been developed to randomize the spacing between the LEs. It should be noted that since the size/dimension of the lattice elements is small compared to the wavelength of the signal that we are using in this study the actual shape of the lattice elements (i.e., cubes) will not have any impact on the final results and other shapes such as spheres may also be used.

Figure 8 shows a sample of the lung model including the randomized lattice structure. The two parameters in this lung model are the LE size and the average spacing between two adjacent LEs. The average spacing will determine the total number of LEs that are contained within the volume of the lungs. The number of LEs along with their size will also determine the total volume of the LEs that exist inside the lungs. This volume indicates the total fluid buildup which impacts the overall mass density of the lungs representing pulmonary edema. This mass density can be calculated according to the formula

$$\rho_{lung}^{(LE\,size,\ LE\,number)} = \big[\rho_{lung}^{(0,\ 0)} \times (V_{lung} - V_{fluid}^{(LE\,size,\ LE\,number)})$$
$$+ \rho_{fluid} \times V_{fluid}^{(LE\,size,\ LE\,number)}\big]/V_{lung}$$

where

$\rho_{lung}^{(LE\,size,\ LE\,number)}$ is mass density of the lungs with excess fluid, $\rho_{lung}^{(0,\ 0)}$ is the mass density of the normal lungs tissue (fully inhaled), ρ_{fluid} is the mass density of the

fluid inside the lungs, $V_{fluid}^{(LE\,size,\ LE\,number)}$ is the total volume of the fluid inside the lungs, and V_{lung} is the total volume of the lungs.

The substance of the fluid that builds up inside the alveoli is approximated to be 0.9 % saline solution. Therefore, ρ_{fluid} = 1.0046 g/cm^3. Electrical properties of lossy liquids such as sodium chloride (NaCl) are very much temperature dependent. Using the polynomial equations in [3], we can calculate conductivity and permittivity of sodium chloride solutions as a function of normality and temperature. The temperature range where the model in [3] was validated with actual data was 5 to 35 °C. However, assuming that the model is still valid for 37 °C (i.e., the normal human body temperature) and using a concentration/molarity of 0.154 mol/L for sodium chloride solutions, the dielectric properties of 0.9 % saline solution at body temperature (37°C) can be estimated to be 72.47 F/m in permittivity and 1.86 S/m in conductivity.

The average mass density of the normal (fully inhaled) lung tissue (ρ_{tissue}) is known to be 0.2 g/cm^3 [4]. Assuming a dimension of 6 mm for an individual cubic-shaped macro-alveolus/LE (equivalent to a volume of 216×10^9 μm^3) and considering that the volume of the 3D computational lung model in our simulation is V_{lung} = 2309760.59 mm^3, Table 1 shows the resulting mass densities of the lungs for various concentration of LEs. This mass density is used as a parameter to study the impact of fluid accumulation on the forward transmission coefficient of the wireless channel between the two

Table 1. *Mass density of lungs for various numbers of macro-alveolus.*

Average LE Spacing (mm)	No. of LEs	V_{fluid} (mm^3)	$\rho_{lung}^{(LE\,size,\ LE\,number)}$ (g/cm^3)
-	0	0	0.2
9.5	604	105791.19	0.24
5.5	1495	263640.72	0.29
4	2284	400857.96	0.34
3	3173	552898.45	0.39
2.3	3895	678601.29	0.44

antennas shown in Figure 7. The possible range of this mass density parameter is from 0.2 g/cm^3 (indicating healthy lung tissue at the fully inhaled state) to an extreme 1.0 g/cm^3 where the whole lung volume is basically considered to be filled with 0.9 % saline solution. Table 1 only shows the relevant information up to 0.45 g/cm^3. It is assumed that the detection of variations in mass densities in the lower range (i.e., up to 0.45 g/cm^3) are more important and should be the main focus for home or remote monitoring applications. Beyond that the patient should most likely transfer to a hospital or clinic to receive proper examination and care.

Using a pair of specially designed and optimized antennas and placing them across the right lung as shown in Figure 7, the forward transmission coefficient of the wireless channel (S_{21}) has been measured for various parameters of the model or mass densities of the lung shown in Table 1. It is observed that with increasing mass density, the S_{21} parameter is gradually reduced. Fluid accumulation in the alveoli will change the effective dielectric properties of the lung tissue, which in turn impacts the scattering parameters of the wireless channel between the antennas. A noticeable 4.5 dB loss in S_{21} is observed when the mass density increases to 0.44 g/cm^3. This change can be easily measured and can serve as a warning sign for the onset of pulmonary edema. We conjecture that the size of the macro-alveolus does not significantly impact the S_{21} results. The validity of this conjecture has been verified for a few computationally feasible sizes of the macro-alveolus (i.e., 6 mm, 5 mm, and 4 mm).

Time domain responses of the wireless channel can also be obtained using the Inverse Fourier Transform (IFT) after applying a Hann windowing on the S_{21} results. Figure 9 shows these channel responses when an envelope detector is used to remove the mid-band carrier frequency of 403 MHz from the IFT output. The impact of the change in the lung mass density is also observed through the peak amplitude of the channel responses. The reduction of this amplitude indicates an increase in the lung mass density due to excess fluid accumulation. Therefore, through constant monitoring of the amplitude of the time domain channel response, the onset of fluid buildup in the lungs can be detected. It should be noted that only the relative reduction of the peak amplitude with respect to the baseline response of the channel (i.e., response of the healthy lung) is important. This is because the absolute numerical value of the baseline response could be different for people with different body sizes and shapes. Monitoring of the channel response can be done by patients at their homes or remotely if wireless Internet connectivity is provided.

The results obtained in this project demonstrate the feasibility of a simple wireless wearable device for detecting pulmonary edema [5]. The main concept used in this research could also serve as a starting point for future research on wearables devices that can assist in detecting abnormalities inside the human organs or tissues. It is envisioned that the wearable antennas used in this research along with their corresponding electronics could be integrated in the fabric of a wearable band or a vest. The band/vest can be worn at home by patients at risk or individuals in remote areas with no convenient access to a medical facility. The information provided by the wearable device can also be transmitted through
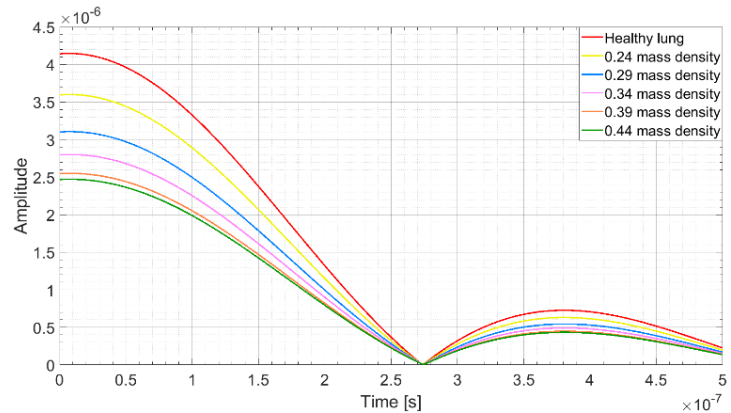
Figure 9. *Time Response of the channel for various mass densities.*

a telemedicine network, therefore reducing the frequency of unnecessary trips to medical facilities. This is especially important during a pandemic lockdown.

This study has been conducted with a static computational model of the lungs. A dynamic simulation platform that can continuously measure the signal during natural breathing cycle (inhalation and exhalation) could be quite challenging. The dielectric properties used for the model assumes fully inhaled lungs. Therefore, additional circuitry might be needed to ensure the measurements are taken at the right instances during the breathing process. More detailed computational models corresponding to fully inhaled and exhaled human lungs are needed in order to study possible advantages of performing measurement at each state. Array antennas may also be used to create a more directional propagation path through the lungs. This could further enhance the resolution of the channel responses. Ultimate verification of the methodology presented in this project requires physical experiments involving patients that are susceptible to pulmonary edema.

References

[1] K. Sayrafian, J. Hagedorn, W. B. Yang, and J. Terrill. A Virtual Reality Platform to Study RF Propagation in Body Area Networks In *IEEE 3rd International Conference on Cognitive Infocommunications* (CogInfoCom), Kosice, Slovakia, December 2-5, 2012. DOI: 10.1109/CogInfoCom.2012.6421944

[2] M. Ochs, J. R. Nyengaard , A. Jung , L. Knudsen , M. Voigt , T. Wahlers , J. Richter , H. Jørgen, and G. Gundersen. The Number of Alveoli in the Human Lung. *The American Journal of Respiratory and Critical Care Medicine* **169**:1 (2003). PubMed: 14512270

[3] A. Peyman, C. Gabriel, and E. H. Grant. Complex Permittivity of Sodium Chloride Solutions at Microwave Frequencies. *Bioelectromagnetics* **28**:4 (2007), 264-74. DOI: 10.1002/bem.20271

[4] J. Van Dyk, T. J. Keane, and W. D. Rider. Lung Mensity as Measured by Computerized Tomography:

Implications for Radiotherapy. *International Journal of Radiation Oncology, Biology, Physics* **8**:8 (1982), 1363-72. DOI: 10.1016/0360-3016(82)90587-9

[5] K. Krhac, K. Sayrafian, U. Bengi, and S. Dumanli-Oktar. A Wearable Wireless Monitoring System for the Detection of Pulmonary Edema. In *IEEE Global Communications Conference* (GLOBECOM), Madrid, Spain, December 7-11, 2021.

Participants

Kamran Sayrafian (ACMD); Katjana Krhac, Uzay Bengi, Sema Dumanli (Boğaziçi University, Turkey)

Probability Distributions in Peak Matching for DART-MS

Direct Analysis in Real Time Mass Spectrometry (DART-MS) is a technology that can be used for chemical identification. One application is screening for controlled substances in forensic seized drug evidence. This rapid procedure produces complex mass spectra because the spectral signatures of all compounds in the sample are observed simultaneously. A major advantage of DART-MS over other methods is its ability to produce multiple measurements of mass spectra so quickly. As DART-MS has become more commonplace in drug screening, differentiating the mass spectra of structurally similar compounds has become increasingly important. Unfortunately, it can be difficult to distinguish between structurally similar compounds using DART-MS. Even with high resolution mass spectra, existing measures of mass spectral similarity cannot always discriminate between structurally similar compounds. One approach to solving this problem is to utilize the statistical nature of the peaks in mass spectra coming from DART-MS in order to determine similarity. Using this method, we have developed a similarity measure between spectra coming from DART-MS that accurately distinguishes structurally similar compounds.

Matthew Roberts

Direct Analysis in Real Time Mass Spectrometry (DART-MS) is a technology that can be used for chemical identification. One application is screening for controlled substances in forensic seized drug evidence. With DART-MS, heated gas metastables are used to both desorb and ionize an analyte. This rapid procedure produces complex mass spectra because the spectral signatures of all compounds in the sample are observed simultaneously. Here we consider high resolution centroided in source collision induced (is-CID) mass spectra coming from DART-MS, so for convenience, we will always be referring to this whenever we mention mass spectra. A major advantage of DART-MS over other methods is its ability to produce multiple measurements of mass spectra so quickly. However, since DART-MS is not performed in a vacuum, resulting spectra have a greater tendency to vary due to contamination [6]. As DART-MS has become more commonplace in drug screening, differentiating the mass spectra of structurally similar compounds has become increasingly important. Figure 10 and Figure 11 show the mass spectra for Methamphetamine and Phentermine at 60V, two structurally similar compounds that can be difficult to distinguish using DART-MS. Even with high resolution mass spectra, existing measures of mass spectral
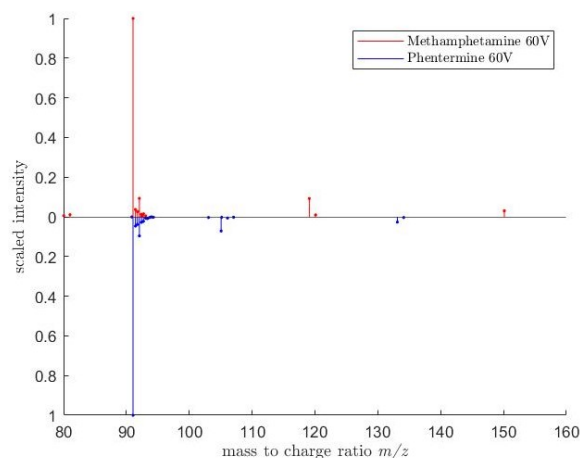
Figure 10. *Measured 60V spectra of Methamphetamine in red and Phentermine in blue. Intensities are scaled to the interval [0,1] with base peak at m/z 91.06 (representing the protonated molecule).*
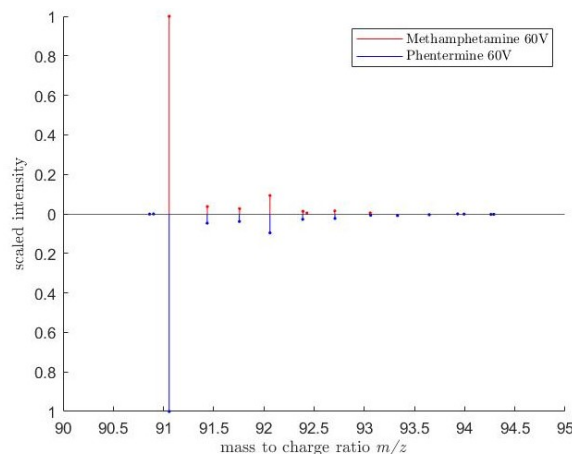
Figure 11. *Measured 60V spectra over the m/z interval [90,95], with Methamphetamine in red and Phentermine in blue. Intensities are scaled to the interval [0,1] with base peak at m/z 91.06 (representing the protonated molecule).*

similarity cannot always discriminate between structurally similar compounds. Because of factors like mass defect in isotopic compounds in high resolution spectra, distinguishment of structurally similar compounds can be inaccurate without additional filtering [5].

One approach to solving this problem is to utilize the statistical nature of the peaks in mass spectra coming from DART-MS in order to determine similarity. Even though DART-MS has greater variation of spectra than other methods of mass spectrometry, the rapid process of DART-MS allows one to collect several measurements of the same compound in very little time. Using these multiple measurements, one is able to quantify the uncertainty in each peak in the spectra. In this way, each
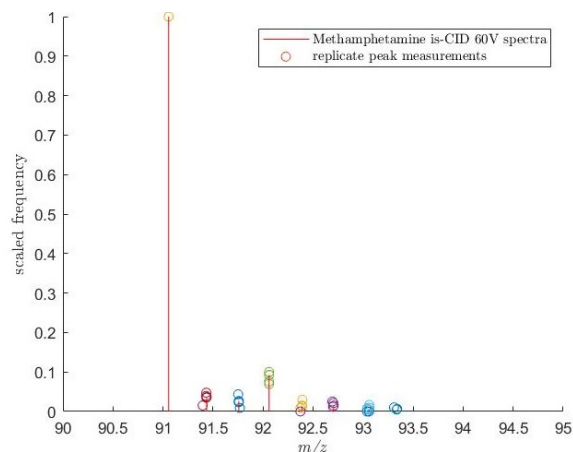
Figure 12. *Measured 60V for Methamphetamine together with the sets Sj for the top 8 largest prominent peaks for Methamphetamine for m/z 90 to m/z 95.*
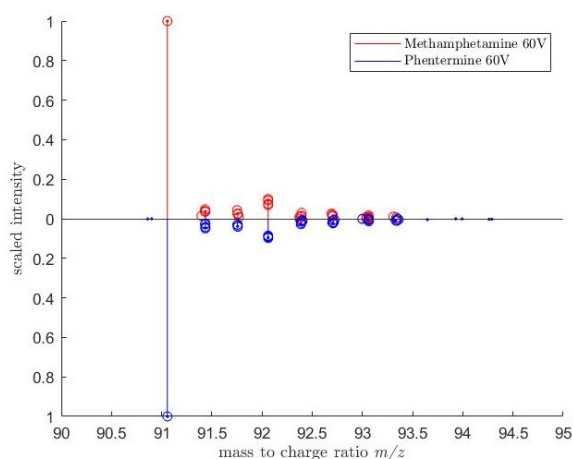
Figure 13. *Measured 60V spectra over the m/z interval [90,95] with Methamphetamine in red and Phentermine in blue. This together with the sets Sj for each compound for the top 8 largest prominent peaks for m/z 90 to m/z 95.*

peak is identified as a probability distribution, and by comparing these probability distributions, one is able to match peaks between two mass spectra appropriately. Using this method, we have developed a similarity measure between spectra coming from DART-MS that accurately distinguishes structurally similar compounds. Additionally, since probability distributions are being used to model spectra, inaccurate similarity measure due to inappropriate binning does not occur like it does with traditional cosine similarity, or other methods of peak matching [4].

In this approach, the first challenge is to understand the statistical nature of each peak in the spectra for a particular compound. By matching peaks between replicate measurements of spectra according to Euclidean distance, groups of matched peaks are formed. Each group of matched peaks produces a sample mean and standard

deviation that together form what we call a peak statistic. This collection of peak statistics is then referred to as a mass spectrum statistic.

Each peak statistic in a mass spectrum statistic is then naturally identified with a two-dimensional normal distribution. In this way, the similarity between two peak statistics can be determined by computing the similarity between their associated normal distributions. Using this similarity measure, a peak matching scheme is then proposed to accurately match peak statistics between two mass spectrum statistics. A similarity score between two mass spectrum statistics can then be computed according to this matching.

Using this method of probability distributions, we have analyzed replicate measurements of 7 pairs of structurally similar compounds. Numerical results show that this method outperforms the traditional dot product method when discriminating these pairs of compounds.

Peak and Spectrum Statistics. The major advantage of DART-MS over other techniques in mass spectrometry is its ability to produce multiple measurements of a substance in very little time. The idea is to utilize these multiple measurements to produce a statistic of the mass spectrum for that substance. Let $a_1, a_2, \cdots, a_n$ be n replicate mass spectral measurements for a substance. In this case, a_i is a collection of mass spectra points in $\mathbb{R}^2$ and may vary in size depending on the measurement. Here, for each point (x, y) in the spectrum, the first coordinate x is the mass-to-charge ratio m/z, and the second coordinate y is the relative intensity, a number from 0 to 1. For the compounds we consider, a spectral measurement has usually between 15 and 40 points.

Starting with the peak with greatest intensity of all the measurements, say peak p_1 from measurement a_i, we determine the peak from each measurement a_j that is closest to peak p_1 in Euclidean distance. This will form the set of peaks $S_1 = \{p_{1,1}, p_{1,2}, \cdots, p_{1,n}\}$, where $p_{1,j}$ is the peak coming from a_j for each j. From this it is clear that $p_{1,i} = p_1$ since $p_1 \in a_1$. The idea is that S_1 consists of the replicate measurements for peak p_1 when grouped this way.

Once the set S_1 has been determined, these peaks are removed from the spectra $a_1, a_2, \cdots, a_n$. Starting with the remaining peak p_2 of greatest intensity of all the remaining spectra $a_1, a_2, \cdots, a_n$., the process is repeated to produce the set $S_2 = \{p_{2,1}, p_{2,2}, \cdots, p_{2,n}\}$, of replicate measurements of peak p_2. In this way, the sets $S_1, S_2, \cdots, S_m$ can be generated. Here,

$$S_j = \{p_{j,1}, p_{j,2}, \cdots, p_{j,n}\}$$

where $p_{j,k}$ is the peak in a_k closest to peak p_j in Euclidean distance. Figure 12 and Figure 13 show the 60V is-CID spectra of Methamphetamine and Phentermine together with the sets S_j of replicate peak measurements over the m/z interval [90, 95].
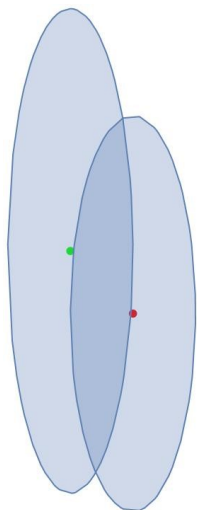
Figure 14. *Peak statistics P and Q with mean and 1 standard deviation in x and y directions.*
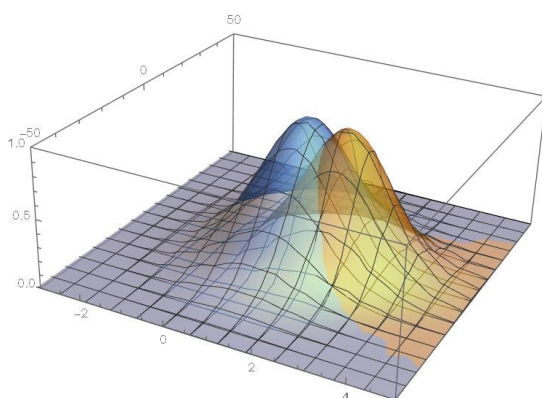
Figure 15. *Plot of probability distributions $f_P(x, y)$ and $f_Q(x, y)$.*

It can be a challenge to determine what is the appropriate value of m. In other words, how many peaks $p_1, p_2, \cdots, p_m$ should be considered when making the sets $S_1, S_2, \cdots, S_m$? In our analysis, we have used 15 peaks in general, but the largest m is equal to the smallest of the replicate spectra $a_1, a_2, \cdots, a_n$. The idea is to choose m such that S_j does not contain artifact peaks for any j. We will refer to the peaks $p_1, p_2, \cdots, p_m$ as prominent peaks.

In essence, S_j is the set of replicate measurements of prominent peak p_j. Therefore, the sample mean $\bar{p}_j = (\bar{p}_{x,j}, \bar{p}_{y,j})$ and the standard deviation $s_{p_j} = (s_{p_{x,j}}, s_{p_{y,j}})$ of S_j gives information about the location and variance of prominent peak p_j. The peak statistic P_j is defined by the 4-tuple

$$P_j = \left(\bar{p}_{x,j}, \bar{p}_{y,j}, s_{p_{x,j}}, s_{p_{y,j}} \right)$$

and the collection of m peak statistics

$$A = \{P_1, P_2, \cdots, P_m\}$$

is defined to be the spectrum statistic of $a_1, a_2, \cdots, a_n$ using m prominent peaks.

The next step is to now utilize the spectra statistics of two compounds to accurately distinguish them from one another. This is done by interpreting each peak statistic as a probability distribution and then match peak statistics between spectra statistics according to their similarity.

Peak Similarity and Peak Matching with Peak Statistics. One of the difficulties with peak matching between mass spectra is determining an appropriate metric in which to compare the peaks [4]. With a peak statistic, additional information is available to explain how the peak may vary from measurement to measurement. Therefore, a peak statistic explains the distribution of possible measurements for that given peak. In this way, a peak statistic $P = \left(\bar{p}_x, \bar{p}_y, s_{p_x}, s_{p_y} \right)$ can be identified with the 2D normal probability distribution

$$f_P(x, y) = \frac{1}{2\pi s_{p_x} s_{p_y}} \, e^{-\frac{1}{2}\left[\left(\frac{x - \bar{p}_x}{s_{p_x}}\right)^2 + \left(\frac{y - \bar{p}_y}{s_{p_y}}\right)^2\right]}$$

Similarly for a peak statistic $Q = \left(\bar{q}_x, \bar{q}_y, s_{q_x}, s_{q_y} \right)$, we can identify it with the 2D normal probability distribution

$$f_Q(x, y) = \frac{1}{2\pi s_{q_x} s_{q_y}} \, e^{-\frac{1}{2}\left[\left(\frac{x - \bar{q}_x}{s_{q_x}}\right)^2 + \left(\frac{y - \bar{q}_y}{s_{q_y}}\right)^2\right]}$$

By comparing the distributions f_P and f_Q, one has a natural way of determining the similarity between peak statistics P and Q. This comparison is done by determining the cosine similarity between f_P and f_Q in the space $L^2(\mathbb{R}^2)$. Just how a mass spectrum can be interpreted as a vector in $\mathbb{R}^n$, the functions f_P and f_Q are vectors in the function space $L^2(\mathbb{R}^2)$ and their cosine similarity is determined by the cosine of the angle between them. In the space $L^2(\mathbb{R}^2)$, the inner product $\langle f_P, f_Q \rangle_{L^2(\mathbb{R}^2)}$ is a generalization of the dot product between vectors in $\mathbb{R}^n$ and is given by

$$\langle f_P, f_Q \rangle_{L^2(\mathbb{R}^2)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_P(x, y) \, f_Q(x, y) \, dy \, dx$$

Therefore, the norm $\|f_P\|_{L^2(\mathbb{R}^2)}$ in $L^2(\mathbb{R}^2)$ generated by the inner product is given by

$$\|f_P\|_{L^2(\mathbb{R}^2)} = \sqrt{\langle f_P, f_Q \rangle_{L^2(\mathbb{R}^2)}} = \sqrt{\int_{-\infty}^{\infty} f_P(x, y)^2 \, dy \, dx}$$

In the same way the cosine of the angle between vectors in $\mathbb{R}^n$ is computed using the dot product and norm in this space, the cosine of the angle between f_P and f_Q in $L^2(\mathbb{R}^2)$ is computed by

$$\cos(\theta_{f_P, f_Q}) = \frac{\langle f_P, f_Q \rangle_{L^2(\mathbb{R}^2)}}{\|f_P\|_{L^2(\mathbb{R}^2)} \|f_Q\|_{L^2(\mathbb{R}^2)}}$$

$$= \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_P(x,y) \, f_Q(x,y) \, dy \, dx}{\sqrt{\int_{-\infty}^{\infty} f_P(x,y)^2 \, dy \, dx} \, \sqrt{\int_{-\infty}^{\infty} f_Q(x,y)^2 \, dy \, dx}}$$

Here θ_{f_P, f_Q} is the angle between f_P and f_Q, and $\cos(\theta_{f_P, f_Q})$ is computed exactly by the formula

$$\cos(\theta_{f_P, f_Q}) = \sqrt{\frac{4 s_{p_x} s_{q_x} s_{p_y} s_{q_y}}{\left(s_{p_x}^2 + s_{q_x}^2\right)\left(s_{p_y}^2 + s_{q_y}^2\right)}}$$

$$\times \, e^{-\frac{1}{2}\left[\left(\frac{p_x - q_x}{\sqrt{s_{p_x}^2 + s_{q_x}^2}}\right)^2 + \left(\frac{p_y - q_y}{\sqrt{s_{p_y}^2 + s_{q_y}^2}}\right)^2\right]}$$

Splitting this into x and y components we have

$$\cos(\theta_{f_P, f_Q}) = \left(\sqrt{\frac{2 s_{p_x} s_{q_x}}{\left(s_{p_x}^2 + s_{q_x}^2\right)}} \, e^{-\frac{1}{2}\left(\frac{p_x - q_x}{\sqrt{s_{p_x}^2 + s_{q_x}^2}}\right)^2} \right)$$

$$\times \left(\sqrt{\frac{2 s_{p_y} s_{q_y}}{\left(s_{p_y}^2 + s_{q_y}^2\right)}} \, e^{-\frac{1}{2}\left(\frac{p_y - q_y}{\sqrt{s_{p_y}^2 + s_{q_y}^2}}\right)^2} \right)$$

Identifying peak statistics with probability distributions is a natural way for determining how well two peak statistics match each other. The similarity measure between two peak statistics P and Q is denoted by $\theta(P, Q)$ and is given by the formula

$$\theta(P, Q) = \cos(\theta_{f_P, f_Q})$$

where all of the terms are defined as above. The similarity measure $\theta(P, Q)$ is a value between 0 and 1, with 0 if there is no similarity and 1 if the peak statistics are identical.

Using the similarity measure θ, we can now present a scheme for peak matching that accurately matches peaks between spectra statistics. Let $A = \{P_1, P_2, \cdots, P_{m_1}\}$ and $B = \{Q_1, Q_2, \cdots, Q_{m_2}\}$ be two spectra statistics. To compare A and B, we consider at most $m = min(m_1, m_2)$ peak statistics, or prominent peaks. For our results, we have used a value of m = 15 unless $min(m_1, m_2)$ was smaller. Peak matching is then performed in which pairs of peaks are chosen such that they maximize similarity measure θ. No peak statistic can be in two pairs, so optimal pairs of peaks are determined starting from peaks with largest relative intensities and moving to the smallest.

Table 2. *Seven pairs of structurally similar compounds.*

	Pairs of Similar Compounds Considered
1	Cotinine, Serotonin
2	Phenibut, MDA
3	MMDPPA, Methylone
4	5-methoxy MET, Norfentanyl
5	Cocaine , Scopolamine
6	HU-210 , Testosterone Isocaprinate
7	Methamphetamine, Phentermine

Table 3. *Similarity scores $\phi(A, B)$ where A and B are the spectra statistics for each pair of compounds at energy levels 30V, 60V, and 90V.*

Similarity Score of $\phi(A, B)$	30V	60V	90V
Cotinine , Serotonin	9.6E-8	0.0088	0.0019
Phenibut, MDA	2.12e-6	3.17e-5	3.92e-13
MMDPPA, Methylone	0.0192	0.0073	6e-4
5-methoxy MET, Norfentanyl	0.6457	4.5e-5	6.65e-24
Cocaine, Scopolamine	0.0012	0.0439	0.0078
HU-210, Testosterone Isocaprinate	0.0630	2.68e-5	4.96e-6
Methamphetamine, Phentermine	0.0096	0.5485	0.6661

Table 4. *Maximum cosine similarity values for each of the 7 pairs of structurally similar compounds at each of the energy levels 30V, 60V, and 90V.*

Max Cosine Similarity $\mathbb{R}^n$	30V	60V	90V
Cotinine, Serotonin	0.2007	0.0360	0.1355
Phenibut , MDA	0.5725	0.0235	0.1014
MMDPPA, Methylone	0.6551	0.0201	0.0456
5-methoxy MET, Norfentanyl	0.9852	0.0055	0.0095
Cocaine, Scopolamine	0.6091	0.0559	0.1278
HU-210, Testosterone Isocaprinate	0.7843	0.3202	0.2066
Methamphetamine, Phentermine	0.7048	0.9922	0.9988

To start, let $P_{i_1} = \left(\bar{p}_{x,i_1}, \bar{p}_{y,i_1}, s_{p_{x,i_1}}, s_{p_{y,i_1}}\right)$ be the peak statistic with greatest relative intensity out of both A and B. That is, P_{i_1} had the largest second coordinate value $\bar{p}_{y,i_1}$ of any of the peak statistics in either A or B. Next, the peak statistic Q_{j_1} is selected from B such that the similarity measure $\theta(P_{i_1}, Q_{j_1})$ is maximized. Once the pair (P_{i_1}, Q_{j_1}) has been determined, these peak statistics are set to the side, and the process repeats for the remaining peak statistics in A and B. In this way, the peak statistic pairs $(P_{i_1}, Q_{j_1}), (P_{i_2}, Q_{j_2}), \cdots, (P_{i_m}, Q_{j_m})$ are created from first to last. The similarity score $\phi(A, B)$ between A and B is then determined by

$$\phi(A, B) = \frac{\sum_{k=1}^{m} \bar{p}_{y,i_k} \bar{q}_{y,j_k} \theta(P_{i_k}, Q_{j_k})}{\sum_{k=1}^{m} \bar{p}_{y,i_k} \bar{q}_{y,j_k}}$$

In this way, similarity measure between peak statistics is weighted by the product of their intensities. The similarity score $\phi(A, B)$ between A and B is a number between 0 and 1, where 0 means no similarity and 1 means identical.

Results. Using the method of probability distributions, we have considered 7 pairs of structurally similar compounds with similar is-CID spectra. These compounds are listed in Table 2. For each of these compounds, 5 measurements of is-CID spectra were collected at 30V, 60V, and 90V.

Using these replicate measurements of DART-MS, we were able to generate a spectra statistic for each compound at each of the three fragmentation levels. For each pair of compounds, spectra statistics A and B are generated, and the similarity scores $\phi(A, B)$ are recorded in Table 3 for each of the 3 fragmentation levels.

We compare these results to the classical method of cosine similarity using the dot product. For a given pair of compounds from Table 2, we compare the similarity of replicate measurements $a_1, a_2, \cdots, a_5$ of one compound to replicate measurements $b_1, b_2, \cdots, b_5$ of the other compound. Here, the spectra a_i and b_j are identified as vectors in $\mathbb{R}^n$ using classical binning with a bin width of 0.1 and with the last bin at 899.9 m/z to 900 m/z, which makes $n = 9000$. We then consider the maximum cosine similarity score between any two a_i and b_j coming from the two sets of replicate measurements [3]. This similarity is determined by the formula

$$cos\left(\theta_{a_i, b_j}\right) = \frac{a_i \cdot b_j}{\|a_i\|_2 \|b_j\|_2}$$

Here, θ_{a_i, b_j} is the angle between a_i and b_j in $\mathbb{R}^n$, and the norm $\|\cdot\|_2$ is the Euclidean norm in $\mathbb{R}^n$. The maximum cosine similarity values between any two replicate measurements a_i and b_j are recorded for each pair of compounds at energy levels of 30V, 60V, and 90V in Table 4.

As can be observed, the method of peak matching with probability distributions outperforms the classical cosine similarity method for distinguishing every pair of compounds at every single fragmentation level. Because of the limited measurements, the similarity score method of ϕ could only be tested on one pair of spectra statistics for each pair of compounds in Table 2, but the results are convincing to look at further testing of this method.

References

[1] E. Sisco, A. S. Moorthy, and L. M. Watt. Creation and Release of an Updated NIST Dart-MS Forensics Database. *Journal of the American Society for Mass Spectrometry* **32**:3 (2021), 685-689.

[2] A. S. Moorthy and E. Sisco. A New Library-Search Algorithm for Mixture Analysis using DART-MS. *Journal of the American Society for Mass Spectrometry* **32**:7 (2021), 1725-1734.

[3] A. S. Moorthy and E. Sisco. The Min-Max Test: An Objective Method for Discriminating Mass Spectra. *Analytical Chemistry* **93** (2021), 13319-13325.

[4] A. J. Kearsley and A. S. Moorthy. Identifying Fentanyl with Mass Spectral Libraries. ChemRxiv. Cambridge: Cambridge Open Engage, 2021.

[5] L. Sleno. The Use of Mass Defect in Modern Mass Spectrometry. *Journal of Mass Spectrometry* **47** (2021), 226-236

[6] R. B. Cody, J. A. Laramée, J. M. Nilles, and H. D. Durst. Direct Analysis in Real Time (DART) Mass Spectrometry. *JEOL News* **40**:1 (2005): 8-12.

Participants

Matthew Roberts, Anthony Kearsley (ACMD); Arun Moorthy, Edward Sisco (NIST MML)

Automation of Experimental Quantum Dot Control

Electrons confined in arrays of semiconductor nanostructures, called quantum dots (QDs), are a promising quantum computing approach. Due to the ease of control of the relevant parameters, fast measurement of the spin and charge states, relatively long decoherence times, and their potential for scalability, QDs are gaining popularity as candidate building blocks for solid-state quantum devices. However, with an increasing number of QD qubits, the relevant parameter space grows sufficiently to make heuristic control unfeasible. In semiconductor quantum systems, devices now have tens of individual electrostatic and dynamical gate voltages that must be carefully set to localize the system into the single-electron regime and to realize good qubit operational performance. It is thus highly desirable to have an automated protocol to achieve a target electronic state.

In recent years, there has been a considerable effort to automate every phase of device calibration and control, from taking a system from room temperature to quantum operation. This project is an effort to automate certain stages of device control—namely the coarse tunning and setting the desired charge state—by combining script-based algorithms with machine learning (ML) techniques. Our efforts rely on combining ML algorithms trained using synthetic data from a physical model with classical optimization techniques to establish an automated closed-loop system for experimental control. Our work serves as a baseline for future investigation of fully automated device control system and to pave the way for similar approaches in a wide range of experiments in physics.

Justyna P. Zwolak

There are myriad quantum computing approaches, each having its own set of challenges to understand and effectively control their operation. Arrays of electrostatically defined *quantum dots* (QDs) present at the interface of semiconductor devices is one such approach [1]. Unfortunately, the difficulty of working with such systems—where fabrication tolerances are tight, impurities punishing, and material considerations vast—has meant that only a few groups around the world have succeeded in advancing the limit of QD performance.

The current practice of tuning QDs manually or in a semi-automated fashion is a relatively time-consuming procedure, susceptible to random errors and inherently impractical for scaling up and other applications. Even tuning a double QD constitutes a nontrivial task, with each dot being controlled by at least three metallic gates, each of which influences the number of electrons in the
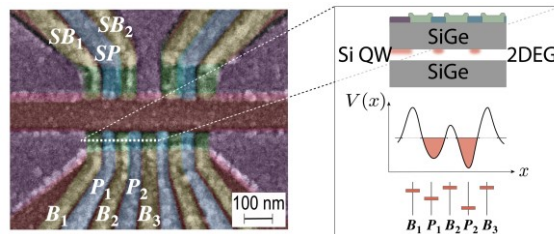
Figure 16. *A false-color scanning electron micrograph of Si/SiGe quadruple QD device. The top gates (SB_1, SB_2, and SP) are used to form the sensing QD while the bottom gates (B_i for i=1,2,3 and P_j for j=1,2) are used to form and control the qubit QDs. The inset, showing a cross section through the device along the dashed white line (top) and a schematic of the electric potential of a tuned double dot (bottom). Adapted from [2].*

dot, the tunnel coupling to the adjacent lead, and the interdot tunnel coupling (see Figure 16). The presence of defects and variations in the local composition of the heterostructure disordering the background potential energy further impedes this process. At the same time, given the progress in the construction of multi-QD arrays in both one and two dimensions (1D and 2D, respectively) [3, 4], it is imperative to replace the current practice of manual tuning to a desirable electronic configuration with a standardized automated method.

Realization of good qubit performance in QDs is achieved via electrostatic confinement of electrons in a two-dimensional electron gas (2DEG) present at the interface of semiconductor heterostructures using dynamically adjusted voltages on multiple electrical gates patterned on top of the device (see Figure 16 for an example of a false-color scanning electron micrograph of Si/SiGe quadruple QD). Reservoir gates (shown in purple in Figure 16) accumulate electrons into leads with stable chemical potential; depletion "screening" gates (shown in red in Figure 16) are used to define 1D transport channel in the 2DEG; barriers define the dot positions by locally depleting carriers within the 1D channel, thereby separating the electron density into disjoint regions while plungers shift the chemical potential in the dots relative to the chemical potentials of the contacts. In other words, the choice of gate voltages determines the number of dots, their position, and their coupling, as well as the number of electrons present in each dot.

The complete process of tuning an unknown QD device can be divided into a sequence of distinct phases shown in Figure 17. Over the past decade, there have been numerous attempts at automating each of them.

Bootstrapping and "sandboxing": Bootstrapping (or initialization) is a pre-tuning process that involves cooling the device down, making local sensing systems operational, and bringing the main device regime into an
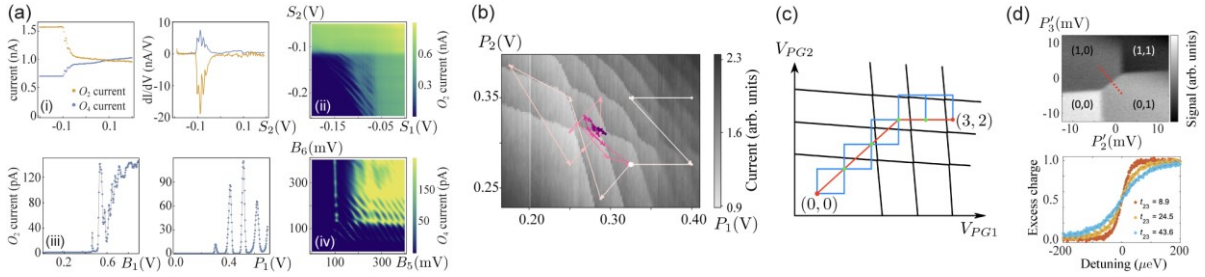
Figure 17. *The phases of the QD tuning flow. (a) Sample plots used in the bootstrapping and "sandboxing" phase. (i) Sweeps (left) and the numerical derivative of the sweeps (right) for the gate used to divide the 2DEG into two half-planes. (ii) Example plot of the qubit dot channel. The large rectangular regions to the right and top represent current flow underneath two screening gates. (iii) A series of sweeps for finger gates showing Coulomb blockade oscillations. (iv) The 2D scan in the space of the two barriers defining the sensor. (b) Visualization of a state tuning run in the space of two plunger gates, with arrows and the intensity of the color indicating the progress of the autotuner. (c) A cartoon of a possible path (red line) connecting the (0,0) and the (3,2) charge state. (d) (top) Charge stability diagram in the space of virtual gates with a red dotted line indicating the inter-dot detuning axis. (bottom) Excess charge extracted from a fit to the sensing dot signal as a function of detuning along the dotted line for three different t_{23}. Adapted from [5].*

appropriate parameter range for taking data (i.e., "sandboxing"). The usual cooling techniques rely on using a Helium-3 cryostat known commonly as a dilution refrigerator.

Once the device is cooled, the functionality testing can begin, where a series of 1D measurements ("sweeps") of individual gates and 2D measurements ("scans") is performed to check the gates and to ensure that two 1D transport channels are formed in the 2DEG: one where the qubit QDs will be formed and the other to host the charge sensing dots, see Figure 17(a-i, ii). Then, the quality assessment and characterization of finger gates takes place, see Figure 17(a-iii). During this stage, the acceptable (safe) ranges of voltages for coarse tuning are determined. The final step of bootstrapping is calibration of the charge sensor, that is, finding a set of voltages applied to the sensing dot gates that maximizes sensitivity of the conductance through the dot to changes in the local electrostatic potential, Figure 17(a-iv).

We are currently developing the bootstrapping tuning protocol. In the first attempt, we developed a script that automates all the pre-calibration steps typically done by an experimentalist. At present, we are refining the protocol to address difficulties we encounter during the first experimental test of the tuner.

Coarse tuning: Once the two 1D channels are formed and the charge sensor is ready, the next step in the tuning process is setting the QD device up in a stable global configuration of known topology in the state space with a known number of charge islands. The process of coarse tuning QD devices involves identifying the state of the device from a series of measurements, followed by adjustment of parameters (i.e., gate voltages) based on the observed outcomes. Figure 17(b) shows a sample 2D scan in the plunger gates space. The parallel charge transition lines (visible in the top right corner of Figure 17(b)) indicate that the device is in a single dot regime while the X-like features (visible in the center of Figure 17 (b)) indicate a regime where two dots are formed. In

general, the process of tuning heavily relies on a visual inspection and classification of image data by a human expert and, as such, does not scale well with the growing array sizes, is prone to random errors, and may result in only an acceptable rather than an optimal state.

In recent years, convolutional neural networks (CNNs) — a class of machine learning (ML) algorithms — have emerged as a "go to" technique for automated image classification, giving reliable output when trained on a representative and comprehensive dataset [6]. Taking advantage of the potential of CNNs, we have proposed an autotuning framework that combines a CNN-based algorithm trained on simulated charge sensor readout data with an optimization routine to eliminate the need for human intervention in tuning semiconductor QD devices [7]. Using a modified Thomas-Fermi approximation, we developed a model for electron transport in gate-defined QDs that mimics the transport characteristics and the charge sensor response of an experimental device [8]. For training purposes, we generated an assembly of 10 010 random charge sensor measurement realizations, with charge sensor response stored as (30×30) pixel maps from the space of plunger gates. By varying between simulations physical parameters of the system (e.g., the device geometry, gate positions, lever arm, and screening length), samples in the training dataset are representative of qualitative features across a wide range of devices. The labels for each simulated measurement are assigned based on the fraction of pixels within given realization in each of the possible states.

The coarse autotuning protocol has shown a lot of promise when tested *in situ*, with 85.7 % success rate over *N*=14 tuning runs [2]. However, since the training data was generated using idealized (i.e., noiseless) simulator, the experimental implementation required device-specific data processing to ensure compatibility with the CNN model. Still, when data was unusually noisy due to, e.g., instability of the charge sensor, the data processing was insufficient, resulting in the CNN
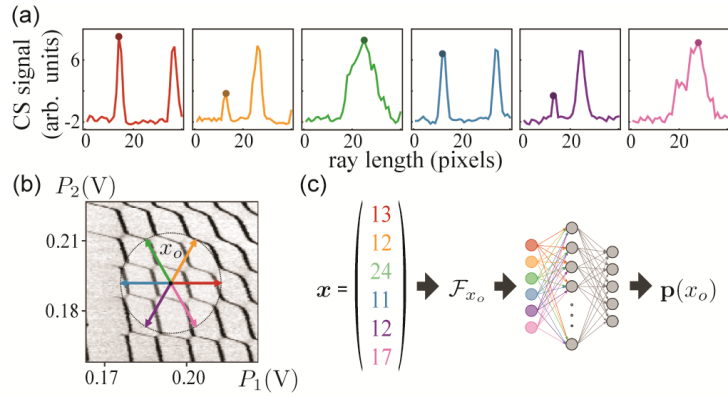
Figure 18. *(a) The processed charge sensor (CS) signal for six evenly distributed rays originating from a fixed point x_o in a double QD state. (b) The distribution of rays overlaying a 2D scan. (c) Visualization of the fingerprinting framework.*

returning an incorrect state vector and, ultimately, tuning failure.

To improve the autotuner performance, recently we incorporated several sources of physical noise into the training dataset. Moreover, to prevent failures due to unusually low data quality, we have extended the autotuning framework to include a data quality control (DQC) "gatekeeper" system that ensures that only reliable data is processed by the state classifier [9]. See page 90 for details. Data assessed by DQC to be of lower quality triggers one of two alternative actions: additional data analysis followed by targeted device recalibration or tuning termination. We are currently preparing for an experimental validation of the revised framework.

With the increasing number of dots, the number of gates that need to be controlled—and thus the number of 2D scans required to assess the state of the device—also grows. With the scalability to larger arrays in mind, we have recently proposed an alternative approach to assessing the state of the device, called the "ray-based classification" (RBC) [10]. Rather than using a full 2D scan capturing a small region of the voltage space, the RBC framework relies on a collection of evenly distributed 1D sweeps (called "rays") measured in multiple directions in the N-dimensional voltage space from a single point

$$x_o = \left(v_{B1}^{(0)}, v_{P1}^{(0)}, v_{B2}^{(0)}, v_{P2}^{(0)}, \ldots, v_{BN}^{(0)}, v_{PN}^{(0)}, v_{B(N+1)}^{(0)} \right),$$

to assess the relative position of transition lines surrounding x_o (see Figure 18 for an example of a sample point in double QD state with six evenly distributed rays). The qualitative information about the voltage space around x_o is encoded in the vector of distances to the nearest transition lines (the so-called "fingerprint"). The state of the device near point x_o is determined using the point fingerprint and a simple deep neural network (DNN) trained using simulated fingerprints data.

The RBC framework has been implemented experimentally showing performance on par with the more-

data-demanding CNN-based classification while requiring up to 70 % fewer measurement points. The RBC framework naturally extends to classifying convex polytopes in higher dimensions [11] which makes it an appealing measurement-cost-effective solution for differentiating between states of multi-QD devices. See page 86 for details.

Charge state tuning: The coarse tuning algorithm terminates once a target state, e.g., double-QD, is reached. However, the specific charge configuration of the system remains at that point unknown. Yet, the various QD qubit realizations — hybrid qubits, resonant exchange qubits, or quadrupolar exchange-only qubits — require a specific number of electrons. The goal of the charge tuning phase is to set the devices into a specific charge configuration.

Current charge sensor measurements provide only the overall direction of change of the number of charges sensed, but not the absolute number of charges. To date, all proposed algorithms aimed at setting a pre-defined charge configuration follow a two-step strategy. Starting within a voltage regime where a double QD is formed (i) empty the QD of all electrons and then (ii) load the desired number of electrons on each QD. Both steps rely on detecting and counting transition lines while electrons are being shuffled, one at the time.

Following the unloading / re-loading strategy, we are currently developing a ray-based approach to setting the desired charge state. In our implementation, the 2D scans are replaced by a series of ray-based measurements to enable more efficient tuning. See page 92 for details.

Establishing controllability: Ideally, changing voltages on a single gate should affect only the parameter it is designed to control (e.g., the electrochemical potential of a specific QD or the tunnel barrier between two adjacent QDs). However, the capacitive crosstalk between the various gate electrodes, i.e., a situation in which a signal on one gate causes a smaller version of the same signal on an adjacent gate because of the capacitance between them, makes it difficult (if not impossible) to vary just a single parameter without affecting the others.

One way to compensate for the capacitive crosstalk is to implement so-called "virtual gates", that is linear combinations of multiple QD gate voltages chosen to address only a single electrochemical potential or tunnel barrier. The utility of virtual gates has been shown in a number of multi-QD experiments, from showing controlled filling of an array of eight QDs using the "n+1 method" to demonstrating shuttling of a single charge through a 9-QDs charge stability space.

Fine tuning: Once the quantum dot system is in the desired configuration in terms of the topology and arrangement of charges, and the targeted controllability is established, there remains a significant additional step before the system can be used as a collection of quantum bits. The final, fine-tuning phase of calibrating QD devices focuses on adjusting the inter-dot tunnel coupling, which directly translates to control over the exchange coupling. As the tunneling rate becomes set, additional spin parameters must then be extracted, and in some cases, tuned. These include both the Zeeman splitting (Larmor precession frequency) and, where appropriate, estimates of the magnetic field gradient seen in each QD.

There are two main aspects to fine-tuning: ensuring that the system is behaving close to the qubit regime and that qubit performance and controllability can be achieved. The output of the fine-tuning system should, ideally, be a range of parameters that enable high fidelity quantum gates on the quantum bits stored within each QD or multi-QD qubit. Recently, we have begun investigating to what extent ML techniques can automate this final phase of tuning QD devices to work as qubits.

Summary: Working with experimental devices with high-dimensional parameter spaces poses many challenges, from performing reliable measurements to identifying the device state to tuning into a desirable configuration. By combining theoretical, computational, and experimental efforts, this interdisciplinary research sheds new light at how modern ML techniques can assist experiments. To use QD qubits in quantum computers, it is necessary to develop a reliable automated approach to control QD devices, independent of human heuristics and intervention. However, more work is needed to develop a fully automated tuner for real-life applications.

References

[1] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. K. Vandersypen. Spins in Few-Electron Quantum Dots. *Reviews of Modern Physics* **79** (2007), 1217–1265.

[2] J. P. Zwolak, T. McJunkin, S. S. Kalantre, J. P. Dodson, E. R. MacQuarrie, D. E. Savage, M. G. Lagally, S.N. Coppersmith, M. A. Eriksson, and J. M. Taylor. Autotuning of Double-Dot Devices *In Situ* with Machine Learning. *Physical Review Applied* **13** (2020) 034075.

[3] D. M. Zajac, T. M. Hazard, X. Mi, E. Nielsen, and J. R. Petta. Scalable Gate Architecture for a One-Dimensional Array of Semiconductor Spin Qubits. *Physical Review Applied* **6** (2016), 054013.

[4] U. Mukhopadhyay, J. P. Dehollain, C. Reichl, W. Wegscheider, and L. M. K. Vandersypen. A 2 × 2 Quantum Dot Array with Controllable Inter-Dot Tunnel Couplings. *Applied Physics Letters* **112** (2018), 183505.

[5] J. P. Zwolak and J. M. Taylor. *Colloquium*: Advances in Automation of Quantum Dot Devices Control. arXiv:2112.09362, 2021.

[6] A. Krizhevsky, I. Sutskever, and G.E. Hinton. Imagenet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems 25* (F. Pereira, C.J.C. Burges, L. Bottou, and K.Q. Weinberger, eds.), Curran Associates, Inc., 2012, 1097–1105.

[7] S. S. Kalantre, J. P. Zwolak, S. Ragole, X. Wu, N. M. Zimmerman, M. D. Stewart, and J. M. Taylor. Machine Learning Techniques for State Recognition and Auto-Tuning in Quantum Dots. *npj Quantum Information* **5**:6 (2019), 1–10.

[8] J. P. Zwolak, S. S. Kalantre, X. Wu, S. Ragole, and J. M. Taylor. QFlow Lite Dataset: A Machine-Learning Approach to the Charge States in Quantum Dot Experiments. *PLoS ONE* **13**:10 (2018), 1–17.

[9] J. Ziegler, T. McJunkin, E. S. Joseph, S. S. Kalantre, B. Harpt, D. E. Savage, M. G. Lagally, M. A. Eriksson, J. M. Taylor, and J. P. Zwolak. Toward Robust Autotuning of Noisy Quantum Dot Devices. Preprint at: *Physical Review Applied*, **17**: 024069, 2022.

[10] J. P. Zwolak, T. McJunkin, S. S. Kalantre, S. F. Neyens, E. R. MacQuarrie, M. A. Eriksson, and J. M. Taylor. Ray-Based Framework for State Identification in Quantum Dot Devices. *PRX Quantum* **2** (2021), 020335.

[11] B. J. Weber, S. S. Kalantre, T. McJunkin, J. M. Taylor, and J. P. Zwolak. Theoretical Bounds on Data Requirements for the Ray-Based Classification. *SN Computer Science* **3** (2022).

Participants

Justyna P. Zwolak (ACMD); Jacob M. Taylor, Thomas McJunkin (NIST PML); Sandesh S. Kalantre (University of Maryland); Mark A. Eriksson, Emily S. Joseph (University of Wisconsin – Madison)

Combinatorial Testing for Software Based Systems

In 1997, the Mars Pathfinder began experiencing system resets at seemingly unpredictable times soon after it landed and began collecting data. Fortunately, engineers were able to deduce and correct the problem, which occurred only when (1) a particular type of data was being collected, and (2) intermediate priority tasks exceeded a certain load, resulting in a blocking condition that eventually triggered a reset. Situations of this type are known as interaction faults. Many real-time failures of software-based systems have been traced to such faults. These are often insidious in that they may remain hidden until the unfortunate combination is encountered during system operation.

Combinatorial testing (CT) is a versatile methodology for detecting interaction faults. CT began as pairwise (2-way) testing in which all pairs of the test values for all pairs of test factors are checked. Thus, pairwise testing can detect faults involving single factors or interactions between two factors. CT is based on an empirical observation, referred to as the interaction rule, that while the behavior of a software system may be affected by many factors, only a few are involved in any given failure. NIST investigations of failures in actual systems have shown that while most faults involved a single factor or interaction between two factors, some faults involved three or more factors [1]. (A fault involving more than six factors has not yet been reported.) Thus, pairwise testing is useful, but it may not be adequate for detecting interaction faults involving more than two test factors.

More than a decade ago, NIST took the initiative to extend pairwise (2-way) CT to higher strength t-way CT for t > 2. NIST has helped make CT practical by developing research tools and techniques for generating combinatorial test suites. CT has now gained significant interest from the international software testing community. Many successful results from the use of CT in aerospace, automotive, and financial service industries, as well as defense, security, and electronic medical systems have since been reported.

Raghu Kacker

A suite of test cases for combinatorial *t*-way testing includes (covers) at least once all possible *t*-tuples of the test values for every set (combination) of *t* factors out of the complete set of all *k* factors that are tested ($k > t$). Use of mathematical objects called covering arrays makes it possible to check all *t*-tuples of the test values with a small number of test cases. Table 5 shows a covering array of 13 rows and 10 columns each having two possible value 0 and 1. Columns correspond to the factors and the rows correspond to the test cases. The

Table 5. *A covering array of 13 rows includes all eight triplets (000, 001, 010, 011, 100, 101, 110, and 111) of the possible values (0 and 1) for every one of the 120 possible sets of 3 out of 10 test factors represented by the columns (for example, see colored entries)*

		Columns								
Rows	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1
3	1	1	1	0	1	0	0	0	0	1
4	1	0	1	1	0	1	0	1	0	0
5	1	0	0	0	1	1	1	0	0	0
6	0	1	1	0	0	1	0	0	1	0
7	0	0	1	0	1	0	1	1	1	0
8	1	1	0	1	0	0	1	0	1	0
9	0	0	0	1	1	1	0	0	1	1
10	0	0	1	1	0	0	1	0	0	1
11	0	1	0	1	1	0	0	1	0	0
12	1	0	0	0	0	0	0	1	1	1
13	0	1	0	0	0	1	1	1	0	1

number of possible sets (combinations) of 3 out of 10 test factors is $(10 \times 9 \times 8)/(3 \times 2 \times 1) = 120$. When each factor has two possible values, each set of 3 factors can have $2^3 = 8$ possible triples of test values $((0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1), (1, 0, 0), (0, 0, 1), (1, 1, 0), (1, 1, 1))$. So, the total number of possible triples of values for all 10 factors is $120 \times 8 = 960$. A test suite based on Table 5 includes ("covers") at least once all 960 distinct triples of the test values of ten factors.

In practice, one wants a minimal covering array, that is an array which covers all possible *t*-tuples of the test values for every set of *t* out of all *k* factors with the least number of rows (test cases). In practice, many factors have dependencies and constraints, and hence not all combinations of the test values may be logically or physically valid. A combinatorial test suite must avoid such forbidden combinations. Generating minimal covering arrays that avoid forbidden combinations is a difficult computational problem [2]. A great deal of research has been done to develop mathematical and computational methods to generate minimal covering arrays of this type. NIST and its collaborators have developed several such algorithms.

NIST-Developed Tools. NIST has developed several research tools to make CT practical. ACTS (for Automated Combinatorial Testing for Software), which was developed in cooperation with the University of Texas at Arlington, includes several algorithms to generate high strength test suites for CT. The ACTS algorithms

are optimized to efficiently avoid forbidden combinations of test settings. More than 4265 have downloaded executable versions of the ACT algorithms from the NIST webpage for CT. (It is difficult to ascertain the number of users because some users have redistributed to others, and some are students who may have used it only once for a single project.)

A second research tool, CCM (for Combinatorial Coverage Measurement), developed jointly by NIST and a guest researcher from CENAM, the national metrology institute of Mexico, describes the incompleteness of a test suite that may not have been developed from a CT viewpoint. Basic combinatorial coverage measurements describe the incompleteness of a test suite relative to a test suite based on a covering array that includes all possible t-tuples of values for every t-factor combination for various values of t. The combinatorial deficiency of a test suite can be remedied by additional tests. Thus, CCM can help guide the expansion of a test suite to satisfy stated combinatorial requirements [3]. The latest version of CCM supports constraints which exclude forbidden combinations of values. A parallel processing version is also available.

Impact of NIST Research. NIST efforts have sparked a surge of research and application of combinatorial testing technology. A 2010 NIST Special Publication on CT was downloaded more than 30 000 times by the end 2014 [4]. In 2013, we published a book with CRC Press on this topic [5]. One of the first large-scale users that we worked with is a group at the U.S. Air Force Base in Eglin, Florida. The behavior of one of their systems depended on the sequential order of certain events. This led to the problem of testing sequences of events, which required development of new mathematical objects called *sequence covering arrays* [6, 7, 8]. Lockheed-Martin, a large U.S. defense contractor, reported (based on eight projects) that use of CT reduced cost of testing by about 20 % with 20 % to 50 % improvement in test coverage [9]. CT methods are now being used in diverse areas such as financial services, automotive, automation, avionics, video coding standards, and for security testing. The NIST webpage for CT cites more than 40 application papers. For testing a software-based system, no single approach is enough. Multiple approaches are generally needed at various stages development and installation. CT complements other approaches. It is now included in software engineering courses taught in more than 18 U.S. universities and 2 Austrian universities at least. NIST efforts on technology transfer of CT tools and techniques received the 2009 Excellence in Technology Transfer Award from the Federal Laboratory Consortium-Mid Atlantic Region.

CT has also gained significant interest from the research community. In 2012, NIST took lead in organizing a workshop on CT[2] in conjunction with the 2012 IEEE International Conference on Software Testing, Verification, and Validation (ICST), a premier conference in this field. Since then, an International Workshop on Combinatorial Testing (IWCT) has become an annual event for sharing advancements in CT tools and techniques, as well as results from practical industrial use of CT. The tenth such IWCT[3] was held on April 12, 2021, in conjunction with ICST 2021[4]. Four project participants (Kacker, Kuhn, Lei, and Simos) were among the co-organizers. The IWCT 2021 received 12 submissions, from which the Program Committee selected 8 papers (5 full papers and 3 short papers). The program contained research papers as well as on applications and use cases of CT.

Recent accomplishments.

Testing complex faults in digital devices in nuclear power applications. We investigated challenges in testing digital devices (with embedded real-time software) in critical nuclear power applications [10]. New methods and novel techniques are needed to adequately test such systems. Advancements in understanding the nature of complex faults, and applying this understanding in testing and verification, make it possible to build embedded cyber-physical systems that are safe and secure [11]. Our experiments and preliminary results indicate that pseudo-exhaustive testing based on combinatorial methods at all levels of interaction (variable, function, and thread interactions) on an embedded device can detect a diverse range of software flaws. We discovered several native defects in a code that were latent for some time. While the available research prototype tools are very capable, the main challenge is building the testbed architecture to conduct further investigations [12].

From fault localization to explaining image classifiers. The core of Artificial Intelligence (AI) based systems is a machine learning (ML) model that is used to perform tasks such as classification and prediction. Often ML models come as a black box to the user, leading to the problem of interpretability. Explainable Artificial Intelligence (XAI) is key to providing confidence and trustworthiness in ML based systems. We observe a fundamental connection between XAI and fault localization in software testing. (Fault-localization refers to detecting from the pass/fail results of testing an error (bug) that triggers a failure.) In fault localization, given a failing scenario, a software developer identifies which part of the input causes the failure. Similarly, in XAI, given a decision made by an ML model, we identify which features of the input causes the decision, in the sense that if these features are removed, then the decision would be different. We developed an approach that uses a CT-based software fault localization approach, called BEN,

---

[2] http://www.research.ibm.com/haifa/Workshops/ct2012/
[3] https://gist.nju.edu.cn/iwct2021/

[4] https://icst2021.icmc.usp.br/

to produce explanations for decisions made by ML models in image classifiers [13]. We report a preliminary evaluation of our approach using a popular (open source) image classifier called VGG16[5] and 50 randomly selected images from the ImageNet dataset. Our results indicate that for 44 (of 50) images, our approach can generate counterfactual explanations by removing a very small number of features in the input image.

Combinatorial approach to testing Deep Neural Network (DNN) used in Autonomous Driving Systems. In Autonomous Driving Systems (ADS), a Deep Neural Network (DNN) model is often used to perform tasks such as pedestrian detection, object detection, and steering control. These DNN models need to be rigorously tested with real-world driving scenarios to detect safety-critical bugs before deployment. We used a combinatorial testing (CT) based approach to generate test images based on image transformations for testing DNNs used in ADS [14]. First, we identified a set of basic image transformations that do not essentially change the "ground truth" of the image being transformed. That is, the prediction result for a transformed image produced by such a transformation is the same (up to a certain threshold) as the original image. Then, we used Combinatorial Testing (CT) methods to generate a t-way test set that covers every t-way combination of these transformations. Each test is a combination of transformations and can be used to create a test image. We consider inconsistent behaviors in two cases. In the first case, the ground truth of a test image remains the same as that of the original image. Thus, we consider that an inconsistent behavior is detected if the prediction result of a test image differs from that of the original image by an amount that is more than a threshold. In the second case, the ground truth of a test image may be different from that of the original image. In this case, we compare the prediction results of the same test image from different DNNs that perform the same prediction. An inconsistent behavior is detected if the prediction results of a test image from different models do not agree with each other. We report an experimental evaluation of our approach using three of the top five models, namely Autumn, Chauffeur, and Rambo from the Udacity self-driving challenge[6].The results indicate that t-way tests can help the practitioners to effectively test DNN models used in ADS.

DNNs consist of "neurons" that relate to each other. Each neuron, if activated, performs a small computation that contributes to the final decision/prediction made by a DNN. Covering a neuron means that we provide a test image that activates the neuron, which allows us to see the impact of the neuron (or more precisely, the computation performed by the neuron) on the final decision/prediction. To some extent, a neuron is like a

statement in the code of software, and neuron coverage is like statement (code) coverage. Achieving more neuron coverage allows us to cover more types of images, which in turn allows us to expose more potential faults in a DNN. DeepTest, is a state-of-the-art open-source tool[7] that aims at generating test images that maximize neuron coverage. We compared the neuron coverage achieved using our combinatorial approach against that achieved with DeepTest. Our combinatorial testing approach typically generates more test images than DeepTest. To facilitate a fair comparison, we use a random sampling approach to select a subset of t-way test images such that both approaches generate the same number of test images. The results suggest that our combinatorial approach performs better than DeepTest in terms of generating valid synthetic images and covering an additional number of neurons [15].

Combinatorically cross-site scripting web application firewalls. Cross-site scripting (XSS) is a common class of vulnerabilities in web applications. So, site operators often seek to protect their assets using web application firewalls (WAFs). These systems employ filtering mechanisms to intercept and reject requests that may be suitable to exploit XSS flaws and related vulnerabilities such as SQL injections. However, they generally do not offer complete protection and can often be bypassed using specifically crafted exploits. We evaluated the effectiveness of WAFs to detect XSS exploits [16]. We developed an attack grammar and used a combinatorial testing approach to generate attack vectors. We compared our vectors with conventional counterparts and their ability to bypass different WAFs. Our results show that the vectors generated with combinatorial testing perform equal or better in almost all cases. They further confirm that most of the rule sets evaluated in this work can be bypassed by at least one of these crafted inputs.

Combination frequency differencing. Combinatorial coverage measures (CCM) are used in a wide range of problems, including fault localization and for evaluating the adequacy of test inputs and input space models. More recently, CCM have been used in applications of artificial intelligence (AI) and machine learning (ML), for explainability and for analyzing aspects of transfer learning. These measures depend on the presence or absence of t-tuples of values in inputs test cases. We extended CCM to a new metric that we refer to as *combination frequency differencing* to compare training sets used in AI and ML algorithms [17].

Current projects.

Combinatorial methods for explainable AI: We continued our work on explainable AI and related

---

[5] https://arxiv.org/pdf/1409.1556.pdf
[6] https://github.com/udacity/self-driving-car

[7] https://arxiv.org/pdf/1708.08559.pdf

combinatorial methods by briefly reviewing recently introduced principles for explainable AI, examining their feasibility, and extending them where appropriate. We further view the combinatorial methods for explainable AI through the lens provided by these four principles, namely explanation, meaningful, explanation accuracy and knowledge limits. A paper has been submitted for a journal publication.

Security Testing of Smart Contracts: A blockchain is a decentralized computer system that uses cryptographic technologies to record information in a verifiable manner. A smart contract is a computer program that executes one or more transactions on a blockchain. When some conditions are met, the transactions are automatically executed, and their results are recorded on the blockchain in a way that cannot be changed. We investigated combinatorial fuzzing (a combination of CT and fuzzing) to improve code coverage and detection of vulnerabilities in testing smart contracts. Combinatorial Fuzzing requires testing tools. We developed a prototype research tool called MagicMirror for combinatorial fuzz testing of smart contracts. We used MagicMirror to test over two thousand real-world smart contracts. MagicMirror performs better than another academic research tool. A research paper has been submitted for presentation at a leading IEEE conference.

We have started a line of work that applies symbolic execution to test smart contracts. Symbolic execution is a powerful program analysis and testing technique, but its application for general software has been limited due to scalability concerns and difficulties to handle complex constraints and interaction with environments. Compared to general programs, smart contracts are small, have relatively simple constraints, and they are designed to execute in a closed environment. We believe that symbolic execution is a very promising technique for security testing of smart contracts. One challenge is that some functions in a smart contract can be difficult to reach when their execution depends on several other functions. This is because symbolic execution typically imposes a depth limit to handle the state explosion problem. We have developed a coverage-driven approach that uses the notion of a function dependency graph to guide the process of symbolic execution. We have built a prototype research tool called SmartExecutor that implements our approach. The experiment results indicate that our approach can significantly reduce the number of states to be explored without compromising code coverage and vulnerability detection results. We have reported our results in a submission to a leading IEEE conference.

Explainable Artificial Intelligence (XAI). XAI focuses on creating approaches and tools that can automatically provide explanations for the decisions made by Machine Learning (ML) models. XAI tries to answer the following two questions: Why does the model make a particular decision? What are the major factors that contribute to the decision? Providing explanations allows ML models to be interpreted, which is key to acceptance of AI-enabled technologies. We have started a new line of research that explores the use of a general software fault localization approach, called delta debugging, to produce counterfactual explanations. We used the delta debugging engine to guide and reduce the search space for generating counterfactual explains. We conducted a proof-of-concept study that compared our approach to a state-of-the-art research tool called DICE. The preliminary results are encouraging, where our approach outperforms DICE in terms of achieving higher proximity and sparsity of the produced explanations. We are in the process of conducting more experiments and are preparing a conference submission to report the results.

Combinatorial Security Testing for IoT systems. We continued our work on testing of Internet of Things (IoT) home automation systems using combinatorial methods for test case generation. We are investigating two testing methodologies. For each approach, we proposed creation of dedicated input parameter models of an IoT home automation system to be used with combinatorial test case generation strategies. Further, we devised an automated test execution framework and automated test oracles for evaluation purposes. We applied both methodologies to a real-world IoT system and evaluated the generated t-way test sets based on the derived input models. Our empirical testing evaluations revealed multiple errors in the tested devices for both methodologies. Moreover, we compared our results to a random testing approach, which we used as a baseline to benchmark the effectiveness of our newly devised IoT testing methodologies. We revised our previous work for a journal submission.

Browser Fingerprinting. We have investigated the applicability of combinatorial sequence testing to the problem of fingerprinting browsers based on their behavior during a Transport Layer Security (TLS) protocol handshake. We are continuing our work on a larger case study which includes more browsers and operating systems. The evaluation of our results using combinatorial techniques and tools from machine learning has been largely completed. We have compared our approach with other established fingerprinting techniques. We are working on a journal submission documenting our results.

Physical Unclonable Function. A PUF is a physical object that for a given input and conditions (challenge), provides a physically defined "digital fingerprint" output (response) that serves as a unique identifier, most often for a semiconductor device such as a microprocessor. PUFs are most often based on unique physical variations which occur naturally during semiconductor manufacturing. A PUF is a physical entity embodied in

a physical structure. Today, PUFs are usually implemented in integrated circuits and are typically used in applications with high security requirements, more specifically cryptography. We are investigating challenge-response relations in available PUFs data through combination frequency differencing (CFD) metrics for potential exploits by clever machine learning algorithms.

References

[1] D. R. Kuhn, D. R. Wallace, and A. M. Gallo, Jr. Software Fault Interactions and Implications for Software Testing. *IEEE Transactions on Software Engineering* **30** (2004), 418-421. DOI: 10.1109/TSE.2004.24

[2] L. Kampel and D. E. Simos. A Survey on the State of the Art of Complexity Problems for Covering Arrays. *Theoretical Computer Science* **800** (2019), 107-124. DOI: 10.1016/j.tcs.2019.10.019

[3] D. R. Kuhn, R. N. Kacker and Y. Lei. Combinatorial Coverage as an Aspect of Test Quality. *CrossTalk: The Journal of Defense Software Engineering*, March/April 2015, 19-2

[4] D. R. Kuhn, R. N. Kacker, and Y. Lei. Practical Combinatorial Testing, NIST Special Publication 800-142, October 2010, National Institute of Standards and Technology. URL: http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-142.pdf

[5] D. R. Kuhn, R. N. Kacker, and Y. Lei. *Introduction to Combinatorial Testing*. CRC Press, 2013.

[6] D. R. Kuhn, J. M. Higdon, J. F. Lawrence, R. N. Kacker and Y. Lei. Combinatorial Methods for Event Sequence Testing. In *Proceedings of the 5-th IEEE International Conference on Software Testing, Verification and Validation Workshops* (ICSTW), Montreal, Canada, April 17-21, 2012, 601-609. DOI: 10.1109/ICST.2012.147

[7] D. R. Kuhn, J. M. Higdon, J. F. Lawrence, R. N. Kacker and Y. Lei. Efficient Methods for Interoperability Testing using Event Sequences. *CrossTalk: The Journal of Defense Software Engineering*, July/August 2012, 15-18. URL: https://apps.dtic.mil/docs/citations/ADA566540

[8] Y. M. Chee, C. J. Colbourn, D. Horsley, and J. Zhou. Sequence Covering Arrays. *SIAM Journal of Discrete Mathematics* **27** (2013), 1844-1861. DOI: 10.1137/120894099

[9] J. Hagar, D. R. Kuhn, R. N. Kacker, and T. Wissink. Introducing Combinatorial Testing in a Large Organization. *IEEE Computer* **48** (2015), 64-72. DOI: 10.1109/MC.2015.114

[10] V. Jayakumar, S. Gautham, D. R. Kuhn, B. Simons, A. Collins, T. Dirsch, R. N. Kacker, and C. Elks. Systematic Software Testing of Critical Embedded Digital Devices in Nuclear Power Applications. In *Proceedings of the 31st International Symposium on Software Reliability Engineering Workshops* (ISSREW 2020), Online, October 12-15, 2020, 85-90. DOI: 10.1109/ISSREW51248.2020.00042

[11] A. Weiss, S. Gautham, A. V. Jayakumar, C. Elks, D. R. Kuhn, R. N. Kacker, and T. B. Preusser. Understanding and Fixing Complex Faults in Embedded Systems. *IEEE Computer* **54**:1 (2021), 49-60. DOI: 10.1109/MC.2020.3029975

[12] A. V. Jayakumar, D. R. Kuhn, B. Simons, A. Collins, S. Gautham, R. Hite, R. N. Kacker, A. D. Rajagopala, and C. Elks. Pseudo Exhaustive Software Testing Framework for Embedded Digital Devices in Nuclear Power. In *Proceedings of the 12th Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies* (NPIC&HMIT 2021), June 14-17, 2021, 812-823.

[13] J. Chandrasekaran, Y. Lei, R. N. Kacker, and D. R. Kuhn. A Combinatorial Approach to Explaining Image Classifiers. In *Proceedings of the 14-th IEEE International Conference on Software Testing, Verification, and Validation Workshop* (ICSTW-2021), Online, April 12-16, 2021, 35-43. DOI: 10.1109/ICSTW52544.2021.00019

[14] J. Chandrasekaran, Y. Lei, R. N. Kacker and D. R. Kuhn. A Combinatorial Approach to Testing Deep Neural Network-based Autonomous Driving Systems. In *Proceedings of the 14-th IEEE International Conference on Software Testing, Verification and Validation Workshops* (ICSTW), Online, April 12-16, 2021, 57-66. DOI: 10.1109/ICSTW52544.2021.00022

[15] J. Chandrasekaran, A. R. Patel, Y. Lei, R. N. Kacker, and D. R. Kuhn. Evaluation of T-Way Testing of DNNs in Autonomous Driving Systems. In *2021 IEEE International Conference on Artificial Intelligence Testing* (AITest), Online, August 23-26, 2021, 17-18. DOI: 10.1109/AITEST52744.2021.00013

[16] B. Garn, D. S. Lang, M. Leithner, D. R. Kuhn, R. N. Kacker, and D. E. Simos. Combinatorically XSSing Web Application Firewalls. In *Proceedings of the 14-th IEEE International Conference on Software Testing, Verification and Validation Workshops* (ICSTW), Online, April 12-16, 2021, 85-94. DOI: 10.1109/ICSTW52544.2021.00026

[17] E. Lanus, L. J. Freeman, D. R. Kuhn, and R. N. Kacker. Combinatorial Testing Metrics for Machine Learning. In *Proceedings of the 14-th IEEE International Conference on Software Testing, Verification and Validation Workshops* (ICSTW), Online, April 12-16, 2021. 81-84, DOI: 10.1109/ICSTW52544.2021.00025

Participants

Raghu N. Kacker, James F. Lawrence (ACMD), D. Richard Kuhn, M. S. Raunak (NIST ITL), Yu Lei (University of Texas at Arlington), Dimitris E. Simos (SBA-Research, Austria) , Eric Wong (University of Texas at Dallas) , Itzel Dominguez-Mendoza (CENAM, Mexico)

Part III

Project Summaries
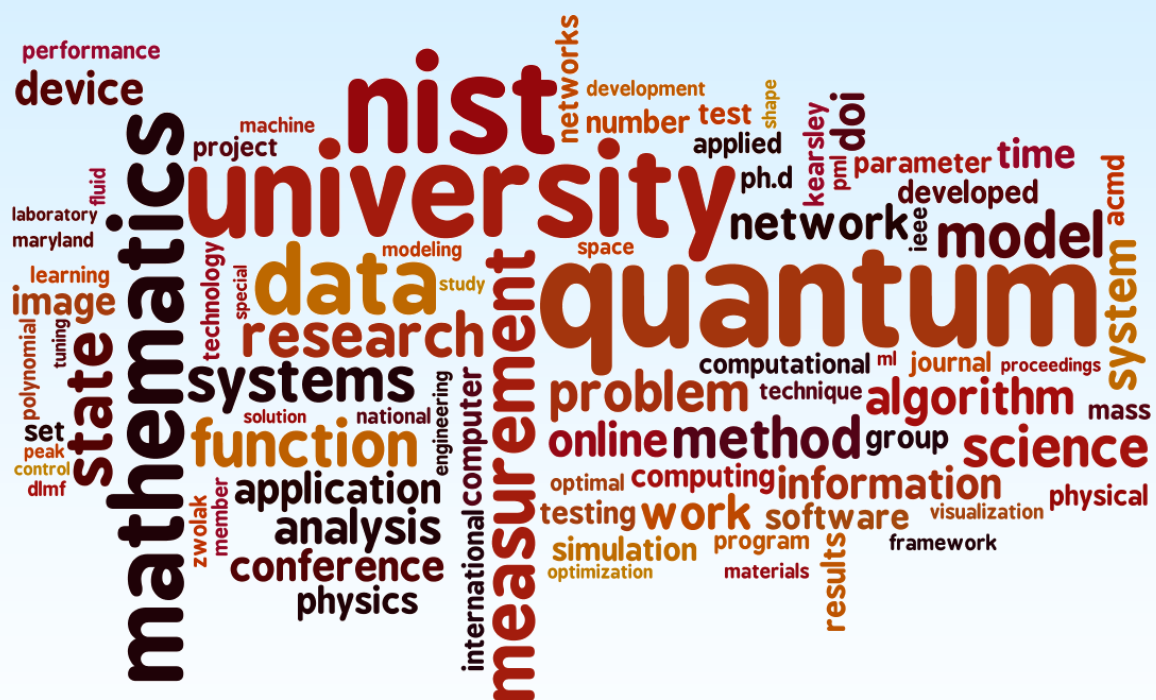
Mathematics of Metrology

Mathematics plays an important role in measurement science. Mathematical models are needed to understand how to design effective measurement systems and to analyze the results they produce. Mathematical techniques are used to develop and analyze idealized models of physical phenomena to be measured, and mathematical algorithms are necessary to find optimal system parameters. Mathematical and statistical techniques are needed to transform measured data into useful information. We develop fundamental mathematical methods and tools necessary for NIST to remain a world-class metrology institute, and to apply these to measurement science problems.

TOMCAT: X-ray Imaging of Nanoscale Integrated Circuits for Tomographic Reconstruction

Bradley Alpert
Dan Swetz (NIST PML)
Joseph Fowler (NIST PML)
Zachary Levine, et al. (NIST PML)
Kurt Larson (Sandia National Laboratory)
Edward Jimenez (Sandia National Laboratory)
Amber Dagel, et al. (Sandia National Laboratory)
Edward Garboczi (NIST MML)
George Barbastathis et al. (MIT)

As the leading semiconductor manufacturing techniques progress through 14 nm, 10 nm, and now 7 nm technology nodes, the ability to fabricate these chips has outrun the ability to image them. This limitation makes a variety of diagnostic needs much more difficult to satisfy. The NIST Quantum Sensors Group (PML), in collaboration with researchers at Sandia National Laboratory, is leading a project for IARPA's RAVEN (Rapid Analysis of Various Emerging Nanoelectronics) program to develop a small-laboratory capability to image integrated circuits by x-ray tomography. In distinction from other RAVEN projects, TOMCAT exploits a scanning electron microscope (SEM) rather than a synchrotron beamline and does not destroy the chip under test. This is enabled by the exquisite energy resolution of NIST-developed cryogenic microcalorimeter spectrometers, comprised of transition-edge sensors (TES), which are being extended to larger arrays (now to 3000 detectors), as well as with better individual-detector throughput (up to 1000 counts/s) and energy resolution (< 10 eV FWHM). The detectors measure fluorescent photons produced when SEM electrons strike a target following their differential attenuation by different materials in the chip.

A principal analysis challenge of the project, enabling tomographic structure recovery in this limited exposure angle, limited-photon regime, is the development of physics-assisted machine learning (PAML)
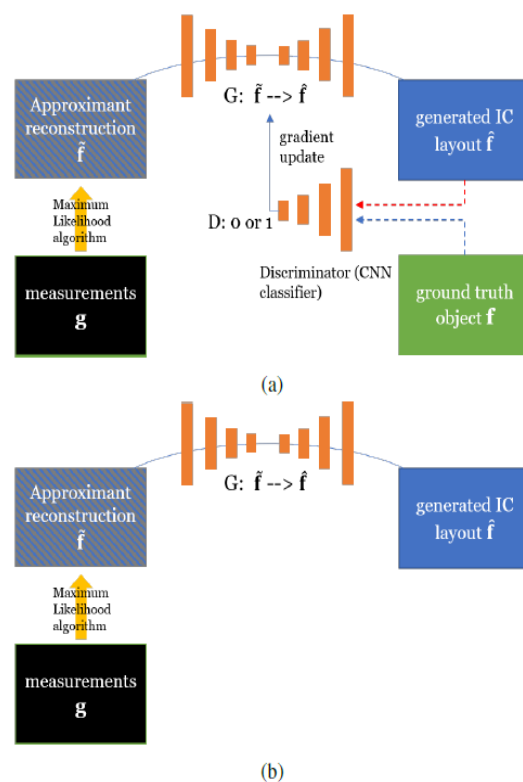
Figure 19. *(a) Supervised training of using modified conditional generative adversarial network uses simulated measurements, paired with their maximum-likelihood reconstructions, and the corresponding ground truth. (b) Testing is on pairs never seen during training.*

customized for the details of photon fluorescence, absorption, and scattering in this instrument configuration. George Barbastathis, who has had considerable success in PAML for optics, is leading this ML work. This year a significant advantage in photon efficiency was demonstrated from ML on simulated x-ray measurements of chip facsimiles generated at random from a simple model of the layered interconnect structure found in semiconductor circuitry.

Future work will generate simulated measurements from sections of actual chip specifications, scale to larger chip areas, and reconstruct shape representations designed both to improve fidelity and reduce the number of unknowns.
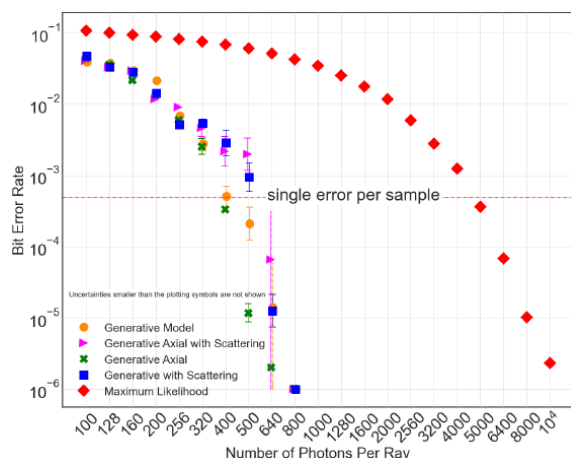
Figure 20. *Four trained generative models all achieve approximately an order of magnitude reduction in required photon counts relative to maximum-likelihood reconstruction.*
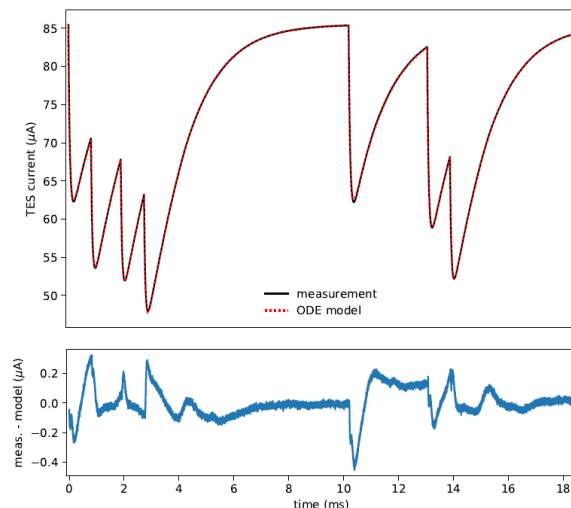
Figure 21. *Example of seven piled-up pulses, showing TES current drop from excitation by Mn x-ray fluorescence photons, with the ordinary differential equations (ODE) model (top) and the difference between the measurement and the model (bottom) with a root-mean-square difference of 1.3 x 10^-3 in relative current.*

[1] Z. Guo, J. K. Song, G. Barbastathis, M. A. Glinsky, C. T. Vaughan, K. W. Larson, B. K. Alpert, and Z. H. Levine. Advantage of Machine Learning over Maximum Likelihood in Limited-Angle Low-Photon X-Ray Tomography. In *Machine Learning for Scientific Imaging 2022 Conference*, IS&T Electronic Imaging, 2022.

[2] Z. Guo, J. K. Song, G. Barbastathis, M. A. Glinsky, C. T. Vaughan, K. W. Larson, B. K. Alpert, and Z. H. Levine. Physics-assisted General Adversarial Network for X-Ray Tomography. In review.

[3] J. W. Fowler, G. C. O'Neil, B. K. Alpert, D. A. Bennett, E. V. Denison, W. B. Doriese, G. C. Hilton, L. T. Hudson, Y.-I. Joe, K. M. Morgan, D. R. Schmidt, D. S. Swetz, C. I. Szabo, and J. N. Ullom. Absolute Energies and Emission Line Shapes of the L X-Ray Transitions of Lanthanide Metals. *Metrologia* **58** (2021), 015016. DOI: 10.1088/1681-7575/abd28a

[4] J. W. Fowler, B. K. Alpert, G. C. O'Neil, D. S. Swetz, and J. N. Ullom. Energy Calibration of Nonlinear Transition-Edge Sensors with Uncertainty Estimates from Gaussian Process Regression. In review.

True Becquerel: A New Paradigm for 21st Century Radioactivity Measurements

Bradley Alpert
Ryan Fitzgerald (NIST PML)
Dan Becker (NIST PML)
Denis Bergeron (NIST PML)
Richard Essex (NIST PML)
Kelsey Morgan (NIST PML)
Svetlana Nour (NIST PML)
Galen O'Neil (NIST PML)
Daniel Schmidt (NIST PML)
Gordon Shaw (NIST PML)
Daniel Swetz (NIST PML)
Mike Verkouteren (NIST MML)
Daikang Yan (NIST PML)

Expanding applications of radioactivity in medicine, energy, and security demand quantification of complex mixtures at uncertainty levels that are currently unachievable. This project will enable measurement of absolute activity (Bq) of radionuclide mixtures, avoiding chemical separation, by analysis of the decay heat signature of gravimetric samples embedded within microcalorimeter detectors. This capability consolidates multiple measurements into one, reducing cost and uncertainty. Success will create a primary realization of the Bq for direct assay of real-world samples at NIST and beyond, resulting in faster clinical trials of new radio-pharmaceuticals and a faster, expanded nuclear forensics "fingerprinting" method for improved decision making.

The project enters its second year of NIST Innovations in Measurement Science funding with initial success in dispensing, weighing, and placement of mg-quantity solutions, and initial ODE fitting of previously measured x-ray pulses in transition-edge sensor (TES) microcalorimeters. Depositing radionuclide solutions in nano-porous substrates for embedding in TES sensors designed for alpha decays and their measurement in new laboratory facilities lies ahead, along with characterization of high-rate radioactive decay events to generate an energy spectrum, and analysis of the spectrum to quantify constituents of the sample at the level of 0.1 % uncertainty.

The analysis challenges include (1) characterization of detector dynamics, to enable determination of decay

energies of events with poor temporal separation, avoiding detector dead time, at an accuracy that reflects the exquisite precision of the TES detector, (2) characterization of the partial energy losses due to transport out of the absorber material of alpha, beta, and gamma rays, and (3) disambiguation of the spectrum into constituents, based on a library of radionuclide decays, with full quantification. Two new tools for detector dynamics characterization are (a) fabricated capability for electronic excitation of the detector with known energy depositions, and (b) algorithms for ODE parameter sensitivity, of much recent attention, to determine an ODE system from its input/output behavior. This machine learning (ML) technique will be combined with more conventional supervised ML for library-based disambiguation of spectra. The uncertainty, and risk, for both techniques is whether the stringent accuracy requirements of the project can be achieved.

[1] R. P. Fitzgerald, B. K. Alpert, D. T. Becker, D. E. Bergeron, R. M. Essex, K. Morgan, S. Nour, G. O'Neil, D. R. Schmidt, G. E. Shaw, D. Swetz, R. M. Verkouteren, and D. Yan. Towards a New Primary Standardization of Radionuclide Massic Activity Using Microcalorimetry and Quantitative Milligram-Scale Samples. *Journal of Research of NIST*, to appear.

Predicting Normal Boiling Points of Chemical Compounds Using Graph Convolutional Neural Network

Chen Qu
Thomas Allison (NIST MML)
Barry I. Schneider
Walid Keyrouz (NIST ITL)
Anthony Kearsley

The normal boiling point (BP) of a substance is the temperature at which a liquid turns into a vapor at atmospheric pressure. It is of great importance to physicists and chemists since BP is routinely used as a measure of purity of chemical substances sold as reagents or after synthesis procedures. Therefore, prediction of BPs using a variety of methods continues to be of significant interest. It has an interesting history dating back hundreds of years.

Notably, Stein and Brown fit a group contribution model with a mean absolute error of 20.4 K (4.3 %) when this model was tested against 6584 compounds not used in fitting [1]. This model was incorporated into the United States Environmental Protection Agency's EPI Suite program that predicts a number of properties for a variety of compounds using molecular structure information.[2]
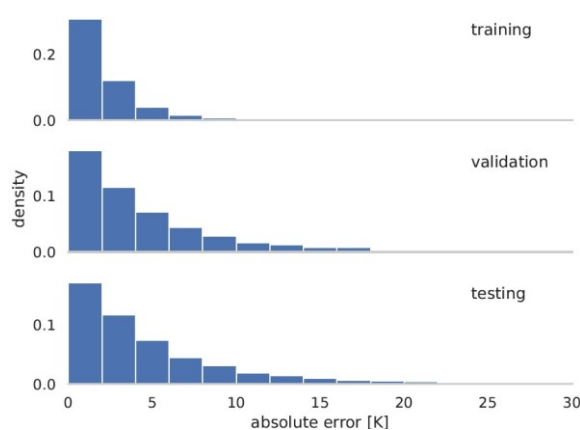
Figure 22. *Plots of experimental versus predicted value of the normal boiling point for the training, validation, and testing data sets.*

In FY 2020, we developed a deep-learning model to predict Kováts retention indices. The graph neural network (GNN) model used there performed well on that problem [3]. Given the nearly linear correlation between retention times and boiling points, we expect that the same methodology will perform well on BP, and we successfully trained a good predictive model for the BP. A brief summary of the methodology is presented below. The details appear in [4].

The data used in this study was extracted from the literature as captured by the NIST Thermodynamics Research Center (TRC) SOURCE Data Archival System [5, 6] This database contains more than 12 000 experimental determinations of the normal boiling point for a wide variety of molecules, providing a robust data set for training a new predictive model for BP. The first step was to remove points which were outside of the molecular set under consideration as well as outliers where multiple measurements of the BP for a single molecule were available. The final set consisted of 3 850 molecules for training and testing the machine learning model.

Our machine learning model is based on the materials graph network (MEGNet) approach developed by Chen et al. [7] This model has been tested for a variety of chemical properties on both molecular and crystalline systems, including the prediction of Kováts retention indices with excellent results [3]. The MEGNet methodology incorporates a graph network architecture that captures molecular structure in a very natural way. In a GNN, the atomic centers correspond to vertices in the graph, and chemical bonds correspond to graph edges.

The inputs to the MEGNet model are atom, bond, and global features of a molecule, obtainable from a 2D MolFile representation of the molecules using the RDKit package. Then these features were successively updated using a "message-passing" network. Finally,

Table 6. *Summary statistics of the 10-fold cross validation procedure. The mean value and standard deviation of the mean absolute error (MAE) and the root mean square error (RMSE) over 10 runs is given for each of the 3 sets used.*

Set	MAE	RMSE
Training	2.30 ± 0.44	2.80 ± 0.51
Validation	5.56 ± 0.52	7.78 ± 1.66
Testing	5.77 ± 0.40	7.81 ± 0.94

the feature vector of the graph representation of the molecule was extracted via a readout layer and was used for the final prediction of the BP.

Training was carried out with hyperparameters that were found to be reasonable as a result of empirical testing during the training phase of the project and in the previous model for retention indices [3]. The model performance was assessed using 10-fold cross validation. The error performance of the machine learning model is depicted in Figure 22. The figure shows that the error in the training set is smaller than the validation and testing sets, which exhibit longer "tails" in their error distributions. This is indicative of overfitting. The fact that the latter error distributions are so similar indicates that the validation set was effective during training, resulting in a model with good predictive capabilities.

The overall statistics of our model from the 10-fold cross-validation are shown in Table 6, with a mean absolute error of 5.77 K and a sample standard deviation of 7.81 K for the testing set.

The current model is performing about a factor of two better than the method of Stein and Brown as implemented in the EPI Suite package [1, 2]. The mean absolute error of the Stein and Brown method over all compounds in our data set is 11.84 K with a median error of 7.99 K and a standard deviation of 12.84 K.

Another significant achievement of our GNN model is that it was capable of detecting errors in the data set. We took note of molecules that consistently exhibited large errors in the training set and then went back to the source literature. This procedure identified nearly 100 errors in the input data set, i.e., we found our model to be effective in detecting true errors in the literature as well as errors resulting from transcriptions of the data.

In summary, we have developed a graph neural network model to predict normal boiling points for a diverse collection of molecules. Our model reproduces the data well, with mean absolute errors of less than 6 K in a 10-fold cross-validation procedure. During our study, we found that the graph neural network model was capable of detecting errors in the source data, and we used this to make corrections to the original data set. Based on our study, we feel that more accurate models will require a larger set of source data with even better accuracy.

[1] S. E. Stein and R. L. Brown. Estimation of Normal Boiling Points from Group Contributions. *Journal of Chemical Informatics and Computational Science* **34** (1994), 581–587.

[2] United States Environmental Protection Agency, Estimation Programs Interface Suite, v 4.11, Washington, DC, USA, 2021.

[3] C. Qu, B. I. Schneider, A. J. Kearsley, W. Keyrouz, and T. C. Allison. Predicting Kováts Retention Indices using Graph Neural Networks. *Journal of Chromatography A* **1646** (2021), 462100.

[4] C. Qu, A. J. Kearsley, B. I. Schneider, W. Keyrouz, and T. C. Allison. Graph Convolutional Neural Network Applied to the Prediction of Normal Boiling Point. *Journal of Molecular Graphics and Modelling*, to appear.

[5] A. Kazakov, C. Muzny, K. Kroenlein, V. Diky, R. Chirico, J. Magee, I. Abdulagatov, and M. Frenkel. NIST/TRC Source Data Archival System: The Next-Generation Data Model for Storage of Thermophysical Properties. *International Journal Thermophysics* **33** (2012), 22–33.

[6] NIST Thermodynamics Research Center, https://www.trc.gov/, 2021. Last accessed: Sep. 23, 2021.

[7] C. Chen, W. Ye, Y. Zuo, C. Zheng, and S. P. Ong. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chemistry of Materials* **31** (2019), 3564–3572.

Numerical Scaling for Preprocessing Mass Spectrometry Library Data

Danielle C. Brager
Edward P. Erisman (NIST MML)
Arun S. Moorthy (NIST MML)
Anthony J. Kearsley

Mass spectrometry is a core technique for elucidating the structure and identity of compounds. After processing, the output of a mass spectrometry analysis is a mass spectrum which is a histogram displaying the intensity of ions vs. mass-to-charge ratio. In this work, we focus exclusively on mass spectra obtained through electron ionization (EI) mass spectrometry which allows us to represent the mass spectra as vectors of equivalent length where the index (or location) of each element in the vector is the mass-to-charge of an ion and the value of each element is the corresponding relative abundance. Given a collection of mass spectra for known compounds, an analyst can infer the identity of unknown compounds using a common tool called mass spectral library searching. During this process, the mass spectrum of an unknown compound is compared with the library spectra using a similarity measure resulting in similarity scores between the query spectrum and each library spectra. Then the library elements are assembled

into a hitlist in order of similarity with the query spectrum where, ideally, top hitlist entries are likely candidates. Finally, the analyst uses the information from the hitlist to infer the identity of the compound.

Oftentimes, hitlists are "congested" which means that top hitlist candidates have similarity scores that only differ by a small percentage making the selection of a candidate spectrum more complicated. We have developed a mathematical pre-processing procedure where a scaling matrix, A, is constructed to maximally separate mass spec similarity scores, thus alleviating congestion in the hitlists. The scaling matrix is generated by solving a nonlinear optimization problem. That is, a nonlinear function is optimized, and the solution yields the optimal scaling matrix that maximally separates the hitlist. The problem is parameterized in a way that allows for minimal constraints, and the formulation always admits a solution. When the hitlist similarity scores are naturally well separated, the weighting is "small" and does not change the resulting scores very much. Equivalently, the optimal weighting matrix is close to the identity matrix of appropriate size. However, when scores are congested, the optimal weighting matrix will change the inter-spectral distances significantly. There are a variety of objective functions that can be formulated depending on desired construction of the scaling matrix.

We demonstrate this procedure using a spectral library consisting of 170 mass spectra of Fentanyl related compounds and separately collected as well as synthetically constructed query spectra. While various similarity measures are used in practice such as Stein and Scott [4], cosine similarity [2], and the hybrid match factor [3], we began with the intuitive Euclidean distance. We compare the Euclidean distance with an optimally scaled l_2 dissimilarity measure using the computed matrix, A. We produced scaled hitlists as well as directly calculated hitlists and compared them to each other. The order of hitlists change, sometimes slightly when scores are already well-separated and sometimes significantly, typically when scores are congested. We visualize the resulting scaled separation in the library using multidimensional scaling (MDS) [1] which is a technique used to represent pairwise measures of similarity amongst a set of n objects as a set of n points in a lower-dimensional Euclidean space while also preserving the relationship between the objects as best as possible. We

are currently implementing this procedure using different similarity measures including Stein and Scott as well as the cosine similarity.

[1] A. J. Kearsley, R. A. Tapia, and M. W. Trosset. The Solution of The Metric STRESS and SSTRESS Problems in Multidimensional Scaling Using Newton's Method. *Computational Statistics* **13**:3 (1998), 369–396.

[2] A. S. Moorthy and A. J. Kearsley. Pattern Similarity Measures Applied to Mass Spectra. In *Progress in Industrial Mathematics: Success Stories: The Industry and the Academia Points of View* (M. Cruz, C. Parés, and P. Quintela, eds.) ICIAM 2019, Valencia Spain. 43-54.

[3] A. S. Moorthy, W. E. Wallace, A. J. Kearsley, D. V. Tchekhovskoi, and S. E. Stein. Combining Fragment-Ion and Neutral-Loss Matching during Mass Spectral Library Searching: A New General-Purpose Algorithm Applicable to Illicit Drug Identification. *Analytical Chemistry* **89**:24 (November 20, 2017), 13261-13268.

[4] S. E. Stein and D. R. Scott. Optimization and testing of mass spectral library search algorithms for compound identification. *Journal of the American Society for Mass Spectrometry* **5**:9 (1994):859–866.

Grassmannian Shape Representations for Aerodynamic Applications

Zachary J. Grey
Olga Doronina (NREL)
Andrew Glaws (NREL)
Ryan King (NREL)

Airfoil shape design is a classical problem in engineering and manufacturing. Our motivation is to combine principled physics-based considerations for the shape design problem with modern computational techniques informed by data-driven approaches. Traditional analyses of airfoil shapes emphasize a flow-based sensitivity to deformations which can be represented generally by affine transformations (rotation, scaling, shearing, translation). Our work focuses on novel representation of shapes which replaces affine-style deformations with a rich set of data-driven deformations over a submanifold of the Grassmannian. The Grassmannian representation,
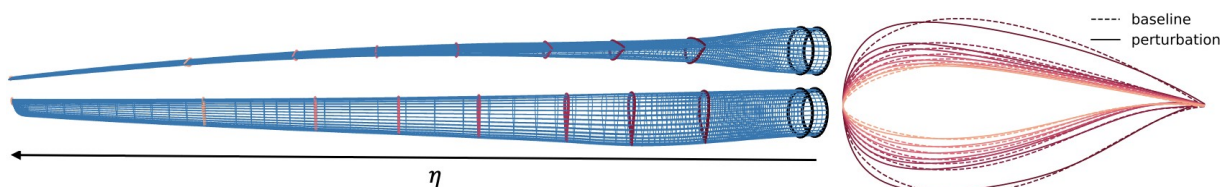
Figure 23. *IEA 15MW blade interpolation. The sparse nominal data (colored cross-sections with thicker line style) are interpolated along a piecewise geodesic construction over the Grassmannian. The refinements along the interpolation are shown as the blue wireframe interpolating a diverse set of shapes form circles through a variety of distinct airfoil classes. Moreover, the blade can be perturbed in a consistent manner (same perturbation for each airfoil cross-section) resulting in smooth deformations (colored curves to the right).*
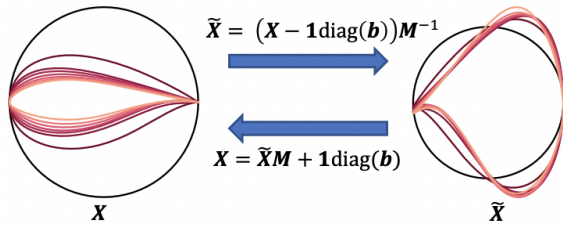
Figure 24. *Landmark affine (LA) standardization used to supplement computations over the Grassmannian and standardize scale variations in the shape to be as-circular-as-possible.*

informed by a database of physically relevant airfoils, offers (i) a rich set of novel 2D airfoil deformations not previously captured in the data, (ii) improved low-dimensional parameter domain for inferential statistics informing design/manufacturing, and (iii) consistent 3D blade representation and perturbation over a sequence of nominal shapes.

Many AI-aided design and manufacturing algorithms rely on shape parametrization methods to manipulate shapes to study sensitivities, approximate inverse problems, and inform optimizations. Two-dimensional cross-sections of aerodynamic structures such as aircraft wings or wind turbine blades, also known as airfoils, are critical engineering shapes whose design and manufacturing can have significant impacts on the aerospace and energy industries. Research into AI and ML algorithms involving airfoil design for improved aerodynamic, structural, and acoustic performance is a rapidly growing area of work.

While airfoil shapes can appear relatively benign, their representation and design are complex due to their extreme operating conditions in use and the highly sensitive relationship between deformations to the shape and changes in aerodynamic performance. The current state-of-the-art for airfoil shape parametrization is the class-shape transformation (CST) method. In this approach, the upper and lower surfaces of an airfoil are each defined using a class function to set the general class of the geometry to an airfoil, and a shape function that usually takes the form of a Bernstein polynomial expansion to describe a specific shape. The coefficients in this polynomial expansion are typically treated as tuning parameters to define new airfoil shapes.

However, defining a meaningful design space of CST parameters across a collection of airfoil types is difficult. That is, it is challenging to interpret how modified CST parameters will perturb the shape and thus difficult to contain or bound CST parameters to produce "reasonable" aerodynamic shapes. Furthermore, CST representations couple large-scale affine-type deformations—deformations resulting in significant and relatively well-understood impacts to aerodynamic performance—with undulating perturbations that are of increasing interest to airfoil designers and manufacturers across industries. This coupling between physically

meaningful affine deformations and undulations in shapes resulting from higher-order polynomial perturbations complicates the design process.

In this work, we explore a data-driven approach that uses a Grassmannian framework to represent airfoil shapes. The resulting set of deformations to airfoil shapes is independent of the very important, and often constrained, affine deformations. Modern airfoil design often incorporates constrained design characteristics of twist (or angle-of-attack) and scale which must be fixed or treated independently of higher-order deformations to a shape such as a rich set of changing inflections. Our approach decouples these two aspects of airfoil design and offers new interpretations of a space of shapes, not previously considered.

A shape can be represented as a boundary defined by a closed injective curve, $c : [0,1] \rightarrow \mathbb{R}^2$. In a computational setting, we represent the 2D airfoil shape as a discrete ordered sequence of *landmarks* $(x_i) \in \mathbb{R}^2$ for $i = 1, \dots, n$. That is, given some curve $c(s)$, we have landmark points $x_i = c(s_i)$ for $0 \leq s_1 < s_2 < \cdots < s_n \leq 1$. Moving along the curve, this discrete sequence of planar vectors defining the airfoil shape can be represented as a matrix $X = [x_1, \dots, x_n]^\top \in \mathrm{R}_*^{n \times 2}$. Important and often rigorously designed deformations of the shapes can be written as affine transformations applied to all landmarks, $M^\top x_i + b$. We introduce a scaling of landmark data, depicted in Figure 24, to parametrize new shape deformations over the Grassmann manifold (Grassmannian). The nature of this revised parametrization "divides out" the effect of the first order matrix operations, XM, thus preserving the affine transformations (combined with a translation invariance) and offering a richer set of parameters describing independent shape deformations.

The innovative characteristic of the proposed approach is representing airfoil shapes as elements of a Grassmann manifold $\mathcal{G}(n, 2) = \mathrm{R}_*^{n \times 2}/GL_2$ paired with a corresponding affine transformation (invertible 2-by-2 matrices and translation) representing a subset of rotation, scaling, and shearing shape deformations. This reinterpretation of the airfoil shape gives rise to a *separable parametrization* making important subsets of deformations independent, allowing designers to make interpretable and systematic changes to airfoil shapes. For example, one may seek to preserve average airfoil scale characteristics M while independently studying all remaining deformations as perturbations over the Grassmannian. Computationally, these new parametrizations become

$$X_0(t, l) = \tilde{X}(t)M(l)$$

where $\tilde{X}(t)$ are representative elements parametrized over coordinates t of a data driven Grassmannian submanifold and $M(l)$ are the separate set of important scale variations as prescribed by airfoil designers. This defines a pair of independent vectors of parameters (t, l)
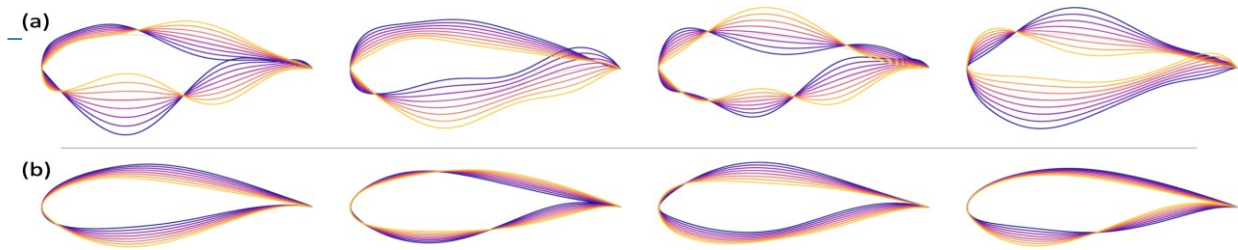
Figure 25. *Four colored parameter sweeps (columns) over two different design spaces (a) and (b). The colors correspond to the distance along the continuous sweep through parameter space resulting in a new airfoil for each color. (a) state-of-the-art CST parameter sweeps linearly interpolating four random pairs of extremal polynomial coefficients inferred from fits over a diverse set of distinct airfoil classes. (b) Grassmannian random parameter sweeps along four geodesics spanning the same data. The Grassmannian parametrization shows improvement in the regularization of connected shapes while CST parameter sweeps result in non-physical (self-intersecting) shapes with undesirable oscillations.*

describing any airfoil shape up to translations. This separability is a desirable characteristic for aerospace/wind-turbine designers who regularly prescribe or fix $M(l) \in GL_2$ but seek undulations $[\tilde{X}](t) \in \mathcal{G}(n, 2)$ in the shape to define a richer set of airfoil perturbations.

The utility in the novel representation is the separable form of the airfoil parametrization such that changes in t are independent of changes in l. Consequently, several advantages are introduced by this representation, including:

- Dimension reduction using Principal Geodesic Analysis to define a dominant submanifold of "undulating" airfoil perturbations from data.

- Improved parameter domain offering regularized shape perturbations devoid of undesirable oscillations in the shape as depicted in Figure 25.

- 3D blade representation, as shown in Figure 23, by interpolation of a sequence of discrete LA standardized airfoils (matrices) over piecewise geodesics of the Grassmannian with subsequent independence against designed affine deformations (twist, scale, and translation).

- Consistent 3D blade perturbations via parallel transport along blade interpolation as shown by the colored curves (right) in Figure 23.

The efficiencies gained by this systematic approach to decouple various types of deformations can be used to describe a 3D wind-turbine blade (Figure 23) using only *five total parameters*. These innovations have enabled improved and significantly more tractable inverse designs via machine learning for next generation wind-turbine blades. This work has been accepted to the AI for Design and Manufacturing Conference [1].

[1] O. A. Doronina, Z. J. Grey, and A. Glaws. Grassmannian Shape Representations for Aerodynamic Applications. To appear. URL: https://arxiv.org/abs/2201.04649

Large Scale Dynamic Building System Simulation

Anthony Kearsley
Amanda Pertzborn (NIST MML)
Aaron Chen (Drexel University)
Jin Wen (Drexel University)

The building sector represents the largest primary energy-consuming sector in the United States, responsible for 41 % of the country's primary energy, in comparison to 28 % for the transportation sector and 32 % for industry. Moreover, buildings consume 74 % of the electricity in the United States, which makes the building sector significant to the overall smart grid infrastructure. Given the rapid development of the smart grid and the potential of buildings to store and generate electricity through demand shifting and transactive control, there is an urgent need to improve the dynamic interactions between buildings and the smart grid, which further calls for robust and accurate dynamic building energy system modeling and simulation.

Traditional dynamic building simulations typically focus on a single building. However, those used for smart grid applications must simulate large and complex building energy systems and their interactions among building clusters that are composed of multiple buildings. These result in large systems of coupled, potentially ill-conditioned nonlinear systems (easily including thousands of equations), that will need to be solved rapidly. Thus, efficient, robust, and accurate solution of large sparse nonlinear algebraic and differential equation systems is becoming more and more essential to meet the demands to simulate large scale multiple building heating ventilation and air conditioning (HVAC) scenarios that are coupled to various complex energy sources either through the smart grid or other means, such as district heating/cooling.

In practice, these simulations are decomposed into geometric zones (see Figure 27) where one expects air flow and thermal conductivity to be similar. Typically, these zones are comprised of rooms or collections of
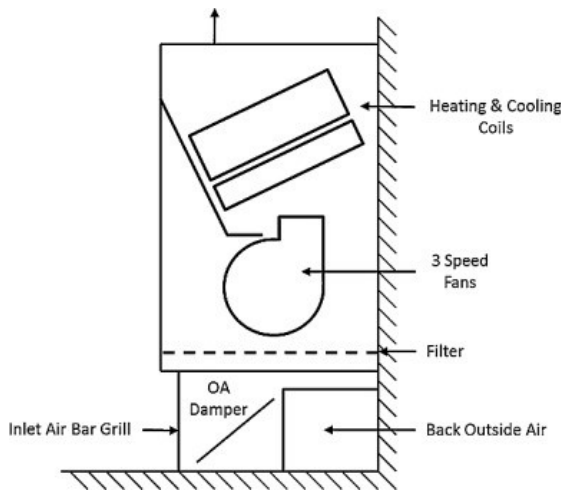
Figure 26. *A simple fan/coil unit schematic.*
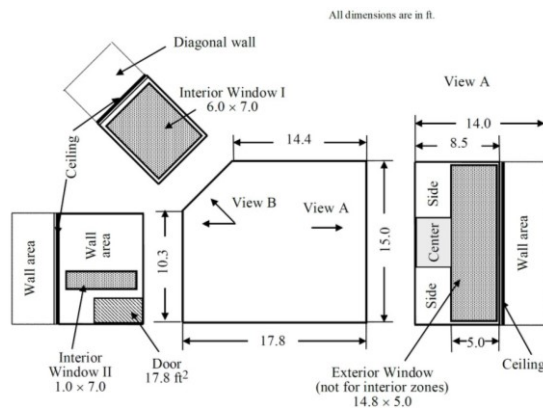
Figure 27. *A floor and window view of a typical zone.*

closely connected rooms. HVAC systems control temperature and air flow in these zones through coupled boundary conditions that are matched in what is called a "super-block." It is interesting that something as simple as a fan/coil unit (see Figure 26) can result in an approximation whose discretized system is quite hard to solve. More detailed mathematical models which would more accurately predict the physics of a simple fan coil unit would result in simulations that would be prohibitively expensive. Less detailed mathematical models yield more well-conditioned systems that can be solved quickly but do not provide the flexibility of, for example, more complicated scenarios through various degrees of occupancy, seasons or other factors that must be considered.

In [1], Chen develops a class of large-scale Krylov methods for solving these nonlinear systems using automatically built preconditioners. A numerical survey of the performance of these methods was completed this year, [5, 6]. Comprised of smaller models, like the one used to predict fan/coil dynamics, the resulting large-scale systems are assembled at the same time as a preconditioner, both of which would be made available to a

building engineer seeking to simulate and evaluate a building HVAC system. Not surprisingly, poor variable scaling [2] can also lead to a deterioration in numerical performance but results presented in [1] suggest that pre-conditioning of inexact Krylov methods can outperform more traditional dense solution techniques [3] as problem sizes grow. Figure 28 demonstrates that the numerical method proposed in [1], here called INB-PSGMRES(m), outperforms Powell's hybrid method (PH) [4] irrespective of seasonal changes. Here m denotes the dimension of the Krylov subspace, traditionally a parameter that is tuned. Of course, the comparison of CPU time is not exactly a straightforward one. The PH method was not designed for problems of this size. However, it provides strong evidence that methods like those proposed in [1, 5] are very much appropriate for these simulations.

[1] Z. Chen. *Advanced Solver Development for Large-Scale Dynamic Building System Simulation*. Ph.D. Thesis, Drexel University, Philadelphia PA, 2019.

[2] Z. Chen, J. Wen, A. J. Kearsley, and A. J. Pertzborn, Scaling Methods for Dynamic Building System Simulation in an HVACSIM+ Environment. In *Proceedings of the 15th International Building Performance Simulation Association (IBPSA) Conference*, San Francisco, CA, US, 2017, 2059-2065.

[3] S. Pourarian, A. J. Kearsley, J. Wen, and A. Pertzborn. Efficient and Robust Optimization for Building Energy Simulation. *Energy and Buildings* **122** (2016) 53-62.

[4] M. J. Powell. A Hybrid Method for Nonlinear Equations. *Numerical Methods for Nonlinear Algebraic Equations* **7** (1970), 87-114.

[5] Z. Chen, J. Wen, A. J. Kearsley, and A. Pertzborn. Evaluating the Performance of an Inexact Newton Method with a Preconditioner for Dynamic Building System Simulation. *Journal of Building Performance Simulation* **15**:1 (2022), 112-127.
DOI: [10.1080/19401493.2021.2007285](10.1080/19401493.2021.2007285)

[6] Z. Chen, A. Kearsley, J Wen, and A. Pertzborn. Smoothing Techniques in Dynamic Building System Simulation. In *2021 International Conference on Instrumentation, Control, and Automation* (ICA) Bandung, Indonesia, August 25-27, 2021. DOI: [10.1109/ICA52848.2021.9625664](10.1109/ICA52848.2021.9625664)

Figure 28. *A comparison of required CPU-time for both Powell's Hybrid Method (PH) and the more complicated technique proposed in [1].*

Numerical Solutions of the Time Dependent Schrödinger Equation

Barry I. Schneider
Heman Gharibnejad(Computational Physics Inc.)
Luca Argenti (University of Central Florida)
Juan Martin Randazzo (CONICET, Argentina)
Nicholas Douguet (Kennesaw State University)
Jeppe Olsen (Aarhus University)
Ryan Schneider (U. of California San Diego)

We are developing numerically robust methods for solving the time-dependent Schrödinger equation (TDSE). There are three related research threads underway.

1. Developing a hybrid basis set approach using B-spline/finite element discrete variable (FEDVR) and Gaussian type orbitals orbital to treat the interaction of attosecond (10^{-18} sec) radiation with molecular targets.

2. Computing the required hybrid basis function matrix elements via a novel 3D numerical grid based on overlapping atomic grids.

3. Examining the performance of various numerical time propagation techniques for the TDSE.

The hybrid basis function approach that has been developed is quite general, but the applications to date have concentrated on describing the single and double ionization of electrons exposed to intense, ultra-fast, laser radiation in many-body atomic and molecular systems. These attosecond (10^{-18} sec) pulses provide a new window to study the electronic motion in atoms and molecules on their natural timescale. To put this in context, the motion of electrons responsible for chemical binding and electron transfer processes in nature have a characteristic timescale of about 100 attoseconds. (It takes an electron 152 attoseconds to go around the hydrogen atom.) These processes can only be described using time-dependent quantum mechanics. Where appropriate, this needs to be coupled to Maxwell's equations to describe macroscopic phenomena. Our overall goal is to image quantum phenomena with sub-femtosecond temporal and sub-Angstrom spatial resolution and to provide coherent control of electron dynamics. Eventually, one can contemplate producing "molecular movies" of this motion in much the same way as it is done in molecular dynamics simulations of heavy particle processes.

The basic methodology as applied to atoms and simple diatomic molecules has been described in [1-11]. Reference [4] provides a detailed review of the work. The essential aspects have been:

- development of the finite element discrete variable method (FEDVR) to spatially discretize the coordinates of the electrons,

- construction of numerical grid capable of efficiently computing the required one ad two electron matrix elements, and

- extensions of the short iterative Lanczos method to propagate the wavefunction in time.

Previous efforts have efficiently parallelized the FEDVR method using MPI, shown that it scales linearly with the size of the FEDVR basis and applied it to selected problems [1-13]. Large scale calculations have been performed on a number of atoms and molecules using resources provided by the NSF Extreme Science and Engineering Discovery Environment (XSEDE).

In more recent research we have begun to employ a mixed basis of Gaussian functions at short range and FEDVR functions at long range to extend our methods to complex polyatomic molecules. This approach has several important advantages over using a single basis over all of space. First, the use of nuclear-centered Gaussians preserves the local atomic symmetry around each nucleus and avoids the poor and often non-convergent behavior of using a single-center FEDVR basis at all distances. Second, once the electron is far enough away from the nuclear cusps, a single center expansion converges quickly and, importantly, can represent the electrons out to very large distances using an approach that is very amenable to domain decomposition. The major issue is to compute the one and two electron integrals between the two types of basis functions. The formalism we have been developing requires as input, transition density matrices extracted from a high-level quantum chemistry code, in order to compute the additional one and two-electron integrals. Quantum chemist Jeppe Olsen from Aarhus University has been collaborating with us on the project. Our research team meets on a weekly basis to discuss the issues and plans. During the past year, the density matrices have been integrated into a new code called ASTRA and some preliminary tests of the code have been performed that indicate things are working as expected. It should be noted that this is a very complex many-body problem and even with the most talented of researchers, it is a long-term effort.

The calculation of the one and two-electron matrix elements over the hybrid basis, must be performed numerically. Given the polycentric nature of the electron distribution and the need to compute these integrals to significantly larger distances than in quantum chemistry calculations, it was necessary to develop an efficient 3D integration scheme. One cannot use methods which fix the coordinate system at a single point in space, as they are at best very slowly convergent and often do not converge at all. To overcome that, a popular approach developed by Axel Becke based on the partition of unity, defines atomic grids, centered on the atoms where the grid points are appropriately weighted to satisfy the partition of unity. The original method of Becke required substantial modification for our purposes. We require a

central grid that can describe the much larger integration region without too much contamination from the points on the nuclear centers. To accomplish that, we define a new partition of unity that constrains the atomic grid points to small atomic spheres. A central grid is then added to take care of the interstitial and longer-range parts of configuration space where the atomic grid points are forced to vanish. The new partition of unity does remarkably well in performing very accurate integration for integrands having nuclear cusps as well as oscillating at larger distances. The work has been completed and a paper [14] describing it has been published.

Lastly, we have been engaged in efforts to generalize the short iterative Lanczos (SIL) method used to integrate the TDSE efficiently and accurately for much longer times [15]. An integral equation formalism has been developed which exploits the fact that part of the Hamiltonian is a linear operator and can be treated using an exponential propagator, which is exactly what the SIL provides. The remainder term, involving a time integral over the residual interaction and the unknown wavefunction, is computed by numerical integration over the large time step using the SIL on each of the terms in the integrand. To do the integral requires that we have a previous approximation to the wavefunction. Thus, the entire scheme is iterative, starting with the wavefunction from the previous time step.

We have been collaborating on this project with a mathematics graduate student, Ryan Schneider, at UCSD. Ryan was initially supported by the Mathematical Sciences Graduate Internship (MSGI) program of the NSF and worked remotely with us for almost three months in 2019. We have continued to work with Ryan since the end of the internship as the work was not finished and Ryan was eager to continue. While Ryan is mainly occupied with his Ph.D. degree, he remains active, working on the problem part time. He has applied to the MSGI program again and we hope to obtain financial support for his research during the summer of 2022.

For completeness, we also reference a review paper [16] by Schneider and Gharibnejad that was published in *Nature Review Physics* which received the ITL Outstanding Journal Paper award in 2020.

[1] J. Feist, S. Nagele, R. Pazourek, E. Persson, B. I. Schneider, L. A. Collins, and J. Burgdörfer. Nonsequential Two-Photon Double Ionization of Helium. *Physical Review A* **77** (2008), 043420.

[2] X. Guan, K. Bartschat, and B. I. Schneider. Dynamics of Two-photon Ionization of Helium in Short Intense XUV Laser Pulses. *Physical Review A* **77** (2008), 043421.

[3] X. Guan, K. Bartschat, and B. I. Schneider. Two-photon Double Ionization of H_2 in Intense Femtosecond Laser Pulses. *Physical Review A* **82** (2010), 041407.

[4] B. I. Schneider, J. Feist, S. Nagele, R. Pazourek, S. Hu, L. Collins, and J. Burgdörfer. Recent Advances in Computational Methods for the Solution of the Time-Dependent Schrödinger Equation for the Interaction of Short, Intense Radiation with One and Two Electron Systems, in Dynamic Imaging. In Quantum Dynamic Imaging, (A. Bandrauk and M. Ivanov eds.), CRM Series in Mathematical Physics, Springer, New York, 2011.

[5] X. Guan, E. Secor, K. Bartschat, and B. I. Schneider. Double-slit Interference Effect in Electron Emission from H_2^+ Exposed to X-Ray Radiation. *Physical Review A* **85** (2012), 043419.

[6] X. Guan, K. Bartschat, B. I. Schneider, L. Koesterke, Resonance Effects in Two-Photon Double Ionization of H_2 by Femtosecond XUV Laser Pulses. *Physical Review A* **88** (2013), 043402.

[7] J. Feist, O. Zatsarinny, S. Nagele, R. Pazourek, J. Burgdörfer, X. Guan, K. Bartschat, and B. I. Schneider. Time Delays for Attosecond Streaking in Photoionization of Neon. *Physical Review A* **89** (2014), 033417.

[8] X. Guan, K. Bartschat, B. I. Schneider, and L. Koesterke. Alignment and Pulse-duration Effects in Two-photon Double Ionization of H_2 by Femtosecond XUV Laser Pulses. *Physical Review A* **90** (2014), 043416.

[9] B. I. Schneider, L. A. Collins, X. Guan, K. Bartschat, and D. Feder. Time-Dependent Computational Methods for Matter Under Extreme Conditions. *Advances in Chemical Physics* **157** (2015), Proceedings of the 240 Conference: Science's Great Challenges, (A. Dinner, ed.), John Wiley.

[10] B. I. Schneider, X. Guan, and K. Bartschat. Time Propagation of Partial Differential Equations Using the Short Iterative Lanczos Method and Finite-Element Discrete Variable Representation. *Advances in Quantum Chemistry* **72** (2016), 95-127.

[11] B. I. Schneider. How Novel Algorithms and Access to High Performance Computing Platforms are Enabling Scientific Progress in Atomic and Molecular Physics. *Journal of Physics: Conference Series* **759** (2016), 012002.

[12] B. I. Schneider. Looking Back at 45 Years of Computational Physics and Chemistry. *Computing in Science and Engineering* **19** (September-October 2017), 4-5.

[13] B. I. Schneider, L. A. Collins, K. Bartschat, X. Guan, and S. X. Hu. A Few Selected Contributions to Electron and Photon Collisions with H_2 and H_2^+. *Journal of Physics* **50** (2017), 214004.

[14] H. Gharibnejad, N. Douguet, B. I. Schneider, J. Olsen, and L. Argenti, A Multi-Center Quadrature Scheme for the Molecular Continuum, Computer Physics Communications, **263** (2021), 107889.

[15] H. Gharibnejad, B. I. Schneider, M. Leadingham, and H. J. Schmale. A Comparison of Numerical Approaches to the Solution of the Time-Dependent Schrödinger Equation in One Dimension. *Computer Physics Communications* 252 (2020), 106808.

[16] B. I. Schneider and H. Gharibnejad. Numerical Methods Every Atomic and Molecular Theorist Should Know. *Nature Reviews Physics* **2** (2020), 89-102.

A Science Gateway for Atomic, Molecular and Optical Science

Barry I. Schneider
Lincoln Carr (Colorado School of Mines)
Klaus Bartschat (Drake University)
Kathryn Hamilton (Drake University)
Oleg Zatsarinny (Drake University)[Deceased]
Igor Bray (Curtin University, Australia)
Armin Scrinzi (Ludwig-Maximilians U., Germany)
Fernando Martin (U. Autonoma de Madrid, Spain)
Jesus Gonzalez Vasquez (U. Autonoma de Madrid)
Jimena Gorfinkiel (Open University, UK)
Robert Lucchese (Lawrence Berkeley National Lab)
Sudhakar Pamidighantam (Indiana University)
Charlotte Froese-Fischer (U. of British Columbia)
Nicholas Douguet (Kennesaw State University)

An international effort has been underway since early 2019 to develop and maintain a Science Gateway for Atomic Molecular and Optical Science (AMOSG) [1,2]. The gateway was renamed from a Science Gateway for Atomic, and Molecular Physics, to a Science Gateway for Atomic Molecular and Optical Science to reflect the broader nature of the ongoing activity The purposes of the gateway are to (1) collect and make available to the community a set of advanced computational tools that are actively being used to study atomic and molecular collisions and the interaction of radiation with atoms and molecules, (2) provide educational materials for beginning and advanced users desiring to learn the ideas and concepts of AMOS, both theoretical and computational and (3) make available to the broad community atomic and molecular data needed for many applications. The availability of collision data is critical to many areas of physics including astrophysics, fusion energy, the study of lighting and microelectronics.

Codes for modeling and simulation of such phenomena have been developed in specific groups by graduate students and postdocs but are often poorly documented, and unavailable outside the group developing the software. This leads to "reinventing the wheel" in too many instances. Maintaining these computational tools, as well as enhancing their capabilities, is one of the major goals of the project and is critical to ensure continued scientific progress in AMOS.

Another important goal is to enable the code developers themselves to compare the calculations of specific well-defined problems using different methodologies. This enables the verification of results of different codes and encourages comparison with experiment, where available. It has already been demonstrated that a few of these codes are often more accurate than experiment and thus provide a predictive capability when experimental results are unavailable.

At the outset, the group acknowledged that, in contrast to some other communities, the AMOS community has lagged behind in developing community supported software packages that are robust and used by others. The group was convinced the time had arrived to change existing practices and make these tools available and easily used by future generations of AMP scientists as well as the developers themselves.

The group wrote a proposal to the NSF XSEDE program to fund some initial development of the gateway. The proposal was successful and, importantly, provided the developers with some hands-on assistance from the Extended Collaborative Support Services arm of XSEDE. This was vital to the success of the effort. In particular, we acknowledge the important contribution of Sudhakar Pamidighantam of IU in making our efforts a success.

There have been some important advances during the current reporting period: (1) the original web pages for the gateway have been completely revamped and new material has been added, (2) the original six major codes chosen for initial deployment have been expanded to include eleven codes and many have already been ported to at least three XSEDE supercomputers, (3) there is an early version of an API available to enable users to perform calculations with at least one of the codes, tRecX, and (4) NIST has awarded a contract to Indiana University to further develop the needed GUI interfaces.

The need for longer term support for the project has been recognized as an essential ingredient to its ultimate success. To that end, the group spent considerable effort during the fall of 2020 in writing an NSF proposal to the Cyberinfrastructure for Sustained Innovation program. That proposal was not successful, but it has been rewritten and resubmitted for the current competition. Professor Lincoln Carr, of the Colorado School of Mines, remains the project PI. Co-PI's are Sudhakar Pamidighantam from IU and Klaus Bartschat from Drake University. In addition, Nicholas Douguet, from Kennesaw State, and Robert Lucchese from LBNL have joined Schneider as Senior Investigators. The proposal was submitted from the Colorado School of Mines as the lead organization. The involvement of the Colorado School of Mines in the project has enabled us to get needed resources to develop the portal pages into what has become a very professional state. The current portal contains good descriptions of the codes, the people involved, links to documentation, a bibliography, a preliminary data repository and some nice graphical material illustrating a few of the calculations that have been done with the codes.

[1] B. I. Schneider, K. Bartschat, O. Zatarinny I. Bray, A. Scrinzi, F. Martin, M. Klinker, J. Tennyson, J. Gorfinkiel, and S. Pamidighanta. A Science Gateway for Atomic and Molecular Physics, 2020. URL: arXiv:2001.02286

[2] B. I. Schneider, K. Bartschat, O. Zatsarinny, K. Hamilton, I. Bray, A. Scrinzi, F. Martin, J. G. Vasquez, J.

Tennyson, J. Gorfinkiel, R. Lucchese, and S. Pamidighantam. Atomic and Molecular Scattering Applications in an Apache Airavata Science Gateway. In *PEARC '20: Practice and Experience in Advanced Research Computing*, July 2020, DOI: 10.1145/3311790.3397342

[3] B. I. Schneider. "Atomic and Molecular Scattering Applications in an Apache Airavata Science Gateway." PEARC '20: Practice and Experience in Advanced Research Computing, Online, July 2020.

Computing Ill-Posed Time-Reversed Dissipative Evolution Equations

Alfred S. Carasso

Ill-posed deconvolution problems and related time-reversed dissipative evolution equations, pervade measurement science and are important in several other technological applications. In numerous scientific measurements, the instrument point spread function is a bell-shaped distribution that may be well-approximated by a Gaussian, or by a heavy-tailed indefinitely divisible probability density, often with parameters that are only tentatively known. Reformulating the integral equation deconvolution problem into an equivalent time-reversed generalized diffusion equation, provides significant advantages. Marching backward in time stepwise, from a positive time T to time $t = 0$, allows the deconvolution to unfold in slow motion, provides the ability to monitor that process, and the possibility of terminating it prior to time $t = 0$, to prevent serious noise contamination and/or development of ringing artifacts. Such an approach, involving backward fractional and logarithmic diffusion equations, has been successfully applied in image deblurring of MRI and PET brain scans, nanoscale scanning electron micrographs, and galactic scale Hubble Space Telescope imagery [1-3].

In other contexts, much success has been achieved in environmental forensics, for example, by solving advection diffusion equations backward in time to locate sources of groundwater contamination [4]. In numerical weather prediction and other geophysical computations, data assimilation involving time-reversed dissipative equations plays a major role in initializing prediction models [5-8].

While iterative methods are often used to solve ill-posed evolution equations, such methods are time-consuming and often develop stagnation points. Recently, a powerful non-iterative direct approach has been developed for time-reversed multidimensional nonlinear dissipative equations, based on stabilizing explicit backward marching finite difference schemes. An appropriate easily synthesized compensating smoothing operator is applied at every time step to quench the instability. The stabilized scheme now becomes unconditionally stable, but slightly inconsistent, and

eventually leads to a distortion away from the true solution. This is the *stabilization penalty*. However, in many problems of interest, the cumulative error is sufficiently small to allow for useful results. In a series of papers [9-16], such stabilized schemes have been successfully applied to interesting classes of time-reversed nonlinear initial value problems for parabolic equations, viscous wave equations, coupled sound and heat ow, thermoelastic vibrations, 2D viscous Burgers equations, and 2D incompressible Navier-Stokes equations. Such computations had not previously been considered possible.

Data Assimilation in 2D Viscous Burgers Equation. The 2D viscous Burgers equation is a coupled system of two nonlinear equations in two unknowns, $u(x, y, t)$, $v(x, y, t)$. In [13], the ill-posed problem of backward recovery from inexact data at an appropriately chosen $T > 0$, is studied. Rigorous uncertainty estimates in [17, 18] require a sufficiently small $T > 0$, as well as sufficiently small solution derivatives on $0 \leq t \leq T$, for useful recovery.

A fundamentally different time-reversed problem is discussed in [16]. In a bounded domain $\Omega \subset R^2$, with homogeneous boundary conditions on $\partial\Omega$, and no forcing term, the following *data assimilation/inverse design* problem associated with that system is considered:

Find initial values $[u(.,0), v(.,0)]$ that can evolve into a close approximation to a desired target result $[u^*(.,T), v^*(.,T)]$ at some suitable $T > 0$.

Here, in contrast to [16], highly non smooth target data are considered that *may not correspond to actual solutions* at time T, and it *may not be possible* to find such initial values. Such a 2D viscous Burgers problem has not previously been studied. For the 1D Burgers equation, data assimilation is discussed in [5-8], using various iterative methods.

Using a direct non-iterative method based on marching backward in time with a stabilized explicit finite-difference scheme, a large class of examples is presented in [16] where, with realistic values of $T > 0$, and Reynolds numbers as high as 18000, useful initial values can be found that evolve into good approximations to the desired target data, with modestly small L^1 relative errors. Importantly, there are also unsuccessful examples.

Numerical experiments presented in [16] involve 8-bit, 256×256 pixel gray-scale images, defined by non-differentiable intensity data that would be quite difficult to synthesize mathematically. Such images pose significant challenges in ill-posed reconstruction and constitute an invaluable exploratory tool in a wide variety of dissipative evolution equations.

In Figure 29, the leftmost column contains the desired non smooth target data at 2.5×10^{-4}, representing images of the Sydney Opera House and Joan Crawford,

DESIRED RESULT AT TIME T   **COMPUTED AT TIME 0**   **ACHIEVED AT TIME T**
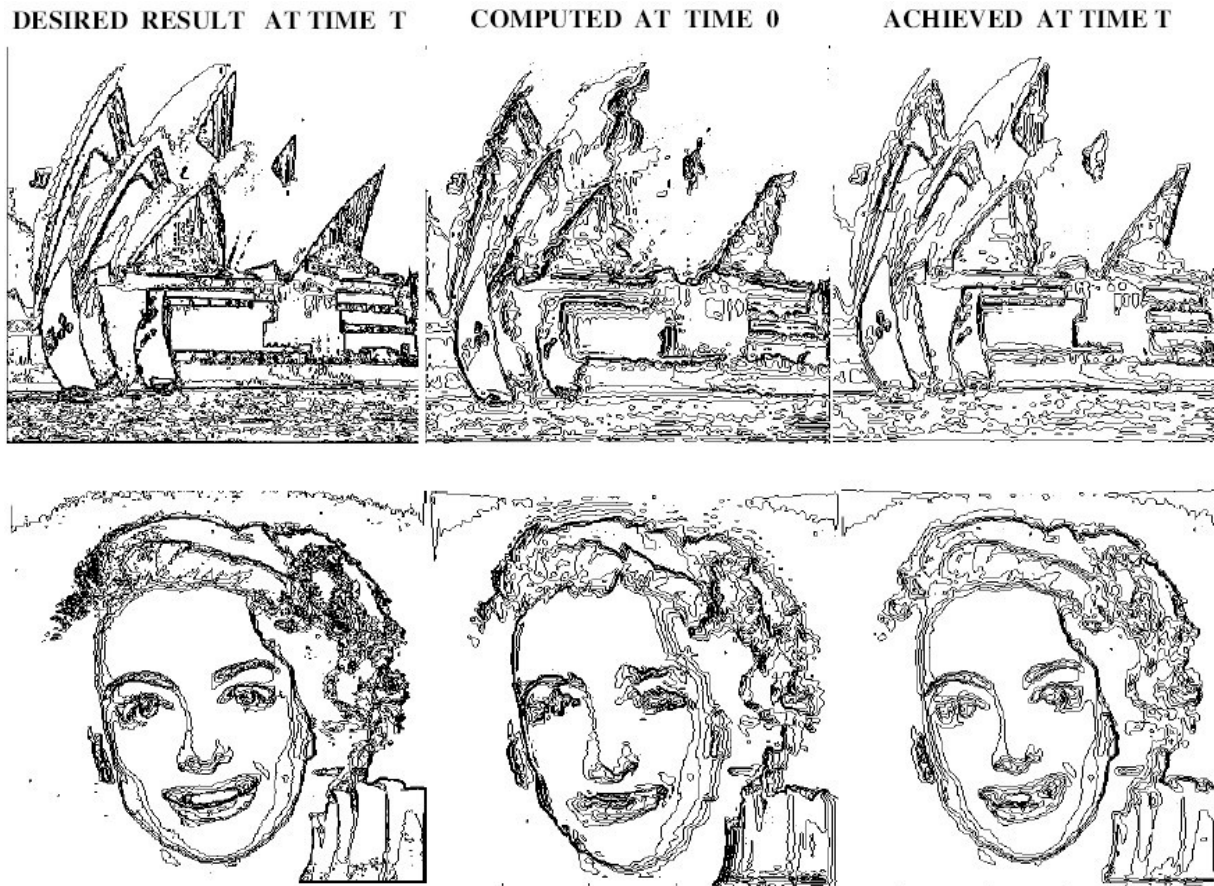
Figure 29. *Successful data assimilation in 2D viscous Burgers equation from target time* 2.5×10^{-4}*, with Reynold's number 17800 using non-differentiable Sydney Opera House and Joan Crawford images. Data in rightmost column approximates desired target data in leftmost column with an* L^1 *relative error less than 9 % for each image.*

with a Reynolds Number of 18000. Marching backward in time with these data leads to candidate initial values at time $t = 0$, shown in the middle column. When marched forward in time, these data now produce a solution to the 2D Burgers equation shown in the rightmost column. However, since the leftmost column was not a solution, the rightmost column cannot be expected to match. Nevertheless, successful data assimilation is achieved, in that the 256×256 intensity data in each of the rightmost images, approximates the corresponding target intensity data in the leftmost images, with an L^1 relative error less than 9 %.

An example of unsuccessful data assimilation with the same T value and Reynolds Number, is shown in Figure 30. There, the desired target alphanumeric image at the top of the leftmost column, is very poorly approximated by the assimilated image at the top of the rightmost column, with an L^1 relative error of 174 %.

[1] A. S. Carasso. The APEX Method in Image Sharpening and the Use of Low Exponent Lévy Stable Laws. *SIAM Journal on Applied Mathematics* **63** (2002), 593-618.

[2] A. S. Carasso, D. S. Bright, and A. E. Vladar. APEX Method and Real-Time Blind Deconvolution of Scanning Electron Microscopy Imagery. *Optical Engineering* **41** (2002), 2499-2514.

[3] A. S. Carasso. Bochner Subordination, Logarithmic Diffusion Equations, and Blind Deconvolution of Hubble Space Telescope Imagery and Other Scientific Data. SIAM *Journal on Imaging Sciences* **3** (2010), 954-980.

[4] J. Atmadja and A. C. Bagtzoglou. State of the Art Report on Mathematical Methods for Groundwater Pollution Source Identification. *Environmental Forensics* **2** (2001), 205-214.

[5] J. Lundvall, V. Kozlov, and P. Weinerfelt. Iterative Methods for Data Assimilation for Burgers' Equation. *Journal of Inverse and Ill-Posed Problems* **14** (2006), 505-535.

[6] D. Auroux and J. Blum. A Nudging-Based Data Assimilation Method for Oceanographic Problems: The Back and Forth Nudging (BFN) Algorithm. *Nonlinear Processes in Geophysics* **15** (2008), 305-319.

[7] D. Auroux, P. Bansart, and J. Blum. An Evolution of the Back and Forth Nudging for Geophysical Data Assimilation: Application to Burgers Equation and Comparison. *Inverse Problems in Science and Engineering* **21** (2013), 399-419.
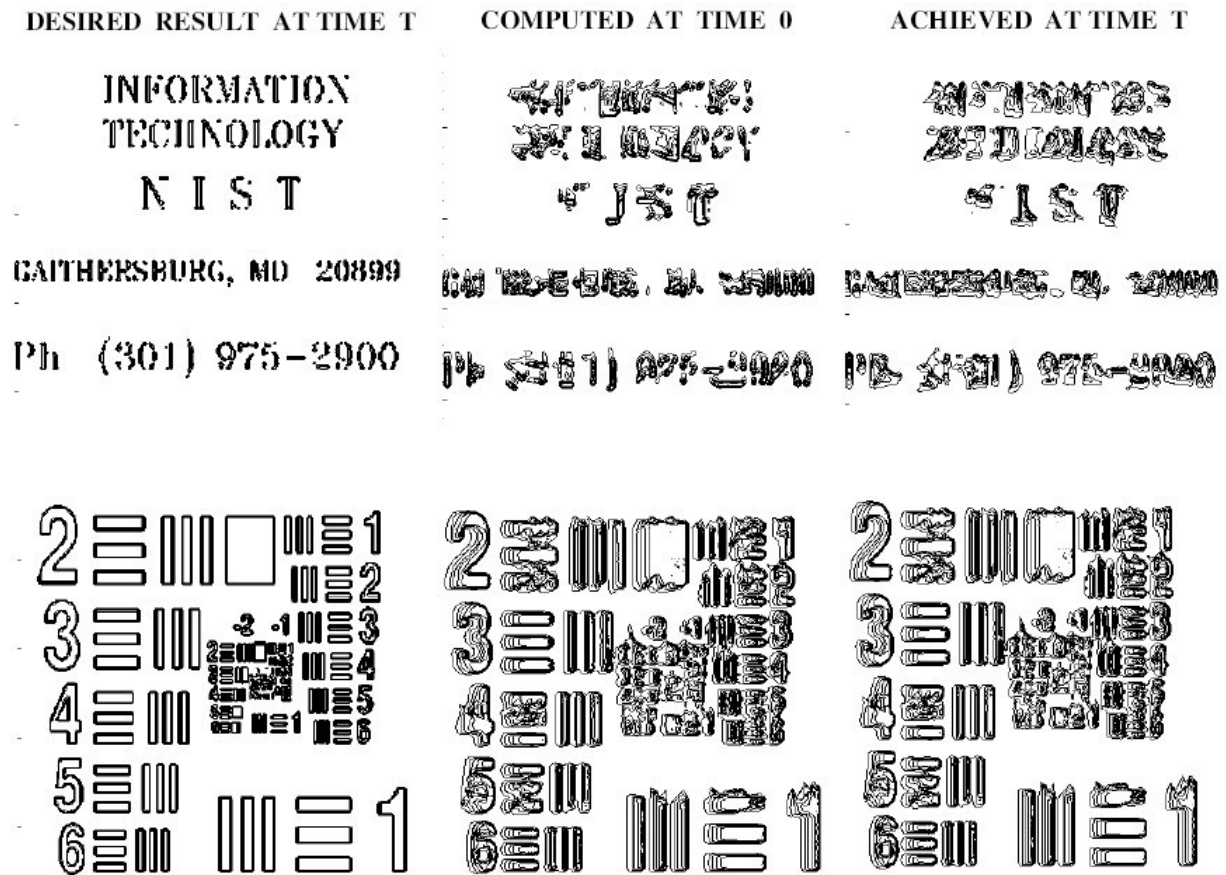
Figure 30. *Failure in data assimilation in 2D Burgers equation from target time* 2.5×10^{-4}*, with Reynold's number 18000 using Alphanumeric and USAF Chart images. Data in rightmost column approximates desired target data in leftmost column with L^1 relative errors of 174 % for the top image and 54 % for the bottom image.*

[8] H. F. de Campos Velho, V. C. Barbosa, and S. Cocke. Special Issue on Inverse Problems in Geosciences. *Inverse Problems in Science and Engineering* **21** (2013), 355-356.

[9] A. S. Carasso. Stable Explicit Time-Marching in Well-Posed or Ill-Posed Nonlinear Parabolic Equations. *Inverse Problems in Science and Engineering* **24** (2016), 1364-1384.

[10] A.S. Carasso. Stable Explicit Marching Scheme in Ill-Posed Time-Reversed Viscous Wave Equations. *Inverse Problems in Science and Engineering* **24** (2016), 1454-1474.

[11] A.S. Carasso. Stabilized Richardson Leapfrog Scheme in Explicit Step-wise Computation of Forward or Backward Nonlinear Parabolic Equations. *Inverse Problems in Science and Engineering* **25** (2017), 1719-1742.

[12] A.S. Carasso. Stabilized Backward in Time Explicit Marching Schemes in the Numerical Computation of Ill-Posed Time-Reversed Hyperbolic/Parabolic Systems. *Inverse Problems in Science and Engineering* **27** (2019), 134-165.

[13] A.S. Carasso. Stable Explicit Stepwise Marching Scheme in Ill-Posed Time-Reversed 2D Burgers Equation. *Inverse Problems in Science and Engineering* **27** (2019), 1672-1688.

[14] A.S. Carasso. Computing Ill-Posed Time-Reversed 2D Navier-Stokes Equations, using a Stabilized Explicit Finite Difference Scheme Marching Backward in Time. *Inverse Problems in Science and Engineering* **28**:7 (2020), 988-1010. DOI: 10.1080/17415977.2019.1698564

[15] A.S. Carasso. Stabilized Leapfrog Scheme Run Backward in Time, and the Explicit $O(\Delta t^2)$ Stepwise Computation of Ill-Posed Time-Reversed 2D Navier-Stokes Equations. *Inverse Problems in Science and Engineering* **29**:13 (2021), 3062-3085. DOI: 10.1080/17415977.2021.1972997

[16] A.S. Carasso. Data Assimilation in 2D Viscous Burgers Equation Using a Stabilized Explicit Finite Difference Scheme Run Backward in Time. *Inverse Problems in Science and Engineering* **29**:13 (2021), 3475-3489. DOI: 10.1080/17415977.2021.2009476

[17] R. J. Knops and L. E. Payne. On the Stability of Solutions of the Navier-Stokes Equations Backward in Time. *Archives of Rational Mechanics and Analysis* **29** (1968), 331-335.

[18] D. N. Hao, V. D. Nguyen, and V. T. Nguyen. Stability Estimates for Burgers-Type Equations Backward in Time. *Journal of Inverse and Ill Posed Problems* **23** (2015), 41-49.

Computational Tools for Image and Shape Analysis

Günay Doğan
Javier Bernal
James Lawrence
Charles R. Hagwood (NIST ITL)
Prashant Athavale (Clarkson University)
Emmanuel Atindama (Clarkson University)
Peter Lef (Clarkson University)
Harbir Antil (George Mason University)
Soeren Bartels (University of Freiburg)
Olakunle Abawonse (SUNY Binghamton)

The main goal of this project is to develop efficient and reliable computational tools to detect geometric structures, such as curves, regions, and boundaries, from given direct and indirect measurements, e.g., microscope images or tomographic measurements, as well as to evaluate and compare these geometric structures or shapes in a quantitative manner. This is important in many areas of science and engineering, where the practitioners obtain their data as images, and would like to detect and analyze the objects in the data. Examples are microscopy images for cell biology or micro-CT (computed tomography) images of microstructures in material science, shoeprint images in crime scenes for footwear forensics. In FY 2021, advances were made in the two fronts of this project: image segmentation and shape analysis. Python implementations of solutions for problems in these areas were released, and documentation and examples were provided. In the following, we provide more details on the specific work carried out.

Image Segmentation. This is the problem of finding distinct regions and their boundaries in given images. It is a necessary data analysis step for many problems in cell biology, forensics, and material science, as well as other fields in science and engineering. In FY 2021, G. Doğan and his collaborators continued to work on multiple strategies for image segmentation.

The first of these strategies builds on implicit representations of image regions by embedding them in phase field functions, developed previously by Doğan and his collaborators, H. Antil and S. Bartels. Initial guesses of region candidates are placed on the image, and the corresponding phase field function is evolved in a manner optimizing a data fidelity and boundary regularity term. This algorithm tends to produce well-defined segmentations without stray pixels. Doğan introduced further improvements to this algorithm in FY 2021 and documented its performance with detailed experiments. The algorithm and the experiments are described in [1].

Doğan also continued to work with former graduate intern O. Abawonse on their implementation an image segmentation algorithm that leverages a convexified version of a piecewise constant segmentation energy.

This algorithm is useful for obtaining two-phase segmentations, i.e., background-foreground segmentation of grayscale images. They are currently writing a paper describing this work.

In order to segment more complicated images, such as natural images and others with complex textures, Doğan pursued a different approach, combining texture features and machine learning algorithms. This works in three steps. First a given image is converted into a feature image, using a set of pre-defined texture features. Then a subset of the image features is selected and assigned class labels to represent the distinct regions to be identified. The labels can be assigned by a user or a clustering algorithm. Finally, the labeled subset is used to train a classification algorithm, such as support vector machines (SVM) or random forests, which, in turn, is used to label the remaining image features and the corresponding pixels. The output of the classification algorithm is the desired segmentation. An example can be seen in Figure 31.

Doğan continued his collaboration with P. Athavale, E. Atindama, and P. Lef (who also was a SURF student in ACMD) on preprocessing and segmentation of orientation images of microstructures. Orientation images, more specifically electron backscatter diffraction (EBSD) images, often come with many misorientation pixels, which have the appearance of noise, and may have regions of missing data. To alleviate these issues, they have been developing PDE-based denoising and inpainting algorithms to produce high quality reconstructions of the EBSD images. They further improved their algorithms in FY 2021 in terms of efficiency and quality of reconstructions. A critical novelty was a new preprocessing algorithm that they developed to remove certain artifacts specific to orientation images. A Python implementation of their algorithm is completed. They are currently working on a paper describing these algorithms.

Shape Analysis. A natural approach to analyze and compare objects in images and data is through their shapes, an aspect that is invariant to rotation, translation, and scaling. Shape-based analysis can facilitate object recognition and can provide a more intrinsic way to perform statistics of geometric data. Previously, Doğan, Bernal, and Hagwood developed an efficient optimization algorithm to compute the elastic shape distances between 2d closed curves [2, 3, 4], building on the square-root velocity framework of Srivastava et al. [6, 7]. This shape distance is the fundamental building block for shape analysis.

In FY 2020 and FY 2021, Bernal and collaborators [5] extended the work of Srivastava et al. [6, 7] for the elastic registration of two simple curves, originally in two dimensions, to higher dimensions, and similarly the elastic shape distance to higher dimensions as well. Assuming the first curve that we consider has multiple possible starting points, and the second curve has a fixed
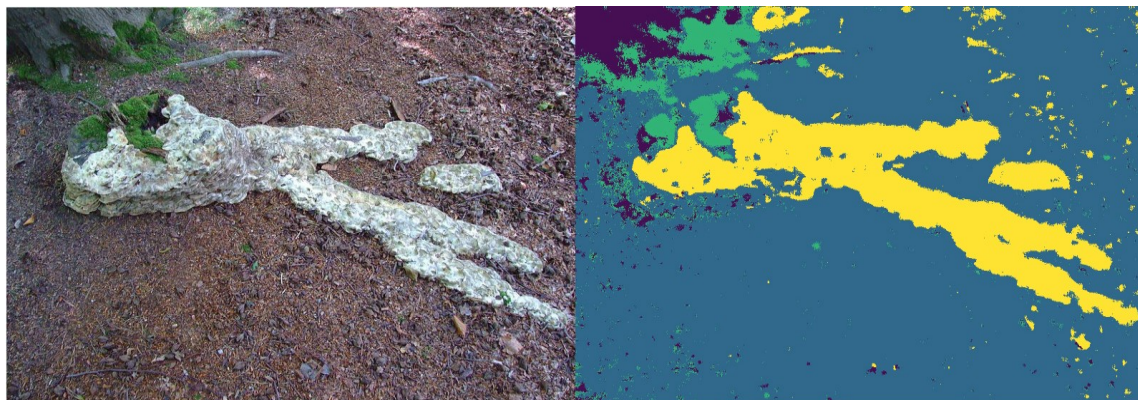
Figure 31. *Segmentation (right) of a natural image (left) with complex texture using a machine learning algorithm applied a derived image of texture features.*

starting point, Srivastava et al. compute the shape distance one starting point (of the first curve) at a time. They minimize an L_2 type distance between the curves using alternating computations of optimal diffeomorphisms of the unit interval and optimal rotation matrices that reparametrize and rotate one of the curves [6, 7]. Building on this previous work, Bernal et al. presented a dynamic programming algorithm to compute optimal diffeomorphisms for curves in any dimensions in linear time complexity [5]. Moreover, they justified a key algorithm for computing optimal rotations in a purely algebraic manner, namely, the Kabsch-Umeyama algorithm for computing optimal rotation matrices, which uses the singular value decomposition of the rotation matrix. In addition, Bernal et al. presented a procedure in [5] to minimize the L_2 type distance that alternates computations of optimal diffeomorphisms with successive computations of optimal rotation matrices for all starting points of the first curve. Carrying out computations this way is not only more efficient on its own, but also if both curves are closed, it allows us to use the Fast Fourier Transform for further efficiency to compute optimal rotation matrices.

Currently, work is underway to extend the shape analysis framework [7, 8] to surfaces, namely, elastic registration of two simple surfaces in three-dimensional (3D) space and then the elastic shape distance between them. Previous work on this problem was carried out by Kurtek et al. [10]. With a simple surface defined as a parametrized surface in 3D space, that is, the range of a one-to-one function from an elementary region in the plane into 3D space, Bernal defined a shape function of a parametrized surface in 3D space, different from the one defined by Kurtek et al., and established some fundamental results about this function. A similar definition and similar results were presented in [5, 6, 7, 8, 9] in the context of the shape function of a parametrized curve in $d-$dimensional space. Based on this definition of the shape function of a parametrized surface, Bernal also defined the elastic shape distance between two

parametrized surfaces. This is done at first by reparametrizing one of the surfaces through the minimization of a double integral in terms of homeomorphisms from the elementary region that is part of the definition of the parametrized surfaces, onto itself (the homeomorphisms having Jacobians of positive determinant), and later by rotating one of the surfaces by minimizing the same double integral in terms of rotations of the surface. He demonstrated that this distance is well defined using arguments similar to those used to justify the definition of the distance between curves in $d-$dimensional space in [7, 8]. Next steps will be focused on developing an algorithm or heuristic to minimize the aforementioned double integral to compute optimal homeomorphisms.

Using different shape representations or different versions of the algorithms lead to different shape dissimilarity metrics, and this brings the question of which metric would perform best. In the previous years, Doğan and Fleisig (UC Berkeley) developed a Python program, VEMOS (Visual Explorer for Metrics Of Similarity) that can be used to evaluate and compare multiple competing similarity/dissimilarity metrics, including shape dissimilarity metrics. VEMOS is useful for applications beyond shape distances; it can be used in a versatile manner to evaluate multiple alternative dissimilarity metrics for heterogeneous data sets, including images, shapes, point clouds and other data types. In FY 2021, Doğan continued to develop VEMOS, made bug fixes, and added new features. A report describing capabilities of VEMOS was published as a NIST technical note [11].

[1] H. Antil, S. Bartels, and G. Doğan. A Region-Based Segmentation Model with Fractional Diffusion for Improved Boundary Regularization. In preparation.

[2] G. Doğan, J. Bernal, and C.R. Hagwood. Fast Algorithms for Shape Analysis of Planar Objects. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (CVPR'15), Boston, MA, June 2015.

[3] G. Doğan, J. Bernal, and C.R. Hagwood. FFT-based Alignment of 2d Closed Curves with Application to Elastic Shape Analysis. In *Proceedings of the 1st International Workshop on Differential Geometry in*

Computer Vision for Analysis of Shapes, Images and Trajectories (DiffCV'15), Swansea, UK, September 2015.

[4] J. Bernal, G. Doğan, and C.R. Hagwood. Fast Dynamic Programming for Elastic Registration of Curves. In *Proceedings of the 2nd International Workshop on Differential Geometry in Computer Vision and Machine Learning* (DiffCVML'16), Las Vegas, NV, July 1, 2016.

[5] J. Bernal, J. Lawrence, G. Doğan, and C.R. Hagwood. On Computing Elastic Shape Distances Between Curves in d-dimensional Space. NIST Technical Note 2164, June 2021, 39 pages. DOI: 10.6028/NIST.TN.2164

[6] A. Srivastava, E. Klassen, S. Joshi, and I. Jermyn. Shape Analysis of Elastic curves in Euclidean Space. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**:7 (2011), 1415-1428.

[7] A. Srivastava and E.P. Klassen. Functional and Shape Data Analysis. Springer, 2016.

[8] J. Bernal. Shape Analysis, Lebesgue Integration and Absolute Continuity Connections. NISTIR 8217, 2018.

[9] S. H. Joshi, E. Klassen, A. Srivastava, and I. H. Jermyn. A Novel Representation for Riemannian Analysis of Elastic Curves in Rⁿ. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (CVPR'07), Minneapolis, MN, June 2007.

[10] S. Kurtek, E. Klassen, Z. Ding, and A. Srivastava. A Novel Riemannian Framework for Shape Analysis of 3D Objects. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (CVPR'10), San Francisco, CA, June 2010.

[11] E. Fleisig and G. Doğan. VEMOS: GUI for Evaluation of Similarity Metrics of Complex Data Sets. NIST Technical Note 2160, June 2021, 34 pages. DOI: 10.6028/NIST.TN.2160

Seminorm Regularization of Linear Inverse Problems

Matthew J. Roberts
Mark S. Gockenbach (University of Delaware)

Linear inverse problems arise in many applications pertaining to measurement. In general, a linear inverse problem is posed as follows: Determine solution x in the equation

$$T\,x\ =\ y,$$

where y is a measurement of y^*. Here, x, y, and y^* are usually functions in a function space such as $L^2(0, 1)$, and this is the setting we will assume. The basic question is the following: if y is a good estimate of y^*, then is x a good estimate of the true solution x^*? Here,

$$T\,x^*=y^*.$$

The inverse problems of interest arise when the solution x does not continuously depend on the data y. This means that an accurate measurement y of y^* can still result in a large distance between x and x^*. An inverse

problem of this type occurs in the case that T is an integral operator over a closed and bounded domain, and more generally, whenever T is a compact operator (see [2] and [5]). One way to deal with this is by using seminorm regularization. We consider the following problem: Determine solution $\tilde{x}$ to the problem

$$\min_{u\,\in\,L^2(0,1)}\|Tu - y\|_{L^2(0,1)^2} + \lambda\|Lu\|\ \ \|_{L^2(0,1)^2}.\quad (1)$$

Here, L is an operator (such as the derivative operator), and λ is a scalar greater than 0. The solution $\tilde{x}$ is the seminorm regularized solution. The idea is to choose an operator L that reflects an undesirable feature in the solution x. For example, if the solution to (1) is highly oscillatory, a good choice for L would be the derivative operator.

The generalized singular value expansion (GSVE) can be used to simultaneously diagonalize the operator pair (T, L) in problem (1) (see [3], sections 2 through 4). This makes determining the solution $\tilde{x}$ of (1) transparent. In practice the operator pair (T, L) is discretized by the pair of finite dimensional operators (T_h, L_h), where, as $h \to 0$, the operator pair (T_h, L_h) goes to (T, L) in some sense. It was shown in [4] that the GSVE of (T_h, L_h) converges to the GSVE of (T, L) as $h \to 0$ if $T_h \to T$ and $L_h \to L$ in the operator norm as $h \to 0$.

Building on this work, together with results from [1], we have been able to derive the rates of convergence of the approximate generalized singular vectors of (T_h, L_h) to the true generalized singular vectors of (T, L). Numerical examples show that these derived rates of convergence are the best possible in the case of variational approximation. Such results will be most useful in error analysis pertaining to seminorm regularization. A manuscript with these results is in review.

In addition, we have been able to derive the rate of convergence of the generalized singular values of (T_h, L_h) to the true generalized singular values of (T, L). We have also been able to prove that there is an improved rate of convergence of the generalized singular values in the case of variational approximation. These results have been documented into a second manuscript that we hope to submit for publication in the near future.

[1] D. K. Crane, M. S. Gockenbach, and M. J. Roberts Approximating the Singular Value Expansion of a Compact Operator. In review.

[2] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems.* Kluwer Academic Publishers Group, Dordrecht, 1996.

[3] M. S. Gockenbach. Generalizing the GSVD. *SIAM Journal on Numerical Analysis* **54** (2016), 2417- 2540.

[4] M. S. Gockenbach and M. J. Roberts. Approximating the Generalized Singular Value Expansion. *SIAM Journal on Numerical Analysis* **56**:5 (2018), 2776-2795.

[5] M. S. Gockenbach. *Linear Inverse Problems and Tikhonov Regularization,* MAA Press, 2016.

Mathematics of Biotechnology

As proof-of-concept academic work in engineering biology meets the market realities of bringing lab science to product initiation, there are questions in how to compare biological products, measure whether desired outcomes are realized, and optimize biological systems for desired behaviors. NIST is working to deliver tools and standards to measure such biological technologies, outputs, and processes from healthcare to manufacturing and beyond. We support this effort with the development and deployment of innovative mathematical modeling and data analysis techniques and tools.

Advanced Data Analysis for Diagnostics, Biometrology, and COVID-19

Paul Patrone
Anthony Kearsley
Prajakta Bedekar
Rayanne Luke
Christopher Heaney (*Johns Hopkins University*)
Nora Pisanic (*Johns Hopkins University*)
Yukari Manabe (*Johns Hopkins University*)
David Thomas (*Johns Hopkins University*)

Figure 32. *Illustration of the classification method developed in [2]. Red x and blue o correspond to samples that are known to be positive and negative for SARS-CoV-2. The measurements are associated with two different methods of quantifying antibodies; see [2] for details. The yellow-green domain is the positive classification domain, whereas dark blue is the negative classification domain. The light blue domain is the holdout domain. The original assay defined the upper two quadrants of the plot as the positive domain, the lower right as the negative domain, and the lower left as the holdout domain. The optimal method maintains testing accuracy while reducing the fraction of holdouts by more than 10 %.*

The on-going COVID-19 health crisis has highlighted the critical need for advanced metrology tools that can be used in both clinical and research settings. For example, accurate diagnostic testing is needed to characterize epidemiological patterns, manage healthcare infrastructure, and identify optimal policy decisions. Motivated by such issues, several staff members in ACMD have been collaborating with a range of internal and external stakeholders on projects related to antibody measurements.

In FY 2021, several of us derived a series of new results showing that diagnostic classification associated with antibody assays can be recast as an optimization problem. From a data analysis standpoint, the underlying measurements return a value r associated with an antibody level, which used to classify a sample as positive or negative, say for having been infected with SARS-CoV-2. A canonical method assigns r to the positive or negative classes depending on whether the data falls into one of two domains D_p or D_n that partition the measurement space. However, we recently demonstrated that the optimal (i.e., minimum error) domains are given by

$$D_p^* = \{r : qP(r) > (1-q)N(r)\}$$
$$D_n^* = \{r : qN(r) > (1-q)P(r)\}$$

where $P(r)$ and $N(r)$ are models quantifying the conditional probability that a positive or negative sample yields a value r, and q is the disease prevalence [1]. We also proved that unbiased estimates of prevalence can be constructed from a counting exercise independent of

sample classification but still dependent on these probability models [1]. Thus, we reduced all diagnostics to a problem of correctly modeling probability distributions.

Following up on this, we began a series of collaborations with the School of Public Health at Johns Hopkins University to improve data analysis associated with saliva-based SARS-CoV-2 antibody tests. A fundamental problem with such assays is the variable nature of the measurement conditions, given that daily activities such as drinking water may dilute samples. As a result, it is sometimes necessary to hold out samples if they are too diluted and/or do not appear to contain a meaningful signal. However, doing so also wastes time and resources if a large amount of the data must be discarded. To address this problem, we formulated a constrained optimization problem that minimizes the size of the holdout domain while maintaining a given classification accuracy for the remaining samples [2]. Interestingly, the solution to this problem is given in terms of the bathtub principle. In the context of diagnostic testing, this principle states that we hold out points from the measurement space with the lowest *local classification accuracy* up to a threshold that depends on the

target classification accuracy. Figure 32 shows the results of this analysis applied to a SARS-CoV-2 assay developed at Johns Hopkins University.

[1] P. N. Patrone, and A. J. Kearsley. Classification Under Uncertainty: Data Analysis for Diagnostic Antibody Testing. *Mathematical Medicine and Biology: A Journal of the IMA* **38**:3 (2021), 396–416. DOI: 10.1093/imammb/dqab007

[2] P. N. Patrone, P. Bedekar, N. Pisanic, Y. C. Manabe, D. L. Thomas, C. D. Heaney, and A. J. Kearsley. Optimal Decision Theory for Diagnostic Testing: Minimizing Indeterminate Classes with Applications to Saliva-Based SARS-CoV-2 Antibody Assays. In review.

Data Analysis for Quantitative Polymerase Chain Reaction Measurements

Paul Patrone
Anthony Kearsley
Robert DeJaco
Peter Vallone (NIST MML)
Erica Romsos (NIST MML)
Patrick Hutchins (US Geological Survey)
Adam Sepulveda (US Geological Survey)

Quantitative polymerase chain-reaction (qPCR) measurements are a mainstay tool for detection of genetic material, with broad applications to diagnostics, environmental testing, and forensics. Indeed, the COVID-19 pandemic has highlighted the critical role of qPCR in identifying active SARS-CoV-2 infections. Despite its importance, however, there are many open questions regarding uncertainty quantification (UQ) and methods for robust analysis of raw data.

In FY 2020, several of us developed new methods for background subtraction and analysis of qPCR data [1]. The main idea behind our approach was to leverage a previously unknown property that all qPCR amplification curves, including their plateau phases, are identical up to horizontal shifts and multiplicative factors. We showed that it was possible to identify positive samples by transforming a test curve onto a positive control via a constrained optimization formulation; see Figure 33. This led to a manuscript and patent application for the resulting methods.

In FY 2021, we engaged in collaborative research and technology transfer with external stakeholders. In particular, the US Geological Survey (USGS) uses qPCR measurements to locate invasive species in freshwater systems by testing for environmental DNA (eDNA). However, given the likelihood of "genetic contamination" in such open systems, UQ of the resulting analysis is critical for justifying remediation decisions.
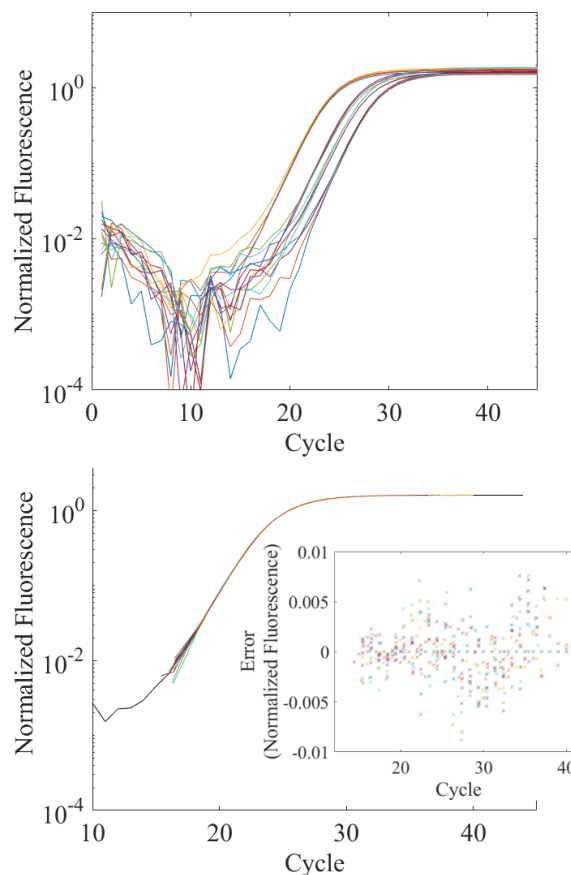
Figure 33. *qPCR amplification curves of the N1 fragment of a SARS-CoV-2 RNA construct. Top: qPCR curves after background subtraction. Bottom: curves after data collapse. The inset shows the error on an absolute scale relative to the master curve. The data collapse is accurate to better than 1 % of the maximum scale of the data.*

This is especially important given that the resulting actions are expensive and have the potential for ecological disruption. To address such questions, we have worked with USGS to identify best practices for data analysis, including determining if certain data analysis routines are statistically consistent with the amount of DNA expected in samples that were tested as part of an interlaboratory study [2].

In addition to this, work continues in doing technology transfer of our qPCR data analysis methods via a CRADA with a private diagnostics lab. On-going research aims to better understand and quantify sources of background in raw data.

[1] P. N. Patrone, E. L. Romsos, M. H. Cleveland, P. M. Vallone, and A. J. Kearsley. Affine Analysis for Quantitative PCR Measurements. *Analytical and Bioanalytical Chemistry* **412** (2020), 7977–7988.

[2] A. J. Sepulveda, P. R. Hutchins, C. Jackson, C. Ostberg, M. B. Laramie, J. Amberg, T. Counihan, A. Hoegh, and D. S. Pilliod. A Round-Robin Evaluation of the Repeatability and Reproducibility of Environmental DNA Assays for Dreissenid Mussels. *Environmental DNA* **2**:4 (2020), 446-459.

Absolute Quantification of Antibodies for SARS-CoV-2

Paul Patrone
Danielle Middlebrooks
Anthony Kearsley
Lili Wang (NIST MML)
Sheng Lin-Gibson (NIST MML)

A key problem associated with antibody measurements for SARS-CoV-2 is the lack of an absolute scale. Typical standards rely on pooled blood samples. Being human derived, however, they have unknown antibody titers, so that subsequent measurements can only be defined on a relative scale. This makes it challenging to understand the degree of protection associated with a given antibody measurement and reduces usefulness of harmonization studies.

To address these problems, NIST has engaged with external stakeholders, including Roche, Abbott, and Regeneron, to develop SARS-CoV-2 monoclonal antibodies (mAbs) as reference materials. Being manufactured proteins, mAbs have well-defined concentrations that are known in absolute terms. However, conformational differences between mAbs relative to the SARS-CoV-2 antigens suggests that these references may yield significant measurement variation when tested against the variety of commercial assays. Thus, it is necessary to quantify the mAb induced uncertainties in order to assess their usefulness as reference materials.

To address this problem, we have undertaken an inter-lab study that combines elements of dark-uncertainty analyses with physics-based intuition to separate distinct sources of measurement variation. While development of the uncertainty quantification is on-going, we have solved several preliminary issues associated with data normalization. The characteristic process for quantifying antibody levels is to measure both a reference material and a test sample; the number of antibodies in the latter is then assumed proportional to the ratio of the corresponding measurement values, e.g., as expressed in fluorescence units. However, this method assumes that the instrument and assay operate in a linear regime wherein each fluorescence signal is proportional to the actual antibody levels. In practice, identifying the linear regime requires serial measurements at increasing sample dilution, and even then, the choice of what constitutes linear is subjective. Moreover, using a single measurement from an entire dilution series wastes valuable information.

To overcome these limitations, we have developed a method that uses constrained optimization to map entire dilution curves onto one another. The key observation is that the total fluorescence signal should be a monotone increasing function of antibody concentration x, which is equal to the undiluted antibody concentration c times the dilution factor d. Because only the first of these is unknown, each dilution curve is a function of a single unknown parameter, which can be determined by solving an optimization problem that collapses all the curves onto one another. By construction, each value of c is a quantification of the number of antibodies. Such estimates are more stable than their single-dilution-factor counterparts, which do not use all of the available data. A manuscript on this work is in progress.

Metrology for Cytometry

Paul Patrone
Anthony Kearsley
Danielle Middlebrooks
Rayanne Luke
Prajakta Bedekar
Gregory Cooksey (NIST PML)
Matthew DiSalvo (NIST PML)
Jalal Sadeghi (NIST PML)

For more than 30 years, flow cytometry, a technique used to measure characteristics of cells, has been a mainstay for cancer detection, drug development, and biomedical research. It has remained a primarily qualitative metrology platform, however, because measurement uncertainties associated with this technique are so large. While exact economic figures are difficult to estimate, this has clearly had a significant impact on the roughly $200 billion of waste in the healthcare industry and contributed to the broader reproducibility crisis in biomedical research [1]. The challenge of making cytometry an accurate and precise metrological tool arises from the competing requirement that it have high throughput. Typical biological samples can have up to hundreds of millions of cells, which must be analyzed over a few hours.

To achieve this throughput, cytometers direct cells through a microfluidic channel at high-speed, past an optical interrogation region that collects fluorescence light from antibodies attached to surface proteins. The total fluorescence collected from each cell should then, in principle, be proportional to the total number of markers on its surface. But in practice, this idealized picture is complicated by the cumulative effects of the physical phenomena involved: fluid-dynamic forces cause cells to move across streamlines and/or have unpredictable trajectories; optical geometric collection efficiencies depending on position in the interrogation region; and signal acquisition and processing tools introduce non-linear effects and measurement uncertainties through discrete sampling. These challenges, in addition to the complexity of exactly replicating the necessary measurement infrastructure at the micron scale, have made it
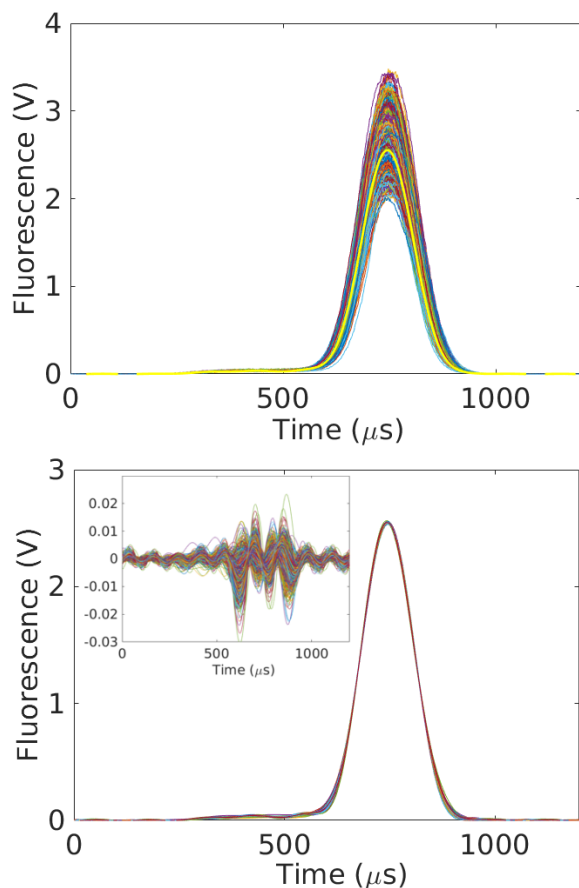
Figure 34. *Top. 1300 cytometry signals corresponding to particles having different velocities and sizes. Bottom. Our theory predicts that the 1300 curves in the top figure can be collapsed onto one another using straightforward linear transformations that depend on the particles' velocities and sizes. The agreement with this theory is on the order of 2% relative to the maximum scale of the measurements. The residuals in the inset can be interpreted as the per-event reproducibility of the measurement process.*

virtually impossible to reproduce measurements on a single cell, a necessary first step towards fully assessing and controlling uncertainties in cytometry.

ACMD, PML, and MML staff were awarded a NIST Innovations in Measurement Science (IMS) award to develop a microfluidic-based cytometer whose design explicitly allows control and study of repeat measurements of cells. Following on work from FY 2020, we are continuing a series of research projects on uncertainty quantification (UQ) for cytometry. The first of these aims to quantify the per-event uncertainty associated with each method. Recently we derived a key result indicating that all cytometry events (e.g., for a fixed cell-marker type) are identical up to a set of straightforward linear transformations that depend on physical parameters such as the cell size, speed, and number of biomarkers; see Figure 34. Using optimization, we can determine these unknown parameters by mapping different signals onto one another. Critically, this also yields multiple, model-based but distinct realizations of

what a measurement would look like could it be reproduced on the same measurand. Using this, we have been able to estimate for the first time the per-event uncertainty; see Figure 34. This analysis also enables new types of data analysis associated with doublet identification, doublet deconvolution, and more generally, cell counting. This work is summarized in a manuscript that has been submitted for publication [2].

The second UQ project aims to identify the probabilities of measurement outcomes for populations that may have some overlapping signals. The key assumption of this analysis is that the probability densities (PDFs) of measurement values for different populations have partially disjoint supports. Given a "mixed" PDF associated with an unknown combination two or more populations, one can only subtract off the PDF of one population from this mixture until the latter becomes negative. This implies an optimization problem wherein one seeks to maximize the amount of subtracted population subject to a positive-definite constraint. Work is ongoing to apply this algorithm to real cell systems and address issues associated with signal noise.

[1] W. H. Shrank, T. L. Rogstad, and N. Parekh. Waste in the US Health Care System: Estimated Costs and Potential for Savings. *JAMA* **322**:15 (2019), 1501-1509.

[2] P. N. Patrone, M. DiSalvo, A. J. Kearsley, G. B. McFadden, and G. A. Cooksey. Reproducibility in Cytometry: Signals Analysis and its Connection to Uncertainty Quantification. In review.

Cell Counting with Flow Cytometry

Ryan M Evans
Anthony J. Kearsley
Sandra M. da Silva (NIST MML)
Kirsten Parratt (NIST MML)

The gut microbiome and its metabolic by-products play an important role in a vast set of health conditions including inflammatory bowel disease, diabetes, obesity, and cancer [1-3]. Assessing gut microbiome health poses a challenge since it often requires an invasive intestinal biopsy, which is typically limited to patients who are suspected of having a disease. Although microbial communities of fecal samples differ from those taken from the intestinal lumen or mucosa of healthy and diseased patients, microbial communities of fecal samples are far easier to collect and may still serve as a useful way of characterizing the gut microbiome.

To this end Sandra Da Silva and Kirsten Parratt of the NIST Materials Measurement Laboratory have shown that flow cytometry is an effective way of quantifying cell types and determining the number of genome equivalents associated with a given stool sample. A typical flow cytometry signal is shown in Figure 35. Each element of the horizontal axis denotes
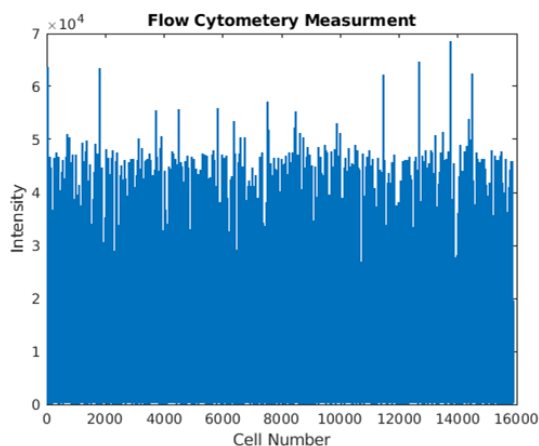
Figure 35. *Flow cytometry measurement.*
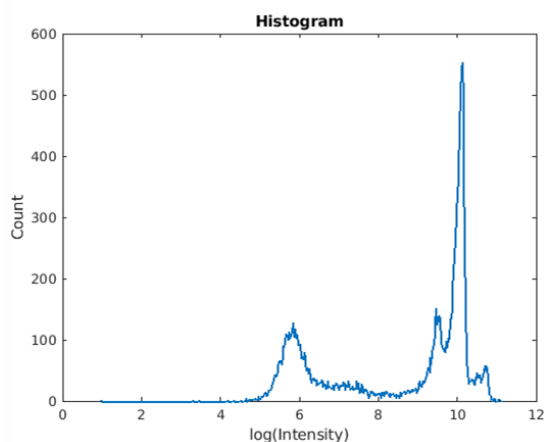
Figure 36. *Flow cytometry measurement histogram.*
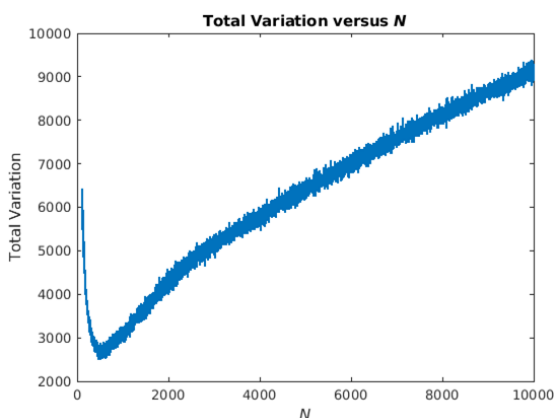
Figure 37. *Total variation as a function of N.*

detection of a biological cell, so there are approximately 16 000 cells in total. Associated with each of these cells is an intensity, which is given on the vertical axis. The number of cells at certain intensities correspond to physiologically important features, and this information can be obtained by mapping the measurement shown in

Figure 35 to a histogram, shown in Figure 36.

To move from the measurement shown in Figure 35 to the histogram in Figure 36, the vertical axis of the former is first partitioned into N subintervals. The number of cells that fall within each subinterval may then be counted to create the plot shown in Figure 36. Essential to this process is the choice of the number of subintervals, and we want to choose this number to minimize the noise. The total variation

$$\sum_{i=1}^{N-1} |y_{i+1} - y_i|$$

was used as an estimate of noise, and for each of the signals N was chosen to minimize this quantity. Figure 37 shows the total variation as a function of N for one of the signals, and the histogram corresponding to the minimum value of N in Figure 37 is shown in Figure 36.

Since the three right-most right peaks of Figure 36 correspond to physiologically relevant features, useful information can be discerned by counting the number of cells associated with them. This can be done by using a mathematical procedure that determines the number of peaks, such as the one developed by Wallace, Kearsley, and Guttman in [4]. However, inspection of Figure 36 reveals that there is still noise which will serve as a source of uncertainty in these counts. To remove this noise stochastic regression denoising was applied, as developed by Kearsley, Gadhyan, and Wallace in [4]. Moreover, stochastic regression was applied in such a way as to determine the *optimal* separation of signal from noise, as performed in [5] in the context of biological field effect transistors by Melara *et al*. The results are the decomposition shown in Figure 38 and the noise estimate shown in Figure 39. The orange curve in Figure 38 will provide a more accurate measure of the number of cells associated with physiologically relevant figures, and the noise estimate in Figure 39 can be used to help construct uncertainty bounds associated with these counts.

Da Silva and Parratt have taken a series of measurements that correspond to different samples and applying these ideas to each of them will give an accurate approximation of how many cells are associated with physiologically important features of each sample. The measurements considered herein are one-dimensional, however 2D measurements may provide even more information and may be used to confirm our one-dimensional estimates, so analysis of two-dimensional measurements may be a subject of future investigation.

[1] I. Ahmed, R. Greenwood, B. de Lacy Costello, N. M. Ratcliffe, and C. S. Probert. An Investigation of Fecal Volatile Organic Metabolites in Irritable Bowel Syndrome. *PLoS One* **8**:3 (2013), e58204.
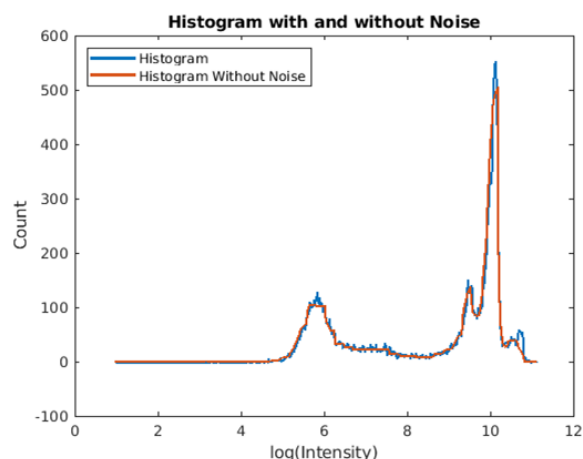
Figure 38. *Histogram before and after stochastic regression denoising.*
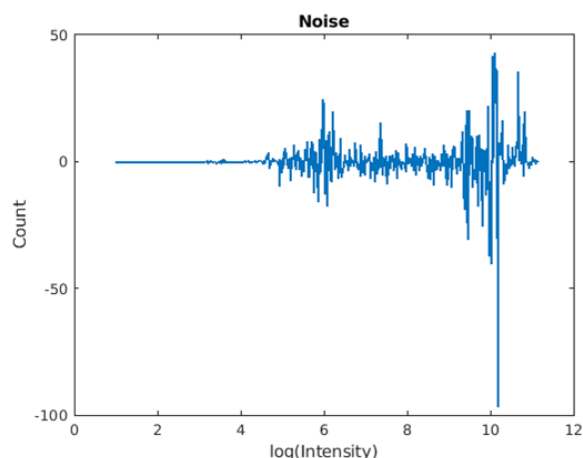
Figure 39. *Noise associated with histogram in Figure 36.*

[2] S. Anand and S. S. Mande. Diet, Microbiota and Gut-Lung Connection. *Frontiers in Microbiology* **9** (2018), 2147.

[3] D. S. Wishart. Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiological Reviews* **99**:4 (2019), 1819-1875.

[4] W. E. Wallace, A. J. Kearsley, and C. M. Guttman. An Operator-Independent Approach to Mass Spectral Peak Identification and Integration. *Analytical Chemistry* **76**:9 (2004), 2446-2452.

[5] L. A. Melara, R. M. Evans, S. Cho, A. Balijepalli, and A. J. Kearsley. Optimal Bandwidth Selection in Stochastic Regression of Bio-FET Measurements. *In review.*

Model Performance Predicting Pure Triglyceride Thermodynamic Properties

Anthony J. Kearsley
Arun S. Moorthy (NIST MML)
Julia Seilert (Technische Universität Berlin, Germany)
Eckhard Flöter (Technische U. Berlin, Germany)

Consisting of a glycerol backbone with three fatty acids attached (see Figure 40), triglycerides provide structure and function to vegetable and animal fats which play an important role in food science and engineering. Melting and solidification of fats in multi-component systems directly impacts product characteristics such as melting range and solid-phase composition. These compounds are also a primary element of stores of adipose. Triacylglycerols form an important part of the human diet serving as an energy source when stored in fat tissues. They provide thermal and mechanical protective layers surrounding important organs. Their primary source in the human diet are plant oils.

There is a significant interest in predicting how these fats melt, as this has a tremendous effect on storage/transportation and commercial viability of food products that contain fats (i.e., almost all food products). Thus, building mathematical models that accurate predict fat melting on the product development cycle can have a significant economic impact on the food.

Together with Technische Universität Berlin, we have investigated mathematical model fits, performance, and predictive power when imposing thermodynamic constraints on the predictions. A well-known model to predict pure component properties of triglycerides proposed by Wesdorp has been shown to perform well despite making thermodynamically inconsistent predictions for certain test cases. In this study, the underlying parameter set is improved to deliver more physically consistent predictions, i.e., increasing melting point and enthalpy of fusion with increasing stability of the polymorphs, without deterioration of the primary model quality to describe the available experimental data. Interestingly, when a curated dataset containing
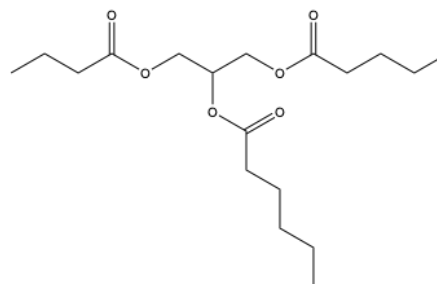
Figure 40. *An example structure drawing of a triglyceride molecule; the most common constituents in dietary fats (e.g., butter, margarines, oils).*

only thermodynamically consistent data is compared to a broader dataset, it appears that the model's efficacy is highly dependent on the quantity of data, specifically the number of unsaturated triglycerides data. Quality and thermodynamic consistency of model predictions and the condition of a reliable description of monoacid triglycerides as a subset is discussed, addressing a potential interdependence.

A manuscript describing this work has recently been published [1]

[1] J. Seilert, A. S. Moorthy, A. J. Kearsley, and E. Flöter. Revisiting a Model to Predict Pure Triglyceride Thermodynamic Properties: Parameter Optimization and Performance. *Journal of the American Oil Chemists' Society* **98**:8 (2021), 837-850. DOI: 10.1002/aocs.12515

Modeling for Biological Field Effect Transistor Experiments

Ryan M. Evans
Luis Melara (Shippensburg University)
Seulki Cho (NIST PML)
Arvind Balijepalli (NIST PML)
Anthony Kearsley

Biological field effect transistors (Bio-FETs) are modern bioelectronics instruments offering novel biomarker measurements. In contrast with traditional measurement techniques that require specialized facilities and expensive equipment, Bio-FETs offer rapid, accurate and low-cost measurements. Since these instruments are hand-held and portable, they promise to yield wider accessibility to critical medical diagnostic tests. During a typical experiment, a chemical reactant bath is injected into a solution-well that contains a buffer fluid. These chemical reactants diffuse through the solution-well and bind with chemical reactants confined to a thin layer known as the biochemical gate on the sensor surface. This produces a time-series signal that can be used to analyze the chemical reaction of interest. See Figure 41 for an experimental Bio-FET signal.

Since estimating parameters associated with these experiments like kinetic coefficients can help us identify biomarkers, accurate mathematical models for these experiments are desired. Previous mathematical models have either not accounted for the time dependent nature of Bio-FET experiments [1–3] or have not included physically relevant transport processes [4–5]. A model recently published by Evans, Kearsley, and Balijepalli is the first dynamic model for Bio-FET experiments that accounts for physically relevant transport effects, and it has been shown that this model compares favorably with experimental data [6]. This model together with experimental developments by Balijepalli and Cho have led to interesting areas of inquiry.
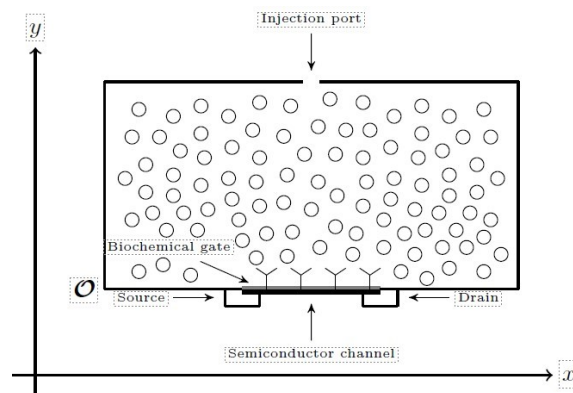
Figure 41. *Idealized cross-sectional schematic of a Bio-FET experiment. Chemical reactants injected at the top of the instrument diffuse to bind with receptors confined to the biochemical gate. In our schematic, chemical reactants in solution are represented as circles spread throughout the instrument, and receptor sites are depicted with Y's on the biochemical gate. This schematic is not to scale. The orientation of the x and y axes have been shown on the left, along with the location of the origin O.*
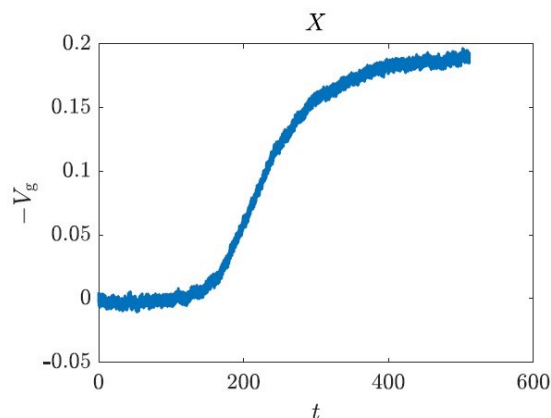
Figure 42. *Experimental Bio-FET signal. The horizontal axis represents time in seconds, and the vertical axis represents voltage that is measured as a result of surface binding.*
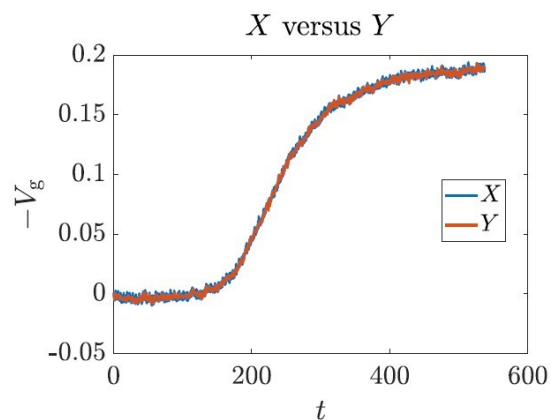
Figure 43. *A comparison between an experimental Bio-FET signal X, and its reconstruction using stochastic regression Y computed with an optimal bandwidth parameter h. These two curves are nearly indistinguishable.*

The first concerns the issue of noise, as illustrated in Figure 42. Bio-FET signals contain noise that interfere with quantitative analysis. Kearsley, Gadhyan, and Wallace have shown that stochastic regression is an effective technique for dealing with noise in the context of mass spectrometry data [7]. This approach centers around modeling the time series X with a stochastic differential equation (SDE) that consists of a linear drift term and a stochastic diffusion term, to wit:

$$dX_t = (a_0(t) + a_1(t)X_t)dt + b(t)dW, \qquad (1)$$
$$X(0) = X_0.$$

The drift coefficients $a_0(t)$ and $a_1(t)$ are determined through solving a local weighted least squares problem, and the diffusion coefficient $b(t)$ is determined through maximum likelihood estimation. Since the model (1) is linear, a Euler-Maruyama discretization of (1) can then be used to extract the estimated deterministic and stochastic components of the signal. The performance of this method relies on the size of the local averaging window, which is set by the bandwidth parameter h. Larger values of h correspond to larger averaging windows, and smaller values of h correspond to smaller averaging windows. If the bandwidth parameter is too large, then the estimated deterministic and stochastic components may not match with experimental data. The same problem may also occur if the bandwidth parameter is too small. So there exists an *optimal* bandwidth parameter h that is neither too large nor too small. Professor Luis Melara, a faculty member in the Department of Mathematics at Shippensburg University of Pennsylvania, together with Evans, Cho, Balijepalli, and Kearsley, have determined optimal bandwidth parameters for a biological field effect transistor experiment [8]. A comparison of an experimental Bio-FET signal X with its reconstruction using stochastic regression Y computed with an optimal band h is shown in Figure 43. In this work it was also shown that this parameter is remarkably robust with respect to the *aspect ratio ϵ*, which is the ratio of the radius of the solution-well height to the biochemical gate diameter.

This project has also led to further interesting collaborations between ACMD and the Microsystems and Nanotechnology Division of the NIST Physical Measurement Laboratory. Balijepalli and Cho of the latter recently fabricated a Bio-FET by conjugating thiol-modified biotin to a gold surface to form a self-assembled monolayer (SAM). The interactions of streptavidin with the SAM were then measured dynamically. Preliminary results show that it is possible to extract rate constants and a diffusion coefficient that characterize the reaction by fitting the model developed by Evans, Balijepalli, and Kearsley to time series data. These encouraging results indicate that the technique may broadly benefit biomolecular interactions of proteins with ligands, antibodies with antigens and numerous other systems to drive new applications in drug discovery and diagnostics of disease.

The model presented in [6] has also led to interesting numerical analysis questions. Therein it is shown that the model takes the form of a nonlinear integrodifferential equation, with a singular kernel that behaves like $1/\sqrt{t}$ as t approaches zero. So finding a numerical approximation to the solution of this nonlinear equation can be difficult for small t when the size of the problem becomes very large, which can result in an ill-conditioned system. Kearsley observed that formulating the problem in a clever way lends it amenable to techniques developed by Professor Josef Sifuentes at the University of Texas Rio Grande Valley. Josef's research has focused on finding accurate preconditioners for Karush-Kuhn-Tucker (KKT) systems for which the generalized minimum residual method (GMRES) converges in as few iterations as possible. Preliminary research shows that his techniques ameliorate the effect of the singularity for very large systems, which aids in finding the solution to large-scale problems.

There are other interesting problems on the horizon for future research, one of which concerns the biotin and streptavidin kinetics, for which cooperativity remains an open question. Takeshi and Cantor have employed gel electrophoresis to argue that biotin-streptavidin is cooperative [9]. However other studies have suggested that sequential binding of biotin to tetrameric streptavidin occurs in a noncooperative fashion, with the four binding sites being thermodynamically equivalent and independent [10–11]. More recent work has suggested that although each of the binding sites on streptavidin may be kinetically independent, sequential binding may induce conformational changes [12]. This is a promising avenue of future investigation. We conclude by noting that experimental advances may also make it possible to use the modeling framework developed in [6] to quantify DNA interactions.

[1] S. Baumgartner, M. Vasicek, A. Bulyha, N. Tassotti, and C. Heitzinger. Analysis of Field-Effect Biosensors Using Self-Consistent 3D Drift-Diffusion and Monte-Carlo Simulations. *Procedia Engineering* **25** (2011), 407-410.

[2] C. Heitzinger, N. J. Mauser, and C. Ringhofer. Multiscale Modeling of Planar and Nanowire Field-Effect Biosensors. *SIAM Journal on Applied Mathematics* **70**:5 (2010), 1634-1654.

[3] D. Landheer, G. Aers, W. R. McKinnon, M. J. Deen, and J. C. Ranuarez. Model for the Field Effect from Layers of Biological Macromolecules on the Gates of Metal-Oxide-Semiconductor Transistors. *Journal of Applied Physics* **98**:4 (2005), 044701.

[4] X. Duan, Y. Li, N. K. Rajan, D. A. Routenberg, Y. Modis, and M. A. Reed. Quantification of The Affinities and Kinetics of Protein Interactions Using Silicon Nanowire Biosensors. *Nature Nanotechnology* **7**:6 (2012), 401-407.

[5] G. Tulzer, S. Baumgartner, E. Brunet, G. C. Mutinati, S. Steinhauer, A. Kck, P. E. Barbano, and C. Heitzinger. Kinetic Parameter Estimation and fluctuation Analysis of CO at SNO2 Single Nanowires. *Nanotechnology* **24**:31 (2013), 315501.

[6] R. M. Evans, A. Balijepalli, and A. J. Kearsley. Transport Phenomena in Biological Field Effect Transistors. *SIAM Journal on Applied Mathematics* **80**:6 (2020), 2586-2607.

[7] A. J. Kearsley, Y. Gadhyan, and W. E. Wallace. Stochastic Regression Modeling of Chemical Spectra. *Chemometrics and Intelligent Laboratory Systems* **139** (2014), 26-32.

[8] L. A. Melara, R. M. Evans, S. Cho, A. Balijepalli, and A. J. Kearsley. Optimal Bandwidth Selection in Stochastic Regression of Bio-FET Measurements. In review.

[9] T. Sano, and C. R. Cantor. Cooperative Biotin Binding by Streptavidin. Electrophoretic Behavior and Subunit Association of Streptavidin in the Presence of 6 M Urea. *Journal of Biological Chemistry* **265**:6 (1990), 3369-3373.

[10] M. L. Jones and Gary P. Kurzban. Noncooperativity of Biotin Binding to Tetrameric Streptavidin. *Biochemistry* **34**:37 (1995), 11750-11756.

[11] L. Deng, E. N. Kitova, and J. S. Klassen. Dissociation Kinetics of the Streptavidinbiotin Interaction Measured Using Direct Electrospray Ionization Mass Spectrometry Analysis. *Journal of The American Society for Mass Spectrometry* **24**:1 (2012), 49-56.

[12] M. J. Waner, J. M. Hiznay, A. T. Mustovich, W. Patton, C. Ponyik, and David P. Mascotti. Streptavidin Cooperative Allosterism Upon Binding Biotin Observed by Differential Changes in Intrinsic Fluorescence. *Biochemistry and Biophysics Reports* **17** (2019), 127-131.

Mathematical Models for Cryobiology

Daniel Anderson
Anthony Kearsley
James Benson (University of Saskatchewan, Canada)
Jessica Masterson (George Mason University)

Cryobiology, the study of biological specimens at cryogenic temperatures, plays an enormous role in a wide range of fields. In the field of medicine, cryobiology is the basis for cryopreservation in assisted reproduction, organ transplantation, biobanking and personalized medicine. Cryo-banking is used in the agriculture industry as well as for initiatives aimed at preserving rare and endangered plant and animal species and in the development of more productive agricultural yields. Applications in forensics arise in the processing and preservation of frozen biological samples that are often important and fragile evidence in criminal investigations. The breadth and depth of these applications reflect the complexity of the biological, chemical, and physical
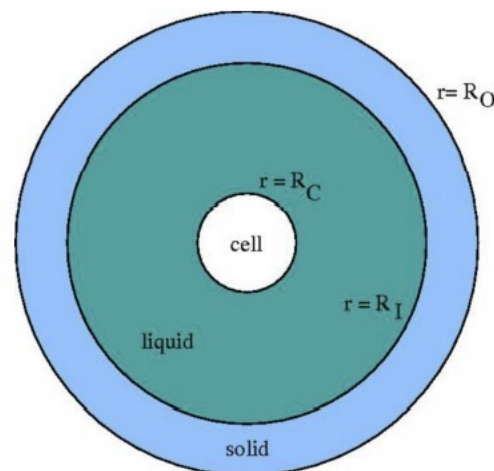
Figure 44. *Spherical cell surrounded by an encroaching solid—liquid interface.*
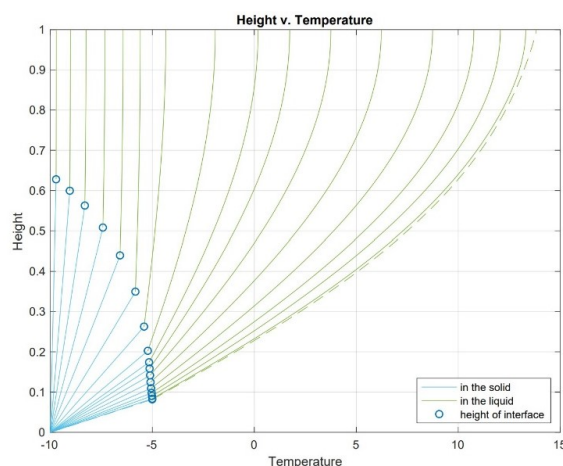
Figure 45. *Time sequence of thermal profiles in solid (blue) and liquid (green) regions controlling the temperature adjacent to a biological sample.*

aspects required to describe and model these problems. Mathematical and computational models can be used to probe these complex systems and, in conjunction with sophisticated control and optimization schemes, can establish more effective protocols for cryopreservation.

Cryopreservation of a human cell is a form of biomimicry that attempts to do in the laboratory, or in silica, what some frog and amphibian species can do naturally. A cryoprotectant, effectively an antifreeze, aims to play the role that substances like glucose play in the winter preservation of frozen frogs. These cryoprotectants, which are added to the extra-cellular environment, help to remove water from a cell before cooling, thus reducing the likelihood of intracellular ice formation. This, however, comes at a price due to the cell's limited ability to withstand the elevated levels of chemicals, including the naturally occurring salts in the cell, possibly damaging the cell, or causing death by chemical toxicity. Thus,

the maintenance of a viable cell during cryopreservation is complicated by two primary factors. Cooling the cell too quickly increases the likelihood of intracellular ice formation and cell damage or death while cooling the cell too slowly can overexpose the cell to high solution concentrations and lead to chemical toxicity. This amounts to a high-stakes goldilocks problem from the cell's perspective.

Mathematical modeling in cryobiology thus requires a detailed understanding of thermal and chemical transport in bulk phases as well as across a semipermeable cell membrane. Additionally, phase transformation of these multicomponent solutions – phenomena that link cryobiology to a wide range of other fields from geophysics to industrial materials processing [6] – must also be included. Cryopreserving a cell requires a delicate balance between two competing damage mechanisms, and thus a delicate optimization problem. We are exploring foundational aspects of biochemical and physical modeling in cryobiology, computational methods for the solution of these models, and applications of these ideas to cryopreservation of cells.

The focus of [1] was to establish the foundations of the chemical thermodynamics necessary to describe transport processes during cryopreservation. This work formulated chemical potentials and related thermodynamics quantities for non-dilute and non-ideal multicomponent solutions of experimental and theoretical interest to cryobiologists. Next, the multiphase, multi-species transport equations were developed along with a consistent characterization of cell membrane dynamics and solid-liquid phase transitions [2]. A critical aspect of our work was to obtain, from first principles, mathematical models that address both spatial and temporal dynamics of chemical species and heat transport. Various aspects of the freezing of a spherical biological cell (see Figure 44) were addressed in [3] and [4]. Based on a numerical algorithm outlined in [4], we explored in [3] the evolution of the thermal fields in the solid, liquid, and intracellular regions along with the concentration of cryoprotectant and the intra- and extracellular salts. These studies incorporated the effects of confinement and partial solute rejection, which had not previously been examined in cryobiology. These observations led to the development of objective functions in [4] which provide a measure of both intracellular undercooling and chemical toxicity. These control functions appear more suited to deriving cooling protocols than previously employed toxicity functions [5].

The mathematical models and computational algorithms established in this foundational series of papers [1-4] provide the launching point for our current work towards optimal control of these systems. Our current work underway with J. Masterson involves the identification of proxies for cell/tissue damage that can be used in optimal control settings. The mathematical framework is PDE-constrained optimization in which the bio/chemical/physical model constrains variables such as solid and liquid thermal fields (see Figure 45) that help define objective functions such as those measuring tissue undercooling. Our first objective is to address optimal protocols in the freezing/cooling stage. A future goal is to extend this framework to explore strategies that can optimize protocols for cryo-recovery where warming and/or melting occur. The central role played by phase transformation and transport in multicomponent systems links these cryobiological processes and their mathematical descriptions to related ones that occur under vastly different conditions in geophysics and industrial materials processing [6].

[1] D. M. Anderson, J. D. Benson, and A. J. Kearsley. Foundations of Modeling in Cryobiology I: Concentration, Gibbs Energy, and Chemical Potential Relationships. *Cryobiology* **69** (2014), 349-360. DOI: 10.1016/j.cryobiol.2014.09.004

[2] D. M. Anderson, J. D. Benson, and A. J. Kearsley. Foundations of Modeling in Cryobiology II: Heat and Mass Transport in Bulk and at Cell Membrane and Ice-liquid Interfaces. *Cryobiology* **91** (2019), 3-17. DOI: 10.1016/j.cryobiol.2019.09.014

[3] D. M. Anderson, J. D. Benson, and A. J. Kearsley. Foundations of Modeling in Cryobiology III: Heat and Mass Transport in a Ternary System. *Cryobiology* **92** (2020), 34-46. DOI: 10.1016/j.cryobiol.2019.09.013

[4] D. M. Anderson, J. D. Benson, and A. J. Kearsley. Numerical Solution of Inward Solidification of a Dilute Ternary Solution Towards a Semi-permeable Spherical Cell. *Mathematical Biosciences* **316** (2019), 108240. DOI: 10.1016/j.mbs.2019.108240

[5] J. D. Benson, A. J. Kearsley, and A. Z. Higgins. Mathematical Optimization of Procedures for Cryoprotectant Equilibration using a Toxicity Cost Function. *Cryobiology* **64** (2012), 144-151. DOI: 10.1016/j.cryobiol.2012.01.001

[6] D. M. Anderson, P. Guba, and A. J. Wells. Mushy Layer Convection. *Physics Today*, February 2022, to appear.

Artificial Intelligence for Low-Field Magnetic Resonance Imaging

Andrew Dienstfrey
Zydrunas Gimbutas
Adele Peskin (NIST ITL)
Joe Chalfoun (NIST ITL)
Kathryn Keenan (NIST PML)
Kalina Jordanova (NIST PML)

Emerging low-field magnetic resonance imaging (MRI) systems offer the promise of low-cost, point-of-care imaging that could be conducted in, for example, rural locations, on the battlefield, and eventually even in an ambulance. However, present low-field MRI results in images with low spatial resolution and high noise. These
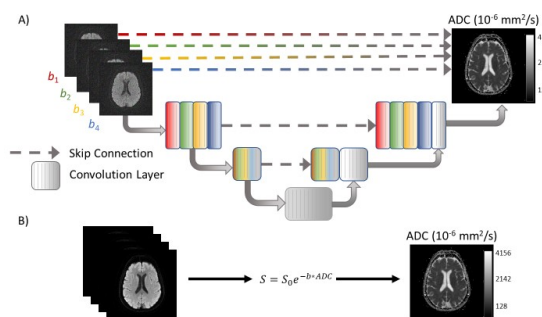
Figure 46. *Schematic of U-Net architecture. (A) The network shown here includes 4 input images for purposes for inverting for ADC map. The same architecture was used with a single input image for purposes of segmentation training and inference. (B) Traditional ADC analysis fits an exponential model on a pixel-by-pixel basis.*

Figure 47. *The new low-field MRI scanner arrived at the Katharine Blodgett Gebbie Laboratory in April 2021. The scanner can be moved using a joystick to drive built-in motorized wheels.*

poor qualities prevent quantitative analyses that are critical for advanced diagnostics. Recent research suggests that artificial intelligence (AI) methods can potentially restore image quality and furthermore may enable quantitative mapping of tissue parameters which are increasingly sought for medical diagnostics. [1] Such claims must be validated via rigorous comparisons to existing measurement standards for high-field MRI. In FY 2021 we initiated a new research project funded under the NIST Use-Inspired AI Program to investigate quantitative characteristics of AI-based analysis in low-field MRI measurement contexts.

Magnetic resonance imaging enables non-invasive imaging of biological soft tissue. Quantitative MRI refers to an extension of MRI in which scanning protocols and data analysis are implemented to obtain an *in vivo* map of measurable quantities whose magnitudes are reported in terms of standard physical units. For example, tumor volumes are reported in units of mm^3, apparent diffusion coefficient (ADC) in $mm^2\ s^{-1}$, and proton spin relaxation times, T_1 and T_2, are reported in seconds.

NIST has an ongoing, multi-year program to provide a robust measurement infrastructure for quantitative MRI. The recent publications [2] documenting the NIST Standard Reference MRI System Phantom, and [3] investigating the use of this phantom to assess long-term stability of T_1 and T_2 measurements are examples or recent outcomes.

While traditional high-field (i.e., greater than 1 Tesla) MRI is a critical medical imaging technology, it also extremely expensive, with U.S. spending exceeding \$10B/year [4]. The root cause of these costs are the helium-cooled, superconducting magnets required to generate 3 Tesla (T) static fields at the heart of such scanners. Direct costs for these exotic magnetic systems exceed \$1M per T. Furthermore, as the magnets weigh several tons and stray magnetic fields present significant safety concerns, high-field MR systems must be sited in structurally reinforced, shielded rooms, thereby incurring substantial indirect costs to house and maintain. Finally, to state the obvious, such a system is not portable. This last fact has consequences for medical outcomes in emergency situations, and in healthcare equity both domestic and international, as MRI scanners require sufficient proximate densities of well-insured patients to cover their substantial costs.

Low-field MRI systems operate at 6.5 mT to 550 mT. Critically, these field strengths do not require superconducting magnets but rather can be generated using, for example, Halbach arrays. Such systems have the potential to disrupt the entire MRI ecosystem. For one, low-field systems are 30x cheaper: a state-of-the-art scanner costs \$80K in comparison to more than \$3M for a traditional 3 T clinical system. Furthermore, low-field systems are portable, and in some cases may be light and small enough to fit in the back of a large van. However, this trade-off does not come for free: traditional reconstruction algorithms applied to low-field signals result in noisy, low-resolution images with limited diagnostic utility. Recent innovations in AI-based reconstruction show promise to resolve these barriers. In FY 2021 we initiated a project in this area using seed funding from a NIST Use-Inspired AI Program award. We briefly describe early-stage work on this project.

Quantitative MRI techniques exist to measure many different parameters. We selected the apparent diffusion coefficient (ADC) as the first measurand to be targeted. As ADC is a measurement of water diffusion it is insensitive to static magnetic field strengths. Thus, previous work in high-field ADC measurement should be directly applicable to low-field contexts.

Our first research question was whether a neural network could take a sequence of MR scans at different "b-values" and invert these to return a spatial map of diffusion coefficient. For training data, we used a dataset of high-field MR scans of the NIST Diffusion Phantom acquired by international collaborators as part of an ongoing study. This standard reference artifact consists of

NIST Diffusion Phantom   **AI Segmentations**
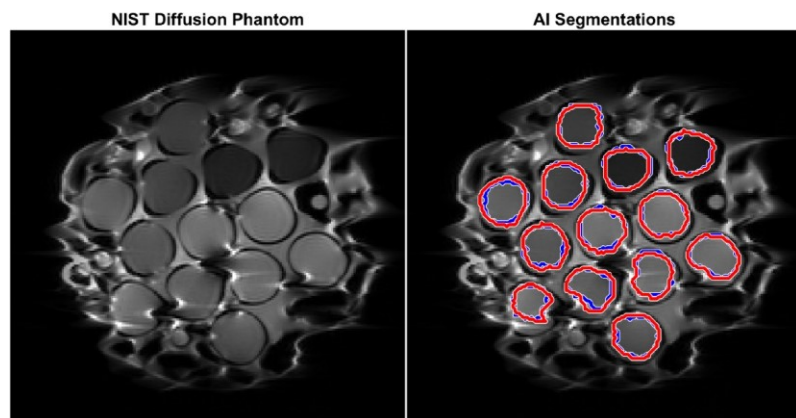
Figure 48. *Results of AI-based segmentation. On the left is an MRI scan of the NIST Diffusion Phantom. On the right we show the same image with the ADC regions inferred by two neural networks developed under this project. The red contours correspond to an FC-DenseNet and the blue to a U-Net.*

an array of tubes containing polymer solutions at different densities. Recent precision measurements conducted at NIST determined values for the ADC of these solutions. Thus, we have the possibility for an SI-traceable reference dataset to be used for training. Prior to executing this training, however, we required a segmentation of the dataset to identify the regions of interest within the MR images which contain the characterized solutions. We trained two neural networks for this task, one with a U-Net architecture and the other an FC-Densenet. The U-Net topology is shown for reference in Figure 46.

Approximately 20 images were segmented by hand by multiple team-members. These images were augmented using standard techniques to create a segmentation training set. Both networks were trained against this set and were used to infer segmentations of the remaining ≈400 images in the ADC phantom dataset.

A sample image and its region segmentation are shown in Figure 48.

The segmented database of MR images with the associated ground truth data for ADC values were then used to train additional neural networks to infer ADC from images. In the next year we plan to implement a detailed quantitative analysis of the segmentation results in comparison to traditional algorithms such as active contours. In addition, we hope to compare the ML-based ADC inference to standard workflow which consists of pixel-by-pixel non-linear regression for the parameters of an exponential model.

Low-field MRI is emerging as a clinical technology but is still very new. We are excited to report that in FY 2021 we completed a CRADA agreement with Hyperfine, a low-field MRI start-up company based in Guilford, CT. Hyperfine's Swoop low-field scanner is the first to be approved by the Federal Drug Administration for clinical medical use. Using $80K of the seed funding from the Use-Inspired AI Proposal we purchased a Swoop system which was delivered to NIST Boulder in April 2021. A picture of the scanner arriving to the Katharine Blodgett Gebbie Laboratory is shown in Figure 47. The NIST-Hyperfine CRADA states that NIST researchers will be granted "expert mode" access to the system and some aspects of its control and analysis software. To date, we have completed over 100 scans using the Hyperfine system and we have held extensive discussions
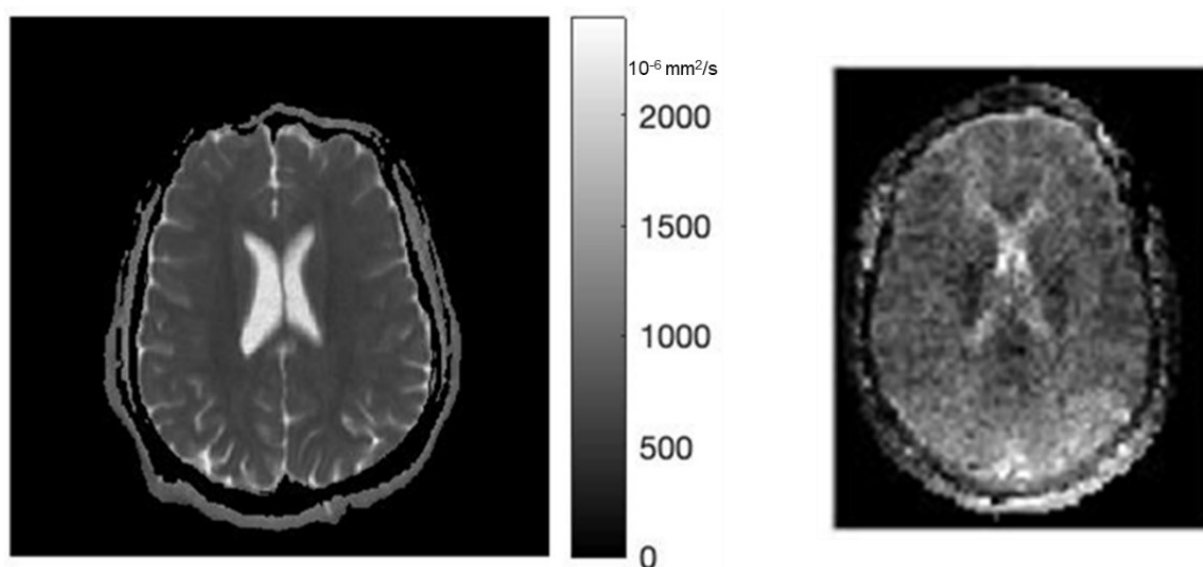
Figure 49. *ADC measurements of human brain acquired by NIST research team. The image on the left shows a high-field measurement using a clinical scanner operated by a collaborator. The same subject was scanned on the NIST-owned, low-field MRI.*

sharing our findings with company representatives. Detailed access to a state-of-the-art, low-field MRI scanner will be invaluable for our current research project.

Finally, in FY2021 all team members completed extensive training required for human-subjects research. A research proposal was approved by the NIST Research Protections Office and an external Internal Review Board, allowing us to conduct brain scans of human subjects on both standard and low-field MRI systems. As of late 2021, we completed scans on eight participants. An example of an ADC measurement of a subject as acquired on a clinical 3T machine in comparison to the NIST low-field system is shown in Figure 49. The lower resolution and higher noise in the low-field measurement is apparent. We will receive scanner protocol and software updates from Hyperfine in the near future.

In the upcoming year we plan to develop computational models of the low-field imaging process which will be used to inform a ML-based analysis of low-field ADC measurement.

[1] M. Figini, H. Lin, G. Ogbole, *et. al*. Image Quality Transfer Enhances the Contrast and Resolution of Low-Field MRI in African Paediatric Epilepsy Patients. arXiv (2020). URL: https://arxiv.org/abs/2003.07216

[2] K. F. Stupic, M. Ainslie, M. A. Boss, *et. al*. Standard System Phantom for Magnetic Resonance Imaging. *Magnetic Resonance in Medicine* **86** (2021) , 1194–1211. DOI: 10.1002/mrm.28779

[3] K. E. Keenan, Z. Gimbutas, A. Dienstfrey, *et. al*. Multi-site, Multi-platform Comparison of MRI T1 Measurement using the System Phantom. *PLOS ONE* (2021). DOI: 10.1371/journal.pone.0252966

[4] M. Rohman. JAMA: U.S. Spends the Most on Healthcare — and Imaging is a Reason Why." *Health Imaging* (2018).

Accuracy of Exposure Determination in Bluetooth-based Automatic Contact Tracing

Kamran Sayrafian
Brian Cloteaux
Vladimir Marbukh

The proximity detection mechanism in current automatic exposure notification systems is typically based on the Bluetooth signal strength from the individual's mobile phone. However, there is an underlying error in this proximity detection methodology that could result in wrong exposure decisions, i.e., false negatives and false positives. A false negative error happens if a truly exposed individual is mistakenly identified as not exposed. This misidentification could result in further spread of the virus by the exposed (yet undetected) individual. Likewise, when a non-exposed individual is
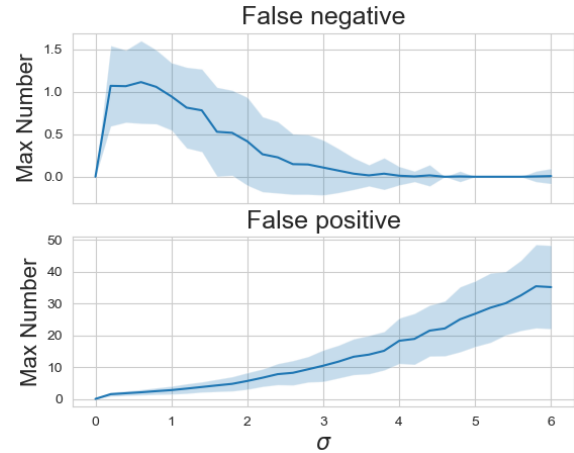
Figure 50. *Number of false exposure determinations versus σ (n=2)*

incorrectly identified as exposed, a false positive error occurs. This could lead to unnecessary quarantine of the individual, therefore incurring further economic cost.

In this project, using a simulation platform and a notion of proximity detection error, we investigate the performance of the system in terms of false exposure determinations [1]. Knowledge of how the Bluetooth-based proximity detection error impacts such false determinations and identification of methodologies that can reduce this impact will be helpful to enhance the effectiveness of an automatic contact tracing system.

Assuming a Lognormal pathloss distribution for the Bluetooth signal and a Gaussian distribution for the shadowing, fading, and scattering, we can obtain the following probability density function for the error in the estimated distance $f_Y(y)$,

$$f_Y(y) = \frac{1}{\sqrt{2n}} \frac{10n}{\sigma log_e(10)} \frac{1}{(d_0 + y)} \times$$
$$e^{-\frac{1}{2} \left(\frac{10n}{\sigma log_e(10)} \right)^2 \left(log\left(\frac{Y}{d_0} + 1 \right) \right)^2}$$

where d_0 is the true distances between the transmitter and receiver pair., n is the pathloss exponent and σ represents the standard deviation of the fading.

To understand the effect of the Bluetooth proximity estimation error on automatic exposure notification, a simulation platform has been developed to test scenarios involving people walking in a plaza or campus area. Initially, a population of agents is created and randomly placed within a closed simulation area. A certain percentage of the population is designated as being infected and contagious. The mobility pattern of the agents will obviously play an important role on the exposure possibilities. To conduct our initial study, we used the mobility algorithms in [2]. The algorithms allow us to select and modify both the individual parameters of each agent's behavior (affecting its speed and movement) as well as the goals that each agent is working towards. A potential difficulty with dynamic approaches in agent

mobility algorithms is the possibility of jamming [3]. Jamming can occur when all the agents in a simulation have similar goals, e.g., all trying to reach the same area within the simulation field. To avoid jamming and thus biasing the results in our simulations, we periodically randomize the goals of each agent. Using this platform, we track the true and estimated distances between any two moving agents at fixed time intervals. The estimated distance is calculated as the sum of the true distance plus a random value according to the probability density function of the error.

Figure 50 shows the max average numbers of false negatives/positives with the corresponding confidence intervals of one standard deviation versus σ. This result considers a population of 135 agents moving within an area of size 162 m × 35 m for 8 hours (i.e., typical length of a workday). These numbers are chosen based on a standard laboratory building inside the campus area of NIST.

The number of infected individuals at the beginning of the simulation is set to 5 % of the population. As observed, false positive counts noticeably increase as σ increases. With $\sigma = 5$ and during 8 hours of simulation, almost a quarter of the population will be erroneously identified as "exposed" (i.e., false positives). This implies that a significant percentage of the population could be required to go to quarantine in addition to the detected exposed people. The mobility pattern used in these simulations considers agents that are constantly moving. This leads to higher probability of agents being outside the 2 m radius of each other than inside. Agents that are outside the 2 m radius can only result in false positive type of erroneous exposure determination. Therefore, higher error intensity (i.e., σ) will increase the likelihood of agents that are farther away from an infected agent to be mistakenly estimated within the 2 m radius. This is the reason for the significant rise in the number of false positives with increasing σ. Similar trends in the results are also observed with lower percentages of infected agents at the beginning of the simulation; however, longer interactions between the agents (e.g., multiple days) might be needed.

We plan to extend the simulation platform by incorporating additional mobility patterns for the agents (e.g., patterns involving congregation) and investigate the effect of each pattern on the resulting false exposure determinations. In addition, a study of filtering functions and their potential relationship to the dynamics of the agents' movement (e.g., speed) might also be beneficial to minimize the impact of $\sigma\sigma$ on false determinations.

[1] K. Sayrafian, B. Cloteaux, V. Marbukh, and C. Emiyah. Evaluation of the Bluetooth-based Proximity Estimation for Automatic Exposure Determination. *IEEE Consumer Communications and Networking Conference (IEEE CCNC)*, Online, Jan. 8-11, 2022.

[2] C. Burstedde, K. Klauck, A. Schadschneider, and J. Zittartz. Simulation of Pedestrian Dynamics Using Two-Dimensional Cellular Automaton. *Physica A* **295** (2001), 507-525. DOI: 10.1016/S0378-4371(01)00141-8

[3] M. Muramatsu and T. Nagatani. Jamming Transition in Two-Dimensional Pedestrian Traffic. *Physica A* **275**:1–2 (2000), 281-291. DOI: 10.1016/S0378-4371(99)00447-1

A Study of Radio Frequency Propagation Inside the Human Skull for Neural Implants

Kamran Sayrafian
Mariella Särestöniemi (University of Oulu, Finland)
Carlos Pomalaza-Raez (Purdue University of Tech.)
Teemu Myllylä (University of Oulu, Finland)
Jari Iinatti (University of Oulu, Finland)

In recent years, there has been a growing interest in the development of non-invasive acute brain injury detection and monitoring systems. At the same time, advances in microelectronics are also contributing to significant progress in invasive monitoring systems such as intra-cranial pressure and brain oxygenation monitoring. As a result, Brain Computer Interface (BCI) as a multidisciplinary research area has drawn considerable attention due to its attractive and transformative applications across many verticals in wireless communications such as E-Health.

Low-power, reliability, and security are essential requirements for wireless monitoring of the brain signals. Some applications may require wireless links that can support data rates up to 100 Mbps. To meet these requirements, in-depth propagation studies at various frequency bands are essential in the design of a wireless brain telemetry system. Most of the existing studies on brain implant communication are based on simple tissue-layer models. Those models are more relevant for applications where the brain implant is relatively close to the skull. For wireless monitoring applications where the implant could be located deeper inside the brain, it is important to study propagation using more realistic models of the human head. It is also necessary to identify which frequency bands would be suitable for different brain monitoring or implant communication
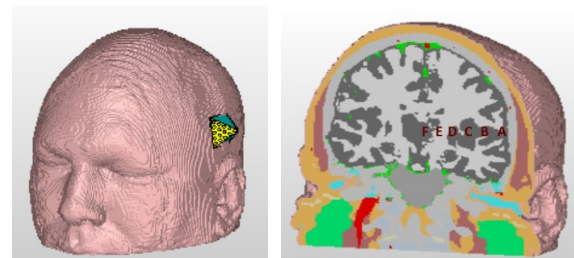
Figure 51. *Left: surface antenna location. Right: implant antenna locations.*
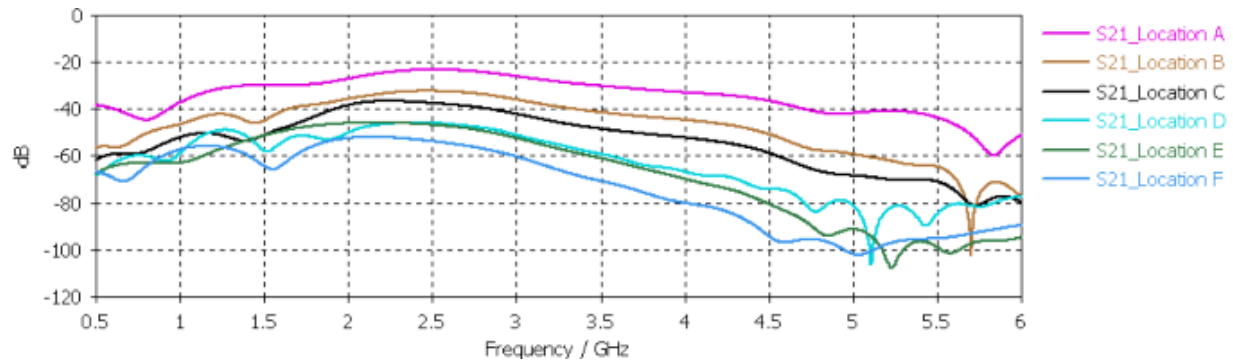
Figure 52. *Forward channel coefficient (S₂₁) for different implant locations.*

applications. This identification should take into the account all desirable requirements for the wireless communication link. The main objective of this project is to study radio waves propagation inside the human skull at several frequency bands.

Using the anatomical 3D head model from the CST's Hugo voxel-based body model, a directional bio-matched mini-horn surface antenna [1], and a double-loop UWB implant antenna [2], channel measurements can be done for a variety of antenna locations on the surface of the head and inside the brain. Sample locations for the surface and implant antenna are shown in Figure 51. The forward channel coefficients (S_{21}) at locations A-F are presented in Figure 52. Assuming that a maximum pathloss of 65 dB is tolerable in order to maintain a reliable communication link, it is observed that the channel attenuations at 914 MHz, 2.45 GHz, and 3.1 GHz are within the acceptable threshold for all considered propagation depths. Maximum attenuations for these frequencies (which occurs at location F) are -54 dB, -58 dB, and -63 dB, respectively. However, at 4 GHz, the channel attenuation is manageable only at the implant locations A-D. At implant locations E and F, a receiver with higher sensitivity would be required as channel attenuation for these locations are over 70 dB. At 5.8 GHz, only locations A and B experience channel attenuation within the 65 dB margin. Results for more antenna locations and the corresponding power flow plots can be found in [3].

While this study provides some insight into the RF propagation inside the human head, more application-specific research is needed to better understand and characterize the propagation media. The wide spectrum of brain telemetry applications will likely result in a diverse set of wireless connectivity solutions, necessitating multiple propagation channel studies. For neural implants, custom-designed antennas considering regulatory and safety issues as well as SAR limitation for the human brain are needed. These factors will impact the achievable propagation depth at a given frequency. Possibility of using array antennas to create efficient communication links will also require more in-depth study of the RF propagation inside the human

brain. Verification of the simulation results with properly designed head phantoms should also be conducted.

[1] J. Blauert and A. Kiourti. Bio-Matched Horn: A Novel 1–9 GHz On-Body Antenna for Low-Loss Biomedical Telemetry with Implants. *IEEE Transactions on Antennas and Propagation* **67**:8 (2019), 5054-5062.

[2] J. Shang and Y. Yu. An Ultrawideband Capsule Antenna for Biomedical Applications. *IEEE Antennas and Wireless Propagation Letters* **18**:12 (2019), 2548-2551.

[3] M. Särestöniemi, C. Pomalaza-Raez, K. Sayrafian, T. Myllylä, and J. Iinatti. A Preliminary Study of RF Propagation for High Data Rate Brain Telemetry. In *16th EAI International Conference on Body Area Networks* (BodyNets 2021), Online, October 25-26, 2021.

Modeling Photoreceptor Dynamics

Danielle C. Brager
Anthony J. Kearsley
Daniel Anderson

Photoreceptors are light-sensing cells in the retina that play an essential role in the vision process. Light enters the eye through the cornea and passes through the pupil before reaching the lens where the light is focused onto the retina. Photoreceptors convert light into electrical signals that are sent to the brain via the optic nerve so that we can see. There are two types of photoreceptors in the human retina - rods and cones. Rods are concentrated in the outer areas of the retina and are responsible for vision at low light levels while cones are concentrated in the macula, an area in the center of the retina, and are responsible for color vision as well as visual acuity. Photoreceptors have an inner segment (IS) and an outer segment (OS). The IS is the photoreceptor's metabolic center and is filled with mitochondria. The OS is filled with stacks of membranes that contain the visual pigment molecules. The photoreceptors undergo continuous renewal and periodic shedding of their OS discs to prevent the toxic effects of accumulated photo-oxidative
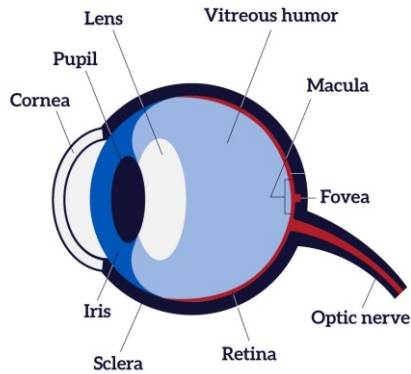
Figure 53. *Components of the eye. Image source : https://www.nei.nih.gov/learn-about-eye-health/healthy-vision/how-eyes-work*

products. Following a circadian rhythm that is governed by light and temperature modulation, discs formed at the OS base are discarded at the tip and phagocytized by the neighboring retinal pigment epithelium (RPE).

Retinal degenerations include a group of disorders that lead to photoreceptor loss. This photoreceptor cell loss is common to degenerative eye disorders such as retinitis pigmentosa, age-related macular degeneration, and cone-rod dystrophy. Research in this field is focused on developing strategies to delay or prevent the onset of photoreceptor degeneration [2, 3]. However, there is currently no cure for diseases linked to photoreceptor degeneration.

Mathematical modeling plays a crucial and beneficial role in this research effort [1]. There exist mathematical and computational models of various regions of the retina in retinal development, health, and disease. However, to our knowledge, there is no mathematical model of the physiology of photoreceptors in a healthy eye, that incorporates the interplay between the spatial density distribution, OS length, and nutrient source. We have developed a novel mathematical model of these processes to be used as a framework for future mathematical modeling in this field.

[1] D. C. Brager. *Modeling and Analyzing the Progression of Retinitis Pigmentosa*. Doctoral dissertation, Arizona State University, 2020.

[2] E. T. Camacho, L. A. Melara, M. C. Villalobos, and S. Wirkus. Optimal Control in the Treatment of Retinitis Pigmentosa. *Bulletin of Mathematical Biology* **76**:2 (2014), 292-313.

[3] E. T. Camacho, S. Lenhart, L. A. Melara, M. C. Villalobos, and S. Wirkus. (2020). Optimal Control with MANF Treatment of Photoreceptor Degeneration. *Mathematical Medicine and Biology: A Journal of the IMA* **37**:1 (2020), 1-21.

Materials Modeling

Mathematical modeling, computational simulation, and data analytics are key enablers of emerging manufacturing technologies. The Materials Genome Initiative (MGI), an interagency program with the goal of significantly reducing the time from discovery to commercial deployment of new materials using modeling, simulation, and informatics, is a case in point. To support the NIST role in the MGI, we develop and assesses modeling and simulation techniques and tools, with emphasis on uncertainty quantification, and collaborate with other NIST Laboratories in their efforts to develop the measurement science infrastructure needed by the materials science and engineering community.

OOF: Finite Element Analysis of Material Microstructures

Stephen A. Langer
Günay Doğan
Andrew C.E. Reid (NIST MML)
Prashant Athavale (Clarkson University)
Shahriyar Keshavarz (Theiss Research)

http://www.ctcms.nist.gov/oof/

The OOF Project, a collaboration between ACMD and MML, is developing software tools for analyzing real material microstructure. The microstructure of a material is the (usually) complex ensemble of polycrystalline grains, second phases, cracks, pores, and other features occurring on length scales large compared to atomic sizes. The goal of OOF is to use data from a micrograph of a real or simulated material to compute its macroscopic behavior via finite element analysis.

The OOF user loads images into the program, assigns material properties to the features of the image, generates a finite element mesh that matches the geometry of the features, chooses which physical properties to solve for, and performs virtual experiments to determine the effect of the microstructural geometry on the material. OOF is intended to be a general tool, applicable to a wide variety of microstructures in a wide variety of physical situations. OOF2 and OOF3D are used by educators and researchers in industry, academia, and government labs worldwide.

There are two versions of OOF, OOF2 and OOF3D, each freely available on the OOF website. OOF2 starts with two dimensional images of microstructures and solves associated two-dimensional differential equations, assuming that the material being simulated is either a thin freely suspended film or a slice from a larger volume that is unvarying in the third dimension (generalizations of plane stress and plane strain, respectively). OOF3D starts with three dimensional images and solves equations in three dimensions. Development this year continued on multiple fronts.

During this past year, A. Reid and S. Keshavarz made progress on the effort to implement crystal plasticity in the OOF software. The plastic computational scheme had been working for some time but was not yet integrated with the OOF mechanism for generating outputs. This work has now been completed, and an example stress-strain curve has been generated for a simple power-law plastic constitutive rule entirely within OOF3D.

There remain some issues with the set-up. It seems to be anomalously sensitive to some constitutive parameters and is currently a relatively low-performance solution. The next task is to identify opportunities for optimization which retain the existing basic functionality and the modular, extensible architecture.

Reid and Keshavarz are also continuing to work with collaborators in the Center for Hierarchical Materials Design to implement a machine-learning-based rapid evaluation scheme for plastic constitutive rules. Current efforts are focused on implementing an early solution into a basic FEM code framework.

G. Doğan has been developing algorithms to help automate segmentation and meshing of microstructure images. Segmentation identifies distinct regions in a microstructure, and a mesh properly aligned with the segmented regions and their boundaries leads to accurate finite element simulations of the microstructure physics. In FY2021, Doğan implemented texture segmentation algorithms using texture features and machine learning. Textures in images are repeated patterns of pixel values, which can be used to distinguish different regions in images. Doğan's implementation relied on predefined texture features to classify pixel locations into different classes using classification algorithms, such as support vector machines (SVM) and random forests.

Doğan worked with P. Athavale and his students, Peter Lef, Emmanuel Atindama from Clarkson University to implement algorithms for restoration of electron backscattering diffraction (EBSD) images. EBSD is an important imaging modality to analyze and understand microstructure images by measuring the crystal orientations at individual locations and generating a map of the orientations. However, EBSD maps are often noisy and incomplete; orientation measurements are missing in some locations, and some measurements contain errors. Doğan and his collaborators implemented a reconstruction algorithm using a weighted total variation equation,

a nonlinear partial differential equation that applies regularizing diffusion selectively to pixels. Their implementation included critical preprocessing components, specific to EBSD images. In this way, they were able to obtain state-of-the-art restorations of EBSD images. More details about Doğan's work on image segmentation and restoration can be found in this report's section on Computational Tools for Image and Shape Analysis.

OOF2 and OOF3D still rely on some old third-party software that will soon become obsolete. S. Langer has continued to work on the switch to the new versions of the aging software. In particular, the GUI toolkit gtk+2 needs to be upgraded to Gtk3, Python 2.7 to Python 3.x, and a substitute for the graphics library libgnomecanvas needed to be written because no Gtk3-compatible third-party software was available.

Much of the year was spent on updated the OOF2 GUI test suite to work with Gtk3. As expected, this uncovered numerous bugs and inefficiencies in both OOF2 and in the testing machinery. This also took an unexpected amount of time. Because the GUI test files needed to be rebuilt for Gtk3, this was a good time to make user-interface improvements that also broke the GUI test files. For example, users had been confused by the old way of determining when newly created objects (images, meshes, etc.) were automatically displayed in the graphics window. Now the user can choose one of three options, and hopefully find one that feels intuitive.

The replacement for libgnomecanvas, called OOFCanvas, is complete. The basic functionality was in place at the beginning of FY 2021, but the code was too tightly tied to OOF2 to be useful elsewhere. Now it is compiled into a stand-alone library which will be distributed separately from OOF2. Its API has been simplified and bugs have been fixed.

Langer and Reid have continued to work with A. Creuziger of MML who is using OOF2 to study the effects of texture (crystal orientation distribution) on the properties of rolled steel. In particular, the size of Creuziger's computations were revealing memory-use inefficiencies in OOF2, which now have been fixed.

OOF2 and OOF3D continue to be used heavily outside of NIST. OOF2 was downloaded about 700 times this past year, and OOF3D was downloaded about 200 times. OOF2 can be run on the NSF nanoHUB facility, where it was used 7499 times in FY 2021.

Micromagnetic Modeling

Michael Donahue
Donald Porter
Robert McMichael (NIST PML)
Cindi Dennis (NIST MML)

http://math.nist.gov/oommf/

Advances in magnetic devices such as field sensors, spin torque oscillators, magnetic nonvolatile memory (MRAM), and thermal sensors are dependent on an understanding of magnetization processes in magnetic materials at the nanometer level. Micromagnetics, a mathematical model used to simulate magnetic behavior, is needed to interpret measurements at this scale. ACMD is working with industrial and academic partners, as well as with colleagues in the NIST MML and PML, to improve the state-of-the-art in micromagnetic modeling.

We have developed a public domain computer code for performing computational micromagnetics, the Object-Oriented Micromagnetic Modeling Framework (OOMMF). OOMMF serves as an open, well-documented environment in which algorithms can be evaluated on benchmark problems. OOMMF has a modular structure that allows independent developers to contribute extensions that add to its basic functionality. OOMMF also provides a fully functional micromagnetic modeling system, handling three-dimensional problems, with extensible input and output mechanisms. Since October 1, 2020, the software was downloaded more than 9300 times by more than 5000 distinct client machines. There are 191 known peer-reviewed journal articles published in 2021 that acknowledge the use of OOMMF. Total OOMMF citations are now more than 3300. OOMMF has become an invaluable tool in the magnetics research community.

Developments in the last year include:

- Substantial reworking of OOMMF networking infrastructure for improved robust operations and lighter resource demands.

- Release of version 2.0a3 of OOMMF with a streamlined, more efficient launcher program.[8]

- Release of contributed OOMMF extension DMI_C2v, computing interfacial induced Dzyaloshinskii-Moriya interaction in systems with C2v symmetry.[9]

- Answering of more than 125 OOMMF user queries through email and other forums.

OOMMF is part of a larger activity, the Micromagnetic Modeling Activity Group (muMAG), formed to address fundamental issues in micromagnetic modeling

---

[8] https://math.nist.gov/oommf/software-20.html

[9] http://ipmras.ru/en/structure/people/tatarsky

through two activities: the development of public domain reference software, and the definition and dissemination of standard problems for testing modeling software. ACMD staff members are involved in development of the standard problem suite as well. This year muMAG published three contributed solutions to the standard problem suite[10] computed by the new Mag-Tense framework[11], demonstrating the continued importance of these benchmarks for micromagnetic modeling software development. A review of the muMAG standard problems has been published in a book chapter [1].

In addition to the continuing development of OOMMF, the project also does collaborative research using OOMMF. M. Donahue is a team member on the NIST Innovations in Measurement Science nanothermometry project Thermal MagIC[12]. Thermal MagIC work includes examination of uncertainty quantification with Mark-Alexander Henn of University of Maryland and antiferromagnetic modeling with University of Colorado graduate student Mingyu Hu (in collaboration with advisor Mark Hoefer). M. Donahue also provides technical guidance on micromagnetic modeling for the DARPA M3IC (Magnetic, Miniaturized, and Monolithically Integrated Components) project[13], which aims to integrate magnetic components into the semiconductor materials fabrication process to improve electromagnetic systems for communications, radar, and related applications. During the reporting period the ACMD micromagnetic modeling project produced two conference presentations. [2, 3]

[1] D. G. Porter and M. J. Donahue. Standard Problems in Micromagnetics. In Electrostatic and Magnetic Phenomena (M. J. Donahue, ed.), World Scientific, 2020, 285–324. DOI: 10.1142/9789813270268_0009

[2] A. Hu, M. Hoefer, E. Iacocca and M. J. Donahue. "Antiferromagnetic Spin Wave Propagation on Nonuniform Magnetic Backgrounds." Magnetism and Magnetic Materials Virtual Conference, November 5, 2020.

[3] M. J. Donahue and D. G. Porter. "High Order Methods for Computing the Demagnetization Tensor for Periodic Boundaries." Magnetism and Magnetic Materials Virtual Conference, November 6, 2020.

Numerical Methods for Reliable Computations with Equations of State

Bradley Alpert
Ian Bell (NIST MML)

Equations of state (EOS) provide quantitative relationships between thermodynamic variables of fluid mixtures, such as temperature, pressure, volume, and internal energy, that enable determination of key fluid properties by exploiting statistical physics to interpolate or extrapolate beyond where laboratory measurements are economical or feasible. The NIST data facilities REFPROP (Reference Fluid Thermodynamic and Transport Properties Database) and TDE (ThermoData Engine), which are widely accessed by the chemical engineering community, are developed from critically evaluated laboratory data in combination with EOS. Recently there has been a greater need to extend reference knowledge to fluids used in semiconductor manufacture.

Increasing reliance on, and automation of, retrieval of such thermodynamic property data imposes heightened demands on the reliability and generality of algorithms both for deriving and for exploiting the EOS. Calculation of the density of a mixture for a given temperature, pressure, and mixture composition from multiparameter mixture models sometimes fails, due to insufficiently accurate root finding, but also insufficiently robust algorithms that may miss certain roots entirely. The initial result of this collaboration was to represent transcendental equations in one dimension with orthogonal polynomial (Chebyshev) expansions
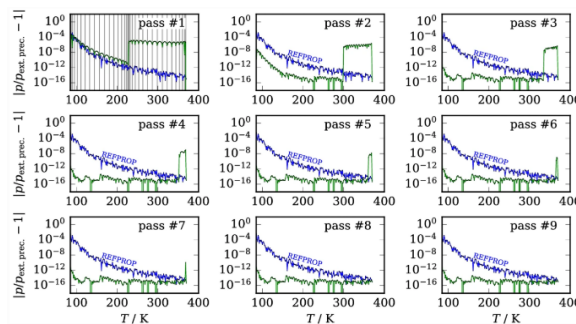
Figure 54. *Absolute value of the relative errors of pressure obtained from either REFPROP 10.0 (blue) or Chebyshev expansions (green) as the latter are adaptively refined. In the first pass, Chebyshev-Lobatto nodes are shown by vertical lines.*

---

10 https://www.ctcms.nist.gov/~rdm/mumag.org.html

11 https://www.magtense.org/

12 https://www.nist.gov/programs-projects/thermal-magic-si-traceable-method-3d-thermal-magnetic-imaging-and-control

13 https://www.darpa.mil/program/magnetic-miniaturized-and-monolithically-integrated-components

that enable reliable root finding. A concomitant requirement for very fast evaluations, achieved here, arises since typical applications may involve an extremely large number of root-finding steps.

This year this earlier-developed Chebyshev expansion capability was exploited adaptively to enable efficient and accurate representation of pure fluid phase equilibria, in particular the singular transitions between fluid phases, which will provide fundamental improvement in the modeling routines offered by REFPROP.

This work is part of a larger program to provide modeling capabilities for both pure fluids and fluid mixtures, including in temperature and density regimes where measurements are costly or infeasible.

[1] I. H. Bell and B. K. Alpert. Efficient and Precise Representation of Pure Fluid Phase Equilibria with Chebyshev Expansions. *International Journal of Thermophysics* **42** (2021), 75. DOI: 10.1007/s10765-021-02824-x

Mathematics of Infrastructure Metrology

Jeffrey T. Fong

To ensure the safe operation of a physical infrastructure or a part of it, be it a steel bridge, a highway network, a chemical processing plant, a nuclear power plant, or a jet airliner, engineers need to formulate and solve problems to determine three generic quantities or properties, namely, strength, reliability, and damage mechanisms.

The first property, *strength*, is encountered primarily during design, manufacturing, and construction to ensure that all parts and connections of the infrastructure have enough strength to work together as a whole without failure. The second property, *reliability*, requires attention of a life-long duration, because the goal is to evaluate the structural integrity of the entire infrastructure from birth to decommissioning. As the infrastructure ages, the strength of each of its parts and connections changes not only by wear and tear, and a host of corrosive environmental conditions, but also by a change of loadings due to natural disasters or new user needs. With input from sensor-based monitors of operating infrastructure today along with laboratory strength and damage test data, it is now possible to estimate the reliability of all properly maintained infrastructure to prevent premature failure. The third problem, as described in [1], is for the engineers and materials scientists to find out the *damage mechanisms* that cause the change in strength and reliability.

The measurement science, or metrology, that supports the science and engineering of physical infrastructure, or, *infrastructure metrology*, is the totality of experimental work, associated measurement data, and their interpretation that make the solution of the

three fundamental problems possible. Those experiments are to simulate the behavior of the material-based components, structures, and systems as they deform and fracture in various environments under static and dynamic loads. The mathematics of infrastructure metrology, therefore, is heavily involved in converting those measurement data into estimates of strength, reliability, and damage mechanisms. Two examples of such efforts in ACMD that were completed in FY 2021 are described in the project descriptions below, one involving the interpretation of laboratory failure test data for estimating minimum allowable strength of a full-size structure or component [2], and the other, the use of laboratory fatigue life and crack propagation test data for estimating the reliability of a cracked or uncracked component [3].

[1] J. T. Fong, ed. *Fatigue Mechanisms*. ASTM STP 675. ASTM International, West Conshohocken, PA, 1979.

[2] J. T. Fong, N. A. Heckert, J. J. Filliben, P. V. Marcal, and S. W. Freiman. Estimation of a Minimum Allowable Structural Strength Based on Uncertainty in Material Test Data. *Journal of Research of NIST* **126** (2021),126036. DOI: 10.6028/jres.126.036

[3] J. T. Fong, P. V. Marcal, R. Rainsberger, N. A. Heckert, J. J. Filliben, S. R. Doctor, and N. A. Finney, Jr. A Multi-Scale Failure-Probability-and-NDE-Based Fatigue Life Model for Estimating Component Co-Reliability of Uncracked and Cracked Pipes. *In Proceedings of the 2021 ASME Pressure Vessels & Piping Division Conference*, Online, July 12-16, 2021. DOI: 10.1115/PVP2021-62169

Estimation of a Minimum Allowable Structural Strength Based on Uncertainty in Material Test Data

Jeffrey T. Fong
James J. Filliben (NIST ITL)
N. Alan Heckert (NIST ITL)
Stephen W. Freiman (NIST MML)
Pedro V. Marcal (MPACT Corp.)

Three types of uncertainties exist in the estimation of the minimum fracture strength of a full-scale component or structure size. The first, called the *model selection uncertainty*, is in selecting a statistical distribution that best fits the laboratory test data. The second, called the *laboratory-scale strength uncertainty*, is in estimating model parameters of a specific distribution from which the minimum failure strength of a material at a certain confidence level is estimated using the laboratory test data. To extrapolate the laboratory-scale strength prediction to that of a full-scale component, a third uncertainty exists called the *full-scale strength uncertainty*. Based on a series of publications in the statistics literature [1-

20], we developed a three-step approach [21] to estimating the minimum strength of a full-scale component using two metrics. The first metric is based on six goodness-of-fit and parameter-estimation-method criteria, while the second metric is based on uncertainty quantification of the so-called A-basis design allowable (99 % coverage at 95 % level of confidence) of the full-scale component. The three steps of our approach are:

(1) Find the "best" model for the sample data from a list of five candidates, namely, normal, two-parameter Weibull, three-parameter Weibull, two-parameter lognormal, and three-parameter lognormal.

(2) For each model, estimate (2a) the parameters of that model with uncertainty using the sample data, and (2b) the minimum strength at the laboratory scale at 95 % level of confidence.

(3) Introduce the concept of *coverage* and estimate the full-scale allowable minimum strength of the component at 95 % level of confidence for two types of coverages commonly used in the aerospace industry, namely, 99 % (A-basis for critical parts) and 90 % (B-basis for less critical parts).

To illustrate the applicability of this uncertainty-based approach to a diverse group of data, we present in [21] results of our analysis, using DATAPLOT [22], for six sets of laboratory failure strength data from four engineering materials, as shown in part in Figure 55, Figure 56, and Figure 57.

[1] N. R. Mann, R. E. Schafer, and N. D. Singpurwalla. *Methods for Statistical Analysis of Reliability and Life Data*. Wiley, New York, 1974.

[2] W. B. Nelson. *Applied Life Data Analysis*. Wiley, New York, 1982.

[3] J. F. Lawless. *Statistical Models and Methods for Lifetime Data*. Wiley, New York, 2nd Ed., 2003.

[4] *MIL-HDBK-17 Composite Materials Handbook Volume 1: Guidelines for Characterization of Structural Materials*. U.S. Department of Defense, Washington, DC, Chapter 8 (Statistical Methods).

[5] J. Aldrich. R. A. Fisher and the Making of Maximum Likelihood 1912–1922. *Statistical Science* **12**:3 1997), 162-176.

[6] D. R. Anderson. Appendix A. Likelihood Theory. In *Model Based Inference in the Life Sciences: A Primer on Evidence*, Springer, New York, 2008, 147–154.

[7] K. V. Bury. *Statistical Models in Applied Science*. Wiley, New York, 1975.

[8] J. J. Filliben. *Simple and Robust Linear Estimation of the Location Parameter of a Symmetric Distribution*. Ph.D. thesis, Princeton University, Princeton, NJ, 1979.

[9] J. J. Filliben. The Probability Plot Correlation Coefficient Test for Normality. *Technometrics* **17**:1 (1975), 111–117.

[10] S. W. Looney and T. R. Gulledge Jr. Use of the Correlation Coefficient with Normal Probability Plots. *The American Statistician* **39**:1 (1985), 75–79.
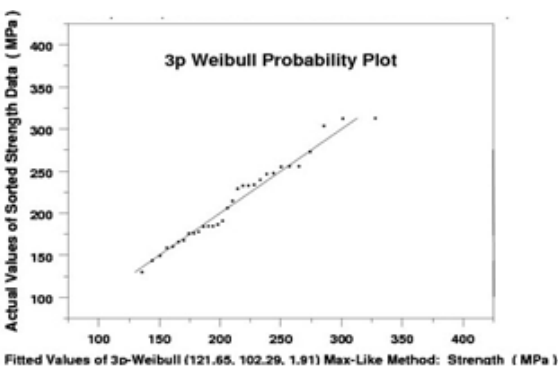
Figure 55. *A three-parameter Weibull plot of a set of 31 biaxial test data for the ultimate tensile strength of a BK-7 glass vs. predicted values.*
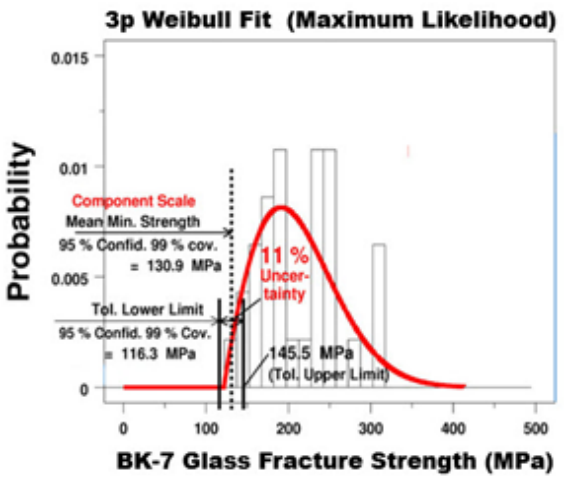
Figure 56. *BK-7 glass: Histogram and fitted 3p Weibull probability density function for the same strength data with lower and upper tolerance limits at 95 % confidence and 99 % coverage (A-basis).*

Figure 57. *Comparison of the A-basis design allowable minimum strength (MPa) selected from the 2pW model (ASTM) approach vs. our approach by making the best choice among five models according to a goodness-of-fit or tolerance limit uncertainty metric.*

[11] R. M. Vogel. The Probability Plot Correlation Coefficient Test for the Normal, Log-Normal, and Gumbel Distributional Hypotheses. *Water Resources Research* **22**:4 (1986), 587–590.

[12] T. W. Anderson and D. A. Darling. Asymptotic Theory of Certain 'Goodness of Fit' Criteria Based on Stochastic Processes. *Annals of Mathematical Statistics* **23** (1952), 193–212.

[13] T. W. Anderson and D. A. Darling. A Test of Goodness of Fit. *Journal of the American Statistical Association* **49** (1954), 765–769.

[14] A. C. Cohen and B. J. Whitten. *Parameter Estimation in Reliability and Life Span Models*. CRC Press, Boca Raton, FL, 1988.

[15] K. Bury. *Statistical Distributions in Engineering*. Cambridge University Press, 1999.

[16] J. I. McCool. *Using the Weibull Distribution: Reliability, Modeling, & Inference*. Wiley, New York, 2012.

[17] P. R. Nelson, M. Coffin, and K. A. F. Copeland. *Introductory Statistics for Engineering Experimentation*. Elsevier Academic Press, London, 2003.

[18] F. Prochan. Confidence and Tolerance Intervals for the Normal Distribution. *Journal of the American Statistical Association* **48** (1953), 550–564.

[19] M. G. Natrella. *Experimental Statistics*. National Bureau of Standards Handbook 91, 1966, pp. 1-14, 1-15, 2-13, 2-14, and 2-15, Tables A-6 and A-7.

[20] H. Rinne. *The Weibull Distribution, A Handbook*. Chapman & Hall/CRC, Taylor & Francis Group, Boca Raton, FL, 2009.

[21] J. T. Fong, N. A. Heckert, J. J. Filliben, P. V. Marcal, and S. W. Freiman. Estimation of a Minimum Allowable Structural Strength Based on Uncertainty in Material Test Data. *Journal of Research of NIST* **126** (2021), 126036. DOI: 10.6028/jres.126.036

[22] J. J. Filliben and N. A. Heckert. *Dataplot: A Statistical Data Analysis Software System*. National Institute of Standards and Technology, Gaithersburg, MD. Available at http://www.itl.nist.gov/div898/software/dataplot.html

[23] ASTM International. *ASTM Designation: C1239-07– Standard practice for reporting uniaxial strength data and estimating Weibull distribution parameters for advanced ceramics. Committee C28 on Advanced Ceramics, reapproved 2008*. ASTM International, West Conshohocken, PA, 2008.

Estimation of a Failure Probability Upper Bound of a Structure Based on Uncertainty in Fracture and Fatigue Test Data

Jeffrey T. Fong
James J. Filliben (NIST ITL)
N. Alan Heckert (NIST ITL)
Stephen Freiman (NIST MML)
Pedro V. Marcal (MPACT Corp)
Robert Rainsberge (XYZ Scientific)
Stephen R. Doctor (PNNL)
Ned Finney (Duke Energy)

The American Society of Mechanical Engineers Boiler and Pressure Vessel Code Committee has recently developed a new Section XI (Nuclear Components Inspection) Division 2 Code [1] entitled *Reliability and Integrity Management* (RIM). RIM incorporates a new concept known as *System-Based Code*, where an integrated approach from design to service inspection is introduced using three new types of statistical quantities: (1) *system reliability index*, or *system reliability target* for any system consisting of structures and components, (2) *structural reliability*, for any structure, and (3) *component reliability* for any component.

In a recent paper [2], we developed a new theory of fatigue and creep rupture for metal alloys at room and elevated temperatures such that the reliability of an uncracked component can be estimated from fatigue and creep rupture test data with simple loading histories. In this paper [5], we extend the theory to include a methodology to estimate the upper bound of failure probability (which is also known as co-reliability or the quantity, 1 - reliability), of a cracked pipe from fatigue crack growth rate test data, probability of detection (POD) data, and nondestructive evaluation (NDE) of initial crack sizing data for simple loading histories.

To illustrate an application of this new modeling approach, we present four numerical case studies using (a) the fatigue failure data of six AISI 4340 steel specimens at room temperature as shown in Figure 58 (after Dowling [3]) for an uncracked steel pipe, and (b) the fatigue crack growth rate data of 17 specimens of 2024-T3 aluminum (after von Euw, Hertzberg, and Roberts [4]) for a cracked aluminum pipe. The four cases are:

(1) Uncracked and uninspected pipe (see Figure 59).

(2) Inspected pipe with a crack-found-location-and-size call (see Figure 60).

(3) Uncracked and inspected pipe with a no-crack-found call and a POD value.

(4) Inspected pipe with a crack-found-location-and-size call and a structural health monitoring (SHM) program.
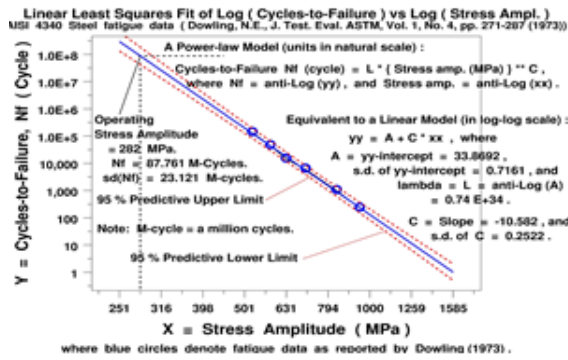
Figure 58. *Linear least squares fit of six fatigue specimen data of AISI 4340 steel at room temperature.*
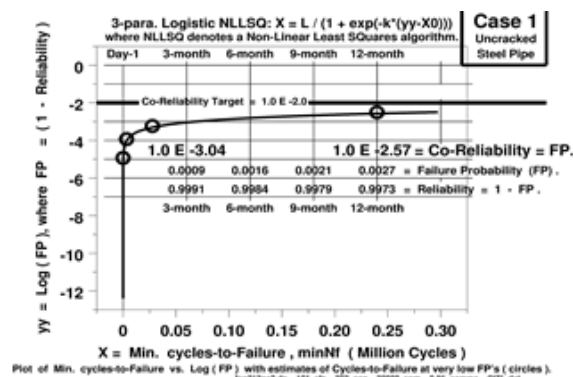
Figure 59. *A co-reliability (failure probability) vs. min. fatigue life plot for Case (1) Uncracked AISI 4340 steel piping at 20 C. Reliability, failure probability, and co-reliability values at 3-mon., 6-mon., 9-mon., & 12-mon. are for an average 20,000 cycles per month of usage and an operational fatigue stress amplitude of 282 MPa at 20 C. A solid line showing a co-reliability target of 1.0 x 10^{-2} is also plotted.*
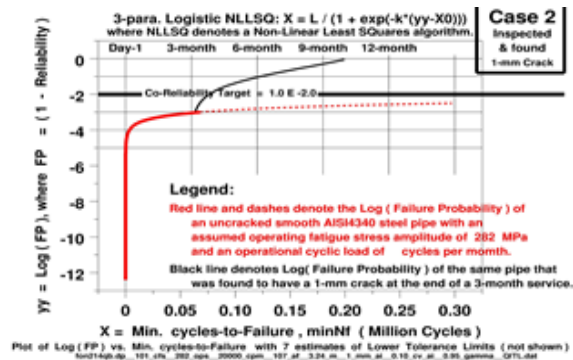
Figure 60. *A co-reliability (failure probability) vs. min. fatigue life plot for Case (2) AISI 4340 steel pipe at 20 C, inspected at the end of a 3-month period, and found to have a "1-mm-crack," assuming an operating stress amplitude of 282 MPa and an average 20,000 cycles per month of usage.*

[1] *ASME Boiler and Pressure Vessel Code, Section XI, Division 2 - Requirements for Reliability and Integrity Management (RIM) Program for Nuclear Power Plants.* Draft for Public Comment. ASME, November 11, 2018.

[2] J. T. Fong, N. A. Heckert, J. J. Filliben, and S. W. Freiman. A Multi-Scale Failure-Probability-based Fatigue or Creep Rupture Life Model for Metal Alloys. *International Journal of PVP* **173** (2019), 79-93. DOI: 10.1016/j.ijpvp.2019.04.003

[3] N. E. Dowling. Fatigue Life and Inelastic Strain Response under Complex Histories for an Alloy Steel. *ASTM Journal of Testing & Evaluation* **1**:4 (1973), 271-287.

[4] E. F. J. von Euw, R. W. Hertzberg, and R. Roberts. Delay Effects in Fatigue-Crack Propagation. In *Stress Analysis and Growth of Cracks, ASTM STP 513*, 1972, 230-259.

[5] J. T. Fong., P. V. Marcal, R. Rainsberger, N. A. Heckert, J. J. Filliben, S. R. Doctor, and N. A. Finney, Jr. A Multi-Scale Failure-Probability-and-NDE-Based Fatigue Life Model for Estimating Component Co-Reliability of Uncracked and Cracked Pipes. In *Proceedings of the 2021 ASME Pressure Vessels & Piping Division Conference*, Online, July 12-16, 2021 . DOI: 10.1115/PVP2021-62169

Extending Zeno

John Nolan
Jack Douglas (NIST MML)
Debra Audus (NIST MML)

https://zeno.nist.gov/

Work on generalized random walks based on multivariate stable laws has been extended. The primary motivation for this is to expand the capabilities of the Zeno program, a NIST package to probe the structure of complex objects, e.g., polymers. Other possible uses are for simulating solutions to fractional diffusion equations, which are used to model fluid movement in porous materials. Our previous focus has been on improvements to the algorithms developed earlier. Speeding up the simulation of hitting distributions for stable processes has been a major focus, because Zeno-type simulations involve millions of iterations. Also, the initial focus was on dimension $d = 3$, but the results are now extended to any dimension. The Walk-On-Spheres method is a method to accelerate simulations in the Brownian case. An extension, called Walk-In-And-Out-Of-Balls, has been developed that takes into account the discontinuous paths of stable motion. The standard estimates have infinite variance, and therefore the Central Limit Theorem does not apply. We are now developing non-standard ways of estimating alpha-capacity and the precision of those estimates.

Quantifying the Relationship Between Adsorption Equilibria and Solute Movement in the Break-through Curve Measurement

Robert DeJaco
Paul Patrone
Anthony Kearsley

In an industrial separation process involving adsorption or chromatography, a fluid mixture is fed into a column packed with immobile solid particles. The solid possesses different affinities for different components of the fluid mixture. As a result, while the mixture flows through the column, different components move along it at different rates and shapes. Eventually, components exit the column at different times, leading to a separated mixture. Since these processes do not require heat, they use less energy than conventional distillation and can separate thermolabile chemicals like biopharmaceuticals.

In recent years, advances in materials science have led to large increases in the space of adsorbent materials that can be accessed experimentally. This has led to the discovery of high-performing materials for a wide variety of separation processes, especially carbon capture. Despite these efforts, industrial implementation of new adsorbent materials is slow. This is in part due to a lack of a quantitative understanding of the nonlinear relationship between physical properties and the movement of solute along the column.

To address this challenge, we are using perturbation theory in the limit of small ϵ, or a fast rate of mass transfer compared to convection. For single-solute adsorption at a rate described by a linear driving force with negligible axial dispersion and thermal variation, our approach provides a quantitative relationship between the solid concentration in equilibrium with a fluid concentration c, or $f(c)$, and the elution profile associated with a step change in feed concentration. The function f is a physical property called the adsorption isotherm. The elution profile, called the break-through curve, is a popular experimental approach used to assess the relationship between physical properties and solute movement. As such, the measurement is essential for performance evaluation and design.

We have found that regular and singular perturbation theory in small ϵ provide a more generic and quantitative description of the conditions under which several theories of solute movement are valid, as well as how they are related. At the same time, our approach has been able to answer several open questions on the relationship between f, ϵ, and the break-through curve measurement.

Local equilibrium theory [1], for example, is so named because the form of the problem is typically attributed to result from the assumption that mass transfer is so fast that the fluid is in thermodynamic equilibrium with the solid at all space and time. This is a somewhat perplexing concept [2], since the solute is moving through the column. Analyzing the limit $\epsilon \to 0$ demonstrates that the local equilibrium concept can be more generically interpreted as the *leading-order behavior*. Dispersive waves (also called rarefaction waves) in local equilibrium theory capture such leading-order behavior. However, shock waves and linear waves are approximated by local equilibrium theory *at the point $\epsilon = 0$*. Since such a condition cannot be achieved experimentally, boundary layer theory can be used to obtain the physically appropriate behavior of shock waves and linear waves for small ϵ.

The means by which shock waves approach a constant-pattern wave (also called a traveling wave) is more quantitatively understood with boundary layer theory, as the former approach immediately neglects all dependence on time. A very long column is not required for constant-pattern behavior to occur, as it is the leading order approximation as $\epsilon \to 0$. Investigation of the dominant behavior also reveals the time scales associated with the constant-pattern wave. A constant-pattern wave is dominant at elution times when $\Delta f / \Delta c$, the change in equilibrium solid concentration divided by fluid concentration across the wave, is neither large nor small. A constant-pattern wave requires that the driving force for adsorption be positive in adsorption mode, but negative in desorption mode. Within the boundary layer, the leading order behavior *cannot* be local equilibrium.

In contrast to the boundary layer associated with a shock wave, the linear boundary layer has received much less attention. In this case, the profile spreads out with time and the width is proportional to $\sqrt{\epsilon}$, a factor larger than that associated with the former. As the adsorption strength increases, the boundary layer decreases in width. The boundary layer approximation is asymptotic to the analytical solution when the boundary layer is sufficiently behind the wave associated with the non-adsorbing solvent.

This work provides a more quantitative understanding of the relationship between physical properties and solute movement, the primary mechanism for separation in adsorption and chromatography. The boundary layer perspective sets the stage for extension to more complex systems (and other separation processes). Such efforts facilitate improved process designs and the selection of high-performing materials from databases of available candidates. We are currently preparing our results for dissemination in peer reviewed journals. In the future, we plan to extend the approach to adsorption isotherms that are weakly nonlinear and to short time behavior.

[1] H. K. Rhee, R. Aris, and N. R. Amundson. *First-Order Partial Differential Equations: Volume 1, Theory and Application of Single Equations*. Prentice-Hall, Inc., 1986.

[2] D. Tondeur. Paradigms and Paradoxes in Modeling Adsorption and Chromatographic Separations. *Industrial and Engineering Chemistry Research* **34** (1995), 2782—2788. DOI:10.1021/ie00047a029

High Performance Computing and Visualization

Computational capability continues to advance rapidly, enabling modeling and simulation to be done with greatly increased fidelity. Doing so often requires computing resources well beyond what is available on the desktop. Developing software that makes effective use of such high-performance computing platforms remains very challenging, requiring expertise that application scientists rarely have. We maintain such expertise for application to NIST problems. Such computations, as well as modern experiments, typically produce large volumes of data, which cannot be readily comprehended. We are developing the infrastructure necessary for advanced interactive, quantitative visualization and analysis of scientific data, including the use of 3D immersive environments, and applying the resulting tools to NIST problems

High Precision Calculations of Fundamental Properties of Few-Electron Atomic Systems

James Sims
Maria Ruiz (University of Erlangen, Germany)
Bholanath Padhy (Khallikote College, India)

NIST has long been involved in supplying critically evaluated data on atomic and molecular properties such as the atomic properties of the elements contained in the Periodic Table and vibrational and electronic energy level data for neutral and ionic molecules contained in the NIST Chemistry WebBook. Fundamental to this work is the ability to predict, theoretically, a property more precisely than even the most precise experiments. It is our goal to be able to accomplish this for few-electron atomic systems.

While impressive advances have been made over the years in the study of atomic structure in both experiment and theory, the scarcity of information on atomic energy levels is acute, especially for highly ionized atoms. The availability of high precision results tails off as the state of ionization increases, not to mention higher angular momentum states. In addition, atomic anions have more diffuse electronic distributions, representing more challenging computational targets than the corresponding ground states.

In the past two decades, there has been breathtaking improvements in computer hardware and innovations in mathematical formulations and algorithms, leading to "virtual experiments" becoming a more and more cost effective and reliable way to investigate chemical and physical phenomena. Our contribution in this arena has been undertaking the theoretical development of our hybrid Hylleraas-CI (Hy-CI) wave function method to bring sub-chemical precision to atomic systems with more than two electrons.

Hy-CI has from its inception been an attempt to extend the success of the Hylleraas (Hy) method to systems with more than three electrons, and hence is an attempt to solve not just the three-body problem but the more general N-body problem [1]. Fundamental to the method is the restriction of one r_{ij} per configuration state function (CSF). (For atomic systems with greater than four electrons, all relatively precise calculations nowadays adopt the Hy-CI methodology of one r_{ij} term per CSF). In the case of three electron lithium systems, we have computed four excited states of the lithium atom to two orders of magnitude greater than has ever been done before [2]. At the four-electron level, to get truly precise chemical properties like familiar chemical electron affinities and ionization energies, it is important to get close to the nanohartree level we achieved for the three-electron atom, a significantly more difficult problem for four electrons than for three. By investigating more flexible atomic orbital basis sets and better configuration state function filtering techniques to control expansion lengths, we have been able to successfully tackle the four-electron case.

Progress to date has included computing the non-relativistic ground state energy of not only beryllium, but also many members of its isoelectronic sequence to eight significant digit precision. With the results from our calculations and a least-squares fit of the calculated energies, we have been able to compute the entire beryllium ground state isoelectronic sequence for Z = 4 through Z = 113 [3]. Li⁻ (with Z=3), nominally the first member of this series, has a decidedly different electronic structure and was not included in those calculations and subsequent discussions, but that omission has been corrected and we have subsequently carried out a large, comparable calculation for the Li⁻ ground state [4].

The first member of the Be isoelectronic ground state sequence, the negative Li⁻ ion, is also a four-electron system in which correlation plays a very important part in the binding. However due to the reduced nuclear charge, it is a more diffuse system in which one of its outer two L shell electrons moves at a greater distance from the nucleus than the other, and hence its nodal structure is different from that of a coupled L shell with an identical pair of electrons. The ground state of the singlet S state of Li⁻ is the same type of problem as the first excited state of Be; it is like Be(2s3s), not Be(2s2s). Completing this calculation has provided the necessary insight to enable the calculation of the Be first excited

state of singlet S symmetry, Be(2s3s), to an order of magnitude better than previous calculations. Armed with this result, we have been able to continue this level of precision to the Be(2s4s) excited state and have calculated the higher, more diffuse Be(2s5s) through Be(2s7s) states as well, and in the process have demonstrated that Hy-CI can calculate the higher, more diffuse Rydberg states with more complicated nodal structures to the same level of precision as less excited states [5].

While our work has demonstrated the efficacy of Hy-CI as a solution to the N-body problem for four or more electrons, this work has also shown the presence of a "double cusp" $r_{12}r_{34}$ term type slow convergence problem at the nanohartree precision level which is ultimately built into Hy-CI for four or more electrons. We have investigated a generalization of the Hy-CI method to an exponentially correlated Hy-CI (E-Hy-CI) method in which the single r_{ij} of an Hy-CI wave function is generalized to a form which pairs an exponential r_{ij} factor with linear r_{ij}, producing a correlation factor which has the right behavior in the vicinity of the r_{ij} cusp, and also as r_{ij} goes to infinity. While this was proposed in 2012 and there have been several papers on E-Hy-CI integrals, there were no computational tests until our calculations. Not only has the E-Hy part (the part that differs from conventional Hy-CI) been tested, but E-Hy-CI calculations have been done for spherically symmetrical and non-symmetrical orbitals as well.

The purpose of this research has been to determine how effective exponential correlation factors can be. By comparing convergence of the E-Hy-CI wave function expansion to that of the Hy-CI wave function without exponential factors, both convergence acceleration and an improvement in the precision for the same basis are demonstrated. This makes the application of the E-Hy-CI method to systems with $N > 4$, for which this formalism with at most a single exponentially correlated and linear r_{ij} factor leads to solvable integrals, very promising. The ground 1 singlet S state non-relativistic energy of He is computed to be -2.9037 2437 7034 1195 9831 1084 hartrees (Ha) for the best expansion.

We followed the success on the ground state of the He atom with calculations on the ground 1 singlet S and the 2 singlet S through 6 singlet S excited S states of the Li$^+$ ion with the same technique, with results comparable to the He atom. This demonstrates the utility of the E-Hy-CI approach for not only ground but also excited states of S symmetry as well [7].

Work is presently in progress to compare the convergence of the E-Hy-CI wave function expansion to that of the Hy-CI wave function without exponential factors for states of non-S symmetry, specifically the 1s2p 2 singlet P state of the He atom and members of its isoelectronic sequence. The computed P state wave functions will be used with the previously obtained S state wave function (and some new ones for Be II, C IV

and O VI) to calculate oscillator strengths, including rigorous quantum mechanical upper and lower bounds, for the lowest singlet S to singlet P transition.

[1] J. S. Sims and S. A. Hagstrom. Combined Configuration Interaction – Hylleraas Type Wave Function Study of the Ground State of the Beryllium Atom. *Physical Review A* **4**:3 (1971), 908. DOI: 10.1103/PhysRevA.4.908

[2] J. S. Sims and S. A. Hagstrom. Hy-CI Study of the 2 ^2S Ground State of Neutral Lithium and the First Five Excited ^2S States. *Physical Review A* **80** (2009), 052507. DOI: 10.1103/PhysRevA.80.052507

[3] J. S. Sims and S. A. Hagstrom. Hylleraas-Configuration Interaction Nonrelativistic Energies for the Singlet S Ground States of the Beryllium Isoelectronic Series up Through Z = 113. *Journal of Chemical Physics* **140** (2014), 224312. DOI: 10.1063/1.4881639

[4] J. S. Sims. Hylleraas-Configuration Interaction Study of the Singlet S Ground State of the Negative Li Ion. *Journal of Physics B: Atomic, Molecular and Optical Physics* **50** (2017), 245003. DOI: 10.1088/1361-6455/aa961e

[5] J. Sims. Hylleraas-Configuration Interaction (Hy-CI) Non-Relativistic Energies for the 3 ^{1}S, 4 ^{1}S, 5 ^{1}S, 6 ^{1}S, and 7 ^{1}S Excited States of the Beryllium Atom. *Journal of Research of the NIST* **125** (2020), 125006. DOI: 10.6028/jres.125.006

[6] J. S. Sims, B. Padhy, and M. B. Ruiz. Exponentially Correlated Hylleraas-Configuration Interaction (E-Hy-CI) Non-Relativistic Energy of the ^{1}S Ground State of the Helium Atom. *International Journal of Quantum Chemistry* **121**:4 (2020), e26470. DOI: 10.1002/qua.26470

[7] J. S. Sims, B. Padhy, and M. B. Ruiz. Exponentially Correlated Hylleraas-Configuration Interaction (E-Hy-CI) Studies of Atomic Systems. II. Non-relativistic Energies of the 1 ^{1}S through 6 ^{1}S States of the Li$^+$ Ion. *International Journal of Quantum Chemistry* **122** (2021), e26823. DOI: 10.1002/qua.26823

Simulation of Dense Suspensions: Cementitious Materials

William George
Nicos Martys (NIST EL)
Steven Satterfield
Judith Terrill

A suspension is a collection of solid inclusions embedded in a fluid matrix. Suspensions play an important role in a wide variety of applications including paints, cement-based materials, slurries, and drilling fluids. Understanding the flow properties of a suspension is necessary in many such applications. However, measuring and predicting flow properties of suspensions remains challenging.

Suspensions can be quite complex, as the inclusions may have a wide range of shapes and a broad size distribution. Further complicating matters is that different matrix fluids may have quite disparate flow behavior. While the simplest type of matrix fluid is Newtonian, where the local stress is proportional to the shear rate, the matrix fluid can also be non-Newtonian, exhibiting quite complex behavior including shear thinning (viscosity decreases with shear rate), shear thickening (viscosity increases with shear rate), viscoelasticity (exhibiting both viscous and elastic properties), or even have a time dependent viscosity (thixotropic). We have two on-going studies on the rheology of cementitious materials, which are dense suspensions in non-Newtonian matrix fluids.

SRMs for Mortar and Concrete. Rotational rheometers, devices that measure fluid properties such as viscosity, are routinely used for homogeneous materials such as oils, but their use on dense suspensions, such as concrete, is relatively new. Since measurements with rheometers involve flow in a complex geometry, it is important that they are calibrated with a well characterized standard reference material (SRM). We are developing such SRMs in collaboration with NIST EL.

NIST produced an SRM for cement paste (SRM 2492) as the first step in the development of a reference material for concrete rheometers. The second step, the development of an SRM for mortar (SRM 2493) was completed in 2017 and is currently available. The material properties of the mortar SRM, such as viscosity, could not be measured in fundamental units with certainty. Thus, simulation was used to determine the viscosity of the mortar. To obtain the necessary fidelity in the simulations, computations at a high-performance computing facility were necessary. Results of these simulations were compared with physical experiments as validation.

In 2019 we released SRM 2497, a standard reference concrete for rheological measurements [1]. The concrete SRM is comprised of the previously released mortar SRM with the addition of suspended 10 mm diameter hard glass spheres. The certified values for SRM 2497 were determined using the simulation results previously computed for the mortar SRM. As an indication of the impact of this work, the use of the cement paste and mortar SRMs are referenced in the new ASTM standard test method for measuring the rheological properties of cementitious materials [2]. A more detailed description of the development and validation of these SRMs has recently been published [3].

We are currently running a suite of simulations of the concrete SRM using the cement paste as the matrix fluid in which both 1 mm and 10 mm diameter hard spheres are suspended. The results of these simulations should match the results from simulations comprising the mortar SRM as the matrix fluid was with suspended 10 mm hard spheres. Depending on the outcome of these
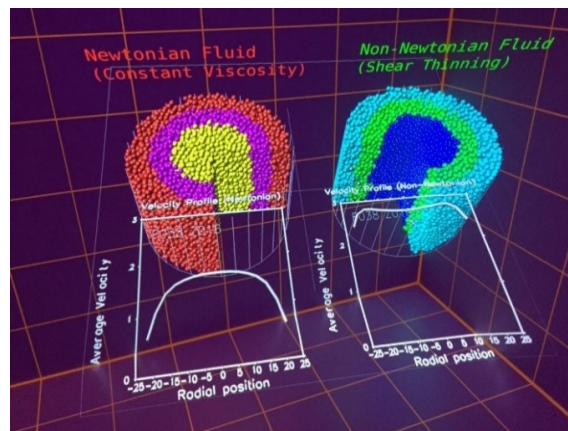
Figure 61. *Side-by-side comparison of identical systems, in the NIST immersive visualization system, with only a single property of the matrix fluid changed, one Newtonian and the other non-Newtonian (shear thinning).*

simulations, the concrete SRM certification may need to be updated. These simulations have not yet reached an equilibrium state and so we continue running them.

Flow of Dense Suspensions in Pipes. Understanding the flow of suspensions in pipes is important for a wide variety of applications. For example, in the construction industry, concrete is often placed by pumping it through extensive pipe systems. However, research on predicting the pumpability of concrete has been limited due to the heavy equipment and large amounts of material needed. Suspension flow in pipes is also important in the developing field of 3D additive manufacturing.

Predicting the flow of this complex fluid, which is composed of a non-Newtonian matrix fluid with suspended solid inclusions flowing under pressure, is challenging. Flow in these systems is also complicated by variety of phenomena such as slip at the wall and shear induced migration, which has only been studied for the simpler case of a suspensions with a Newtonian matrix fluid. A detailed discussion of this topic is available in a NIST Technical Note [4]. It is also the case, especially in the case of 3D additive manufacturing, that the placement of these materials is time sensitive, from the time the material is initially mixed to the time it is pumped and placed [5].

We have been conducting detailed simulations of the flow of suspensions through pipes to enable the development of predictive flow models and to advance measurement science in this field. Through quantitative analysis and visualization of results, we have gained insight into shear-induced migration and slip behavior in these systems. For example, Figure 61 shows a side-by-side comparison of identical systems with only a single property of the matrix fluid changed, one Newtonian and the other non-Newtonian (shear thinning). Over the last year we have been conducting a suite of simulations of shear thinning and shear thickening suspensions flowing

in pipes, varying the properties of the matrix fluid and the driving force.

Studying the flow velocity fields as a function of driving force we have discovered a useful scaling relationship. Given that the matrix fluids have a viscosity, μ, that relates to the strain rate $\dot{\gamma}$ such that

$$\mu \sim \dot{\gamma}^n,$$

we have determined that the system velocity in the pipe, v, is related to the driving force, g, as

$$v \sim g^{\frac{1}{1+n}}$$

So, for example, with a shear thinning matrix fluid with $n = -0.5$, we have

$$v \sim g^{\frac{1}{1-0.5}} = g^2$$

and with a matrix fluid which is shear thickening, with $n = 0.5$, we have

$$v \sim g^{\frac{1}{1+0.5}} = g^{\frac{2}{3}}$$

Notice that this scaling relation depends on the power-law behavior of the non-Newtonian matrix fluid. As a consequence, given a few measurements of the flow velocity versus the driving force of the suspension, one can determine the power law behavior of the suspension, and indeed we can then also determine the power law behavior of the matrix fluid.

In the last year, all of our pipe-flow simulations began to reach steady state, although the simulation of shear thickening mortars in pipes has proven to be the slowest to approach steady state and so we will continue to run those simulations. Our results, which show agreement with the model predictions to within 15 %, were published in 2020 in the Journal of Rheology [6].

In support of NIST researchers investigating 3D additive manufacturing using cementitious materials we have begun developing a simulation of the flow through a pipe and then through an "active mixer" device. This device is being designed to mix additives to a cement paste or mortar at the point of placement, that is, immediately before the material is placed. See Figure 62. The additives to be mixed are injected through the openings in the shafts on the right side (input side) of the rotor.

Nanoscale Suspension Flow. New instrumentation at NIST's NCNR allows for the study of flow in nanoscale parallelepiped channels using neutron scattering. In support of this study, we are extending the capability of the QDPD dense-suspension code to model similar flows, matching interparticle interactions and flow rates, for comparison to these new scattering experiments. This work will also be extended to include suspensions of monoclonal antibodies and other extended or non-spherical objects in the flow geometry [7].

[1] Nicos S. Martys, William L. George, Ryan P. Murphy, Katheen Weigandt. Certification of SRM 2497: Standard Reference Concrete for Rheological Measurements. NIST Special Publication 260-194, April 2019, 116 pages. DOI: 10.6028/NIST.SP.260-194

[2] ASTM International C1874-19. New Test Method for Standard Test Method for Measuring the Rheological Properties of Cementitious Materials using a Coaxial Rotational Rheometer. ASTM International, West Conshohocken, PA, 2019.

[3] N.S. Martys, W.L. George, S.G. Satterfield, B. Toman, M.A. Peltz, S.Z. Jones, and C.F. Ferraris. Standard Reference Materials for Rheological Measurements of Cement-Based Materials. *ACI Materials Journal* **118**:6 (November 2021). DOI: 10.14359/51733132

[4] M. Choi, C. F. Ferraris, N. S. Martys, V. K. Bui, H. R. Trey Hamilton, and D. Lootens. Research Needs to Advance Concrete Pumping Technology. NIST Technical Note 1866, 2015. DOI: 10.6028/NIST.TN.1866

[5] S. Z. Jones, D. P. Bentz, N. S. Martys, W. L. George, and Thomas. Rheological Control of 3D Printed Cement Paste. In *Digital Concrete 2018: First International Conference on Concrete and Digital Fabrication*, Zurich Switzerland, September 9-12, 2018. DOI: 10.1007/978-3-319-99519-9_7

[6] N. S. Martys, W. L. George, R. P. Murphy, and K. Weigandt. Pipe Flow of Sphere Suspensions Having a Power-Law-Dependent Fluid Matrix. *Journal of Rheology* **64**:2 (March/April 2020). DOI: 10.1122/1.5131021

[7] Y. Zhai, N. S. Martys, W. L. George, J. Nayem, and Y. Liu. Intermediate Scattering Functions of a Rigid Body Monoclonal Antibody Protein in Solution Studied by Dissipative Particle Dynamic Simulation. *Structural Dynamics* **8** (March 2021). DOI: 10.1063/4.0000086
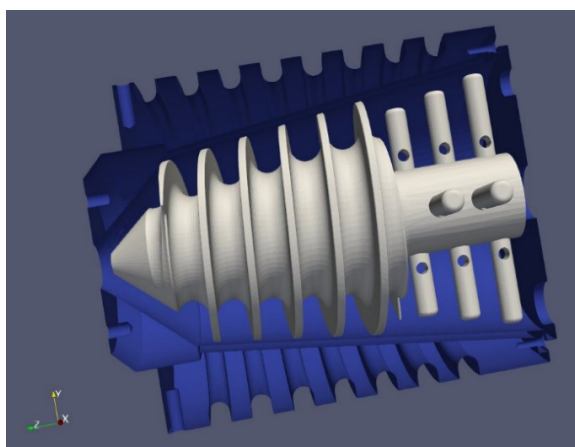
Figure 62. *NIST EL's proposed active mixer for additive manufacturing with cementitious material. The rotor is shown in gray and the housing, which is clipped in half to show the rotor inside, is shown in blue.*

Monoclonal Antibodies Under High Shear

William George
Nicos Martys (NIST EL)
Steven Satterfield
Judith Terrill

Recent advances at the NIST Center for Neutron Research (NCNR) have enabled neutron scattering experiments for the study of the development of long-range structure and phase transitions under pressure driven flow conditions in a tube geometry. This capability was developed, in part, to study therapeutic proteins, such as monoclonal antibodies, as they flow under high shear rate such as would be encountered during direct injection via a hypodermic needle.

Such direct injection is desired because it would simplify the delivery of these therapeutic proteins, which are currently only administered intravenously. However, since such injections can lead to shear rates on the order of 10^6/s there is concern that this could lead to structural degradation of the molecules. Even small distortions of the molecular structure could cause aggregation of the protein. The most dangerous potential consequence of such distortion and agglomeration is that the patient may develop an allergic reaction to the medication. Recorded fatalities have occurred where immunogenic response to such aggregates is suspected. For these reasons it is necessary for the pharmaceutical companies to demonstrate that the structure of protein therapeutics is not degrading before high shear rate delivery methods can be utilized.

In support of this research, we have been extending the capabilities of our dense-suspension simulator, QDPD, to match the physical experiments being performed at the NCNR. Recent physical experiments are in good agreement with preliminary QDPD simulations of similar systems, that is, the flow of suspensions in a tube. The extensions to QDPD will enable the simulation of suspensions of the NIST monoclonal antibody (NISTmAb) Reference Material (SRM 8671) shown in Figure 63. In 2021 we implemented both a spring force and a bending force between the three lobes of the NIST SRM mAb. A third force, torsion, which controls the twisting of the lobes has been implemented this year.

In collaboration with NCNR, NIST EL, and NIST MML, we are now ready to carry out a suite of simulations designed to determine the effect of the flexibility within monoclonal antibodies on the suspension's viscosity for a set of volume fractions and shear rates like that studied at NCNR. Figure 64 shows a snapshot of a simulation of the NIST mAb flowing in a suspension.

In addition to the rheological studies, we will also modify QDPD to generate pertinent data from our NIST mAb simulations that will assist in the analysis of neutron scattering data from NCNR experiments on such
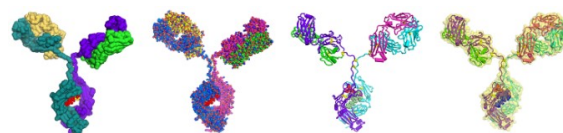
Figure 63. *Several depictions of the NIST monoclonal antibody (NIST mAb) Standard Reference Material (SRM 8671).*
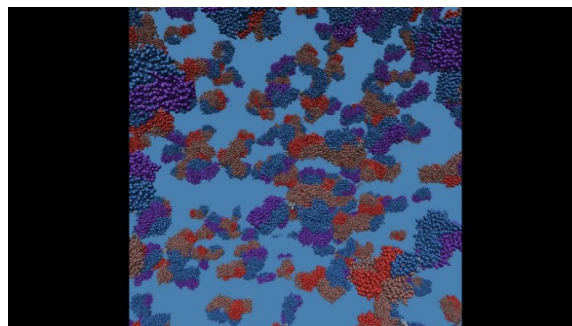
Figure 64. *A snapshot of a simulation of the NIST mAb in suspension. The individual spheres are the amino acids that make-up each mAb.*

systems. This is needed since experimental neutron scattering results so far have been shown to be difficult to interpret, primarily due to their complexity.

Early application of this new simulation capability, in collaboration with researchers at the University of Illinois and the University of Delaware, treating the NIST mAb as a rigid protein, has been published in the journal Structural Dynamics [1].

[1] Y. Zhai, N. S. Martys, W. L. George, J. Nayem, and Y. Liu. Intermediate Scattering Functions of a Rigid Body Monoclonal Antibody Protein in Solution Studied by Dissipative Particle Dynamic Simulation. *Structural Dynamics* **8** (March 2021). DOI: 10.1063/4.0000086

HydratiCA, In Situ Analysis, and Machine Learning

Judith Terrill
Justyna Zwolak
James Filliben (NIST ITL)
Jeffrey Bullard (Texas A&M University)

When large-scale computational simulations have extremely long running times, it is not possible to write out the data during runs. But there are things that need to be studied in the simulations. The only way forward is to do the needed analysis during the runs. The advantage is that we get to do the needed analyses, but the disadvantage is that the analyses need to be done locally in each cell of the simulation without outside interaction.

We use a large-scale simulation code, HydratiCA [1], to develop and test our techniques. HydratiCA is a

stochastic reaction-diffusion model of cement hydration. Hydration transforms paste from a fluid suspension (e.g., cement powder mixed with water) into a hardened solid. This process involves complex chemical and microstructural changes. Understanding and predicting the rates of these changes is a longstanding goal

HydratiCA integrates a partial differential equation forward in time using small time steps on a 3D lattice of sites, and data analysis is challenging because a single timestep generates a large, many-dimensional data set. Furthermore, there are large parts of that data that are either empty or not interesting to the scientists. The original goal was to try to understand and predict the conditions when a lattice site either runs out of material or space and causes the simulation to greatly reduce the time step, and substantially increase the simulation time. To address this, we first searched the simulation for a good descriptor (something which can be used to characterize something else: a dimension of representation space). We discovered that reaction saturation index [2, 3], a quantitative measure of how far a reaction is from equilibrium and also a measure of the thermodynamic driving force for growth, provided valuable information on the state of the simulation at each site. We successfully modeled the saturation index vs. simulation time with Chebyshev polynomials and achieved a dramatic reduction in data because more than a million time steps could easily be reduced to 20 coefficients. This has the additional benefit of smoothing the data which enabled us to calculate reliable first and second derivatives. Now the simulation data could be reduced to a few states. It was then straightforward to find the peaks in the curve that indicated when and where the simulation was running short of resources [4]. The peaks in the saturation index have also been seen in some laboratory experiments [7] and still are not well understood, opening up the possibility of designing laboratory experiments and corresponding simulation runs to do direct comparisons between the experiments and simulation outputs.

Evaluating the Quality of the Fits. Curve fitting goes on in each cell of the simulation. But the goal for realistic runs requires around 1000^3 cells. Because of the size, it is not possible to look at this data. So, the fits need to be done locally at each cell and without outside intervention. And this means the quality of these fits needs to be evaluated locally. There are recognized ways of doing this. But as Matejka and Fitzmaurice [5, 6] and others have demonstrated, single number statistics are not robust enough to be relied upon. The solution is to use images of plots and have a trained classifier "look" at them and report a classification. This year we started assembling a dataset rom data accumulated from decades of experiments at NIST. We have created code to analyze each of the datasets and convert them into images. We have started building and testing a deep

learning classifier to provide us with information on the quality of fits. In the coming year, we will continue the deep learning work. In addition, we will work on updating the HydratiCA parallel code.

[1] J. W. Bullard, E. Enjolras, W. L. George, S. G. Satterfield, and J. E. Terrill. A Parallel Reaction-Transport Model Applied to Cement Hydration and Microstructure Development. *Modelling and Simulation in Materials Science and Engineering* **18**:2 (March 2010), 18.

[2] J. W. Bullard, G. W. Scherer, and J. J. Thomas. Time Dependent Driving Forces and the Kinetics of Tricalcium Silicate Hydration. *Cement and Concrete Research* **74** (2015), 26–34.

[3] A. E. Nielsen and J. M. Toft. Electrolyte Growth in Crystals. *Journal of Crystal Growth* **67** (1984), 278-288.

[4] J. E. Terrill. "Curve Fitting with Validation as a First Step to the Discovery of Physical Laws." Workshop III: Validation, Guarantees in Learning Physical Models: from Patterns to Governing Equations to Laws of Nature, Institute for Pure and Applied Mathematics, UCLA, Los Angeles, CA, October 28, 2019.

[5] A. Cairo. Download the Datasaurus: Never Trust Summary Statistics Alone; Always Visualize Your Data. URL: http://www.thefunctionalart.com/2016/08/download-datasaurus-never-trust-summary.html

[6] J. Matejka and G. Fitzmaurice. Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing. In *Proceedings of the 2017 ACM SIGCHI Conference on Human Factors in Computing Systems*, Denver, CO, May 6-11, 2017.

[7] J. M. Makar, G. W. Chan, and Y. Esseghaier. A Peak in the Hydration Reaction at the End of the Cement Induction Period. *Journal of Material Science* **42** (2007), 1388–1392.

Visualization of Greenhouse Gas Emissions

William Sherman
Simon Su
Judith Terrill
James Whetstone (NIST SPO)
Israel Lopez Coto (NIST SPO)
Anna Karion (NIST SPO)
Kimberly Mueller (NIST SPO)

The NIST Greenhouse Gas (GHG) Measurements Program[14] develops advanced tools and standards for accurately measuring GHG emissions so industries and governments will have the information they need to manage emissions effectively. ACMD's HPCVG is collaborating with James Whetstone, Leader of NIST's Greenhouse Gas Measurements Program and his team

---

[14] https://www.nist.gov/greenhouse-gas-measurements

of climate and weather simulation researchers to produce interactive visualizations of their data.

As this collaboration involves the use of the ParaView open-source scientific visualization tool (see the project summary Transition to Open Source Visualization Software on page 93), we have been exploring how to best integrate the data from the Weather Research and Forecasting (WRF) format into ParaView. Previously we wrote a data translator to serve as a bridge from WRF to ParaView, and while that achieves the goal at one level, it produces duplicate data, and consumes the time of the simulation researcher in converting data.

To further smooth the operation from WRF simulation to ParaView, we have begun studying the best methods for making use of ParaView's client/server model to offload data processing to one of NIST's HPC clusters while performing the visualization on a local desktop, laptop, or even the NIST Immersive Visualization Environment.

As detailed below, ACMD's HPCVG is working to enhance the immersive capabilities of the ParaView visualization tool, including a plugin extension that enables it to integrate directly into the NIST Immersive Visualization Environment. In addition, ParaView provides an interface to consumer head-mounted display (HMD) extended reality systems, enabling another style of immersive interaction with the data. Our ongoing work is to bring these two efforts together, making the transition from large-scale systems best for group interactions to the more personal HMD systems. In the coming year we will add more capability to our greenhouse gas emissions visualizations, and study them in the CAVE, including the addition of direct readers for the WRF format to better integrate data directly into ParaView.

Figure 65. *Overview of the framework. The colored arrows link the preparation, validation, and application phases of the framework. The red path represents the preparation and implementation of the physics-based- approximation model while the blue path represents the ML modules. Adapted from [4].*

Combining Machine Learning with Physics: Enhanced Dark Soliton Detection in BECs

Justyna P. Zwolak
Ian B. Spielman (NIST PML)
Sophia M. Koh (Amherst College)
Shangjie Guo (University of Maryland)
Amilson R. Fritsch (University of Maryland)

Machine-learning (ML)-based image classification has many applications in science, from particle physics data analysis, dark matter searching, quantum state preparation to material property prediction. In atomic physics, ML has been used to locate topological phase transitions in images of atomic density and to characterize particles in disordered fields. In our work, we introduce and demonstrate a hybrid framework, shown in Figure 65, that integrates ML techniques with a science-driven analysis to detect, classify, and assess quality of features in experimental data from many-body atomic physics.

Using cold-atom Bose-Einstein condensates (BECs), we focus on solitonic excitations, robust solitary waves that retain their size, shape, and speed as they travel. These properties arise from an interplay between nonlinearity and dispersion that is present in many physical systems. Since their first observation in water channels, solitons have been found in rivers and seas, BECs, optical fibers, astronomical plasmas, and even human blood vesicles. Due to their inherent stability, solitons in optical fibers have found commercial applications in long-distance, high-speed transmission lines.

While the natural environment does not allow for the controlled study of quantum solitons, BECs are an ideal medium where individual or multiple solitons can be created on-demand, with all their properties, such as position and velocity, tuned according to necessity [1, 2]. Most measurements in BEC experiments produce raw data in the form of images that, in our context, provide information about the excitations' positions within the BEC. The challenge is to identify the number, type, and location of excitations efficiently and reliably. Prior to this work, such information was obtained manually [1], inhibiting the automated analysis of large datasets, which is crucial for soliton dynamics studies.

We previously developed a solitonic excitation detection and positioning system that takes image data as input and determines whether a single soliton is present, and, if so, its location [3]. This algorithm is comprised of a data preprocessor that converts raw data into a ConvNet-compatible format, a ConvNet image classifier that determines if a single soliton is detected, and a traditional least-squares fitting position regressor that locates the soliton within the BEC, when applicable. However, given that the fitting techniques can locate solitons only if the soliton number is known in advance and the ConvNet classifier requires a significant amount of data per expected class for training, the utility of our

soliton detection and positioning system left room for improvement.

Our new hybrid framework not only helps to overcome limitations of our initial algorithm but also goes beyond just recognizing patterns in data. In addition to a ConvNet classifier used previously, our framework employs an ML object detector (OD) to automatically localize the features on interest (in our case, all solitonic excitations within BEC). Complementary to the ML tools, the framework also incorporates a physics-based module that enables further, science-driven analysis of the data. For our implementation, we developed a physics-informed excitation (PIE) classifier that provides a fine-grained classification of individual solitonic excitations into physically motivated categories (e.g., longitudinal solitons or solitonic vortices). Finally, to further assess the quality of the excitations, we also introduce a quality estimator which yields the likelihood that a fit to the one-dimensional (1D) profile of a given excitation has parameters in the range expected for a solitonic excitation.

Our implementation of this framework, SolDet, targets the identification, classification, and tracking of solitonic excitations in image data generated by cold atom experiments. Figure 66 depicts a representative sample of the type we are interested in. As a practical test case, we concentrate on identifying dark solitons, which are spatially localized excitations that appear as vertically aligned atomic density depletions in BECs, see Figure 66(b, c). Deep depletions are caused by kink solitons (and solitonic vortices viewed from the side), as in Figure 66(b-i), whereas shallow and asymmetric depletions can be caused by solitonic vortices, as in Figure 66(b-iii). Our framework is the first to automatically differentiate between these instances and locate all solitonic excitations in each image. Importantly, neither of the ML modules require data labeled with the physically motivated sub-categories, which significantly lessens the burden of manual labeling. In fact, the OD is trained using only the no excitation and lone excitation data.

As part of this research effort, we established and curated a dataset of over 16,000 experimental images of BECs with and without solitonic excitations [5, 6]. The dataset consists of images manually labeled into the three pre-defined categories (i.e., no excitations, lone

excitation, other excitations; 33 % of the data) and unlabeled data. The lone excitation class is in addition tagged with the excitation position, PIE class, and quality score. The remaining 67 % of the data is automatically labeled using the SolDet package. This dataset is available [5] to provide an opportunity for the data science community to develop more sophisticated analysis tools and to further understand nonlinear many-body physics.

Our high-level framework—combining ML methods with physics-based analysis—is an integrated platform for studying experimental data. As we show, complementing ML modules with science-based methodologies can minimize data requirements while also improving output reliability. In addition, the PIE classifier and the accompanying quality metric can serve as a prototype for discovery with general cold-atom data. An article reporting this work has recently been published online [4].

[1] A. R. Fritsch, Mingwu Lu, G. H. Reid, A. M. Piñeiro, and I. B. Spielman. Creating solitons with controllable and near-zero velocity in Bose-Einstein condensates. *Physical Review A*, **101** (2020), 053629.

[2] L. M. Aycock, H. M. Hurst, D. K. Efimkin, D. Genkina, H-I Lu, V. M. Galitski, and I. B. Spielman. Soliton diffusion in Bose–Einstein condensates. *Proceedings of the National Academy of Sciences*, **114** (2017), 2503-2508.

[3] S. Guo, A. R. Fritsch, C. Greenberg, I. B. Spielman, and J. P. Zwolak. Machine-learning Enhanced Dark Soliton Detection in Bose-Einstein Condensates. *Machine Learning: Science and Technology* **2** (2021), 035020.

[4] S. Guo, S. M. Koh, A. R. Fritsch, I. B. Spielman, and J. P. Zwolak. Combining Machine Learning with Physics: A Framework for Tracking and Sorting Multiple Dark Solitons. arXiv: 2111.04881, 2021.

[5] National Institute of Standards and Technology. Dark solitons in BECs dataset (2022). Database: data.nist.gov [Internet]. DOI: 10.18434/mds2-2363

[6] A. R. Fritsch, S. Guo, S. M. Koh, I. B. Spielman, and J. P. Zwolak. Dark Solitons in Bose-Einstein Condensation: A Dataset for Many-body Physics Research. In preparation.
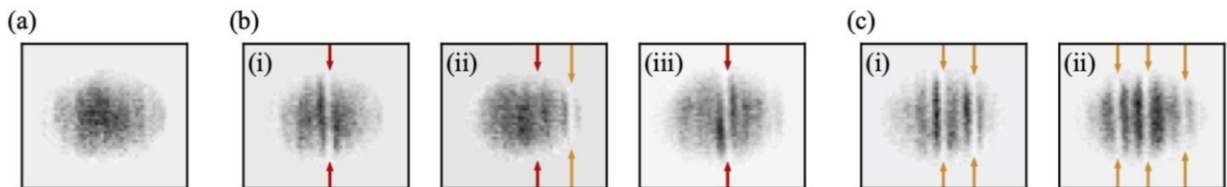
Figure 66. *The pre-processed images from the Dark Solitons in BECs dataset. The red arrows mark the location of the deepest depletion in the density fluctuations, while the orange lines mark the solitons' locations found from our OD. (a) An element of the no-excitation class. (b) Three elements of the single excitation class. (c) Tow representative elements of the other excitations class. Adapted from [4].*

Ray-based Classification Framework for Quantum Dot Devices

Justyna P. Zwolak
Sandesh S. Kalantre (University of Maryland)
Thomas McJunkin (University of Wisconsin-Madison)
Brian J. Weber (ShanghaiTech University)
Jacob M. Taylor (NIST PML)

There are myriad quantum computing approaches, each having its own set of challenges to understand and effectively control their operation. Arrays of electrostatically defined *quantum dots* (QDs) present at the interface of semiconductor devices is one such approach [1]. The still prevalent practice of tuning QDs manually or in a semi-automated fashion is a relatively time-consuming procedure, inherently impractical for scaling up device complexity or other integrated applications. Even tuning a double QD constitutes a nontrivial task, with each dot typically being controlled by at least three metallic gates, each of which influences the number of electrons in the dot, the tunnel coupling to the adjacent lead, and the interdot tunnel coupling. At the same time, given the progress in the construction of multi-QD arrays in both one and two dimensions (1D and 2D, respectively) [2, 3], it is imperative to replace the current practice of manual tuning to a desirable electronic configuration with a standardized automated method.

As part of our efforts to develop an autonomous tuning system for multi-QD devices, we have recently proposed a new framework for classifying simple high-dimensional geometrical structures: *the ray-based classification (RBC)* [4]. Rather than relying on the full N-dimensional data tensor, the RBC utilizes a minimal collection of 1D representations, called *rays*, to assess the geometry of the structure enclosing a given point x_o. The rays are uniquely defined by a set of colinear points

$$\Re(x_o, x_f) = \{x | x = (1-t)x_o + tx_f, t \in [0,1]\},$$

where x_o is the ray origin and x_f is the measurement termination point. A sufficiently long ray is expected to intersect the boundary of the structure enclosing x_o and may cut across more than just one structure (see Figure 67(a) for a depiction). The set of intersection points between a given ray and the structure(s) boundaries F_{x_o, x_f} is called the *feature set* in the RBC framework. The natural order in F_{x_o, x_f} given by the 2-norm distance function allows us to choose a feature nearest to x_o, so-called the *critical feature*.

A collection of M rays of a fixed length r centered at x_o defines the *M-projection* of x_o. As can be seen in Figure 67(b), it is possible for an M-projection to include rays without any features, that is F_{x_o, x_ℓ} could in general be empty for $1 \leq \ell \leq M$. Finally, the *point fingerprint*
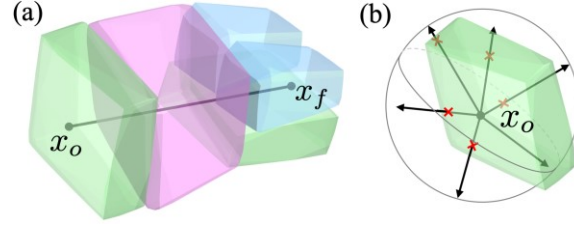
Figure 67. *(a) Visualization of a sample ray $\Re(x_o, x_f)$ from x_o to x_f in $\mathbb{R}^3$. Different colors represent different classes pf polytopes. (b) A sample M-projection centered at x_o. The black arrows represent the $M = 6$ rays and the red marks indicate the critical features used to determine fingerprints. Adapted from [4].*

of x_o, which is the primary object of the RBC, is defined as

$$\mathfrak{F}_{x_o} = [\Gamma(F_{x_o, x_1}), \dots, \Gamma(F_{x_o, x_M})],$$

where $\Gamma: \{F_{x_o, x_i}\}_1^M \to [0,1]$ is a *weight* function given as

$$\Gamma(F_{x_o, x_i}) = \begin{cases} \gamma(x_i^{(c)}) & \text{if } F_{x_o, x_i} \neq \emptyset \\ 0 & \text{if } F_{x_o, x_i} = \emptyset \end{cases}$$

with $x_i^{(c)}$ denoting the critical feature of F_{x_o, x_i} and $\gamma: \mathbb{N}^{>0} \to [0,1]$ being some decreasing function. Recently, we have proved that if sufficiently many rays in appropriate directions are chosen from x_o, the point fingerprint $\mathfrak{F}_{x_o}$ is sufficient to qualitatively determine the geometry of the convex polytope enclosing x_o [5]. Formally, the correspondence between the point fingerprint $\mathfrak{F}_{x_o}$ and the set of classes of polytopes filling the $\mathbb{R}^N$ space can be formulated as follows:

Ray-based classification: Consider a point $x_o \in \mathbb{R}^N$ and its fingerprint $\mathfrak{F}_{x_o}$. Given a set of bounded and unbounded polytopes filling an N-dimensional space and belonging to C distinct classes, $C \in \mathbb{N}$, determine to which of the classes the polytope enclosing x_o belongs.
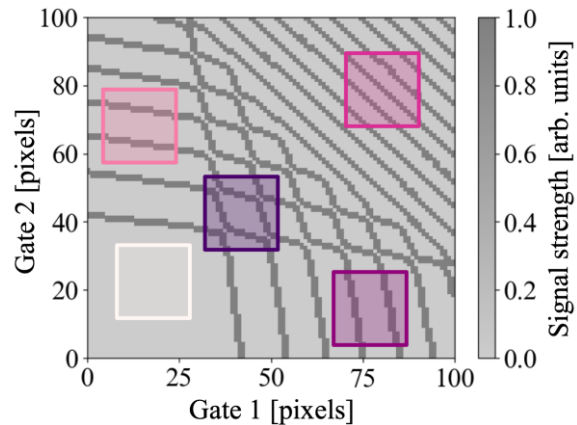
Figure 68. *A sample simulated 2D map showing the different bounded and unbounded polytopes in $\mathbb{R}^2$.*

The RBC framework naturally applies to identifying QD states which manifest themselves as distinct geometrical patterns in charge sensor response as a function of the gate voltages, see Figure 68. The performance of the "fingerprinting" approach in differentiating between the various states of QD devices was initially tested using an ensemble of 20 simulated double dot devices. A grid of 1369 points with 10 and 12 evenly spaced traces per point (a total of 27380 points) was sampled for each device. This data allowed us to compare performance for five M-projections, with $M = 3, 4, 5, 6,$ and 12. The ray length was varied between 10 and 80 pixels, with the simulation resolution chosen such that 30 pixels correspond to an average separation between transition lines. The performance test of the classifier over 50 training and validation tests per combination of ray's number and length revealed that 6 rays of length 60 pixels are sufficient to obtain a classification accuracy of 96.4 % (standard deviation 0.4 %). This corresponds to 60 % data reduction compared to the traditional image-based classification. A preliminary analysis of the RBC framework using simulated triple-dot system suggest even more significant data reduction [4]. Importantly, the model used in these tests was a relatively small deep network with three fully connected layers (256, 128, and 32 units, respectively) and five possible classes (i.e., no dot; left, central, and right single QD; double QD).

The RBC has also been validated experimentally [6]. Figure 69(a) shows the preprocessed charge sensor (CS) signal for six evenly distributed rays measured from a point shown in panel (b). Using a series of measurement of a double QD device, we showed that RBC surpasses the 82 % accuracy benchmark from the experimental implementation of image-based classification techniques from prior work [7], while reducing the number of measurement points needed by up to 70 %, see Figure 69(c). When implemented as part of an autotuning system (off-line), we found 78.7 % tuning success rate for a set of 225 uniformly sampled initial points, with additional 10.2 % of the points landing in an area that moderately resembles double-QD features. For comparison, the success rate for tuning using the 2D scans reported in [7] was 75 %, even though the tuning with RBC was started from a region enclosing approximately 18 electron transitions while the image-based tuning was initiated at most nine transition lines away from target region. The RBC has also been tested in off-line tuning in a three-dimensional (3D) space formed by a series of scans in the plunger gates space taken at different values of the middle barrier gate. We found an overall success rate of 67 % for tuning in 3D over 100 tuning runs initialized within a voltage region where no double QD are formed.

The reduction in measurement cost is a significant gain for time intensive QD measurements and is a step forward toward the scalability of these devices. Our work provides both theoretical motivation for RBC and experimental validation of the efficient state identification and optimization with machine learning techniques for nontraditional measurements in quantum systems with high-dimensional parameter spaces and time-intensive measurements.

[1] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. K. Vandersypen. Spins in Few-Electron Quantum Dots. *Reviews of Modern Physics* **79**:4 (2007), 1217-1265.

[2] [D. M. Zajac, T. M. Hazard, X. Mi, E. Nielsen, and J. R. Petta. Scalable Gate Architecture for a One-Dimensional Array of Semiconductor Spin Qubits. *Physical Review Applied* **6**:5 (2016), 054013.

[3] U. Mukhopadhyay, J. P. Dehollain, C. Reichl, W. Wegscheider, and L. M. K. Vandersypen. A 2 × 2 Quantum Dot Array with Controllable Inter-Dot Tunnel Couplings. *Applied Physics Letters* **112**:18 (2018), 183505.

[4] J. P. Zwolak, S. S. Kalantre, T. McJunkin, B. J. Weber, and J. M. Taylor. Ray-based Classification Framework
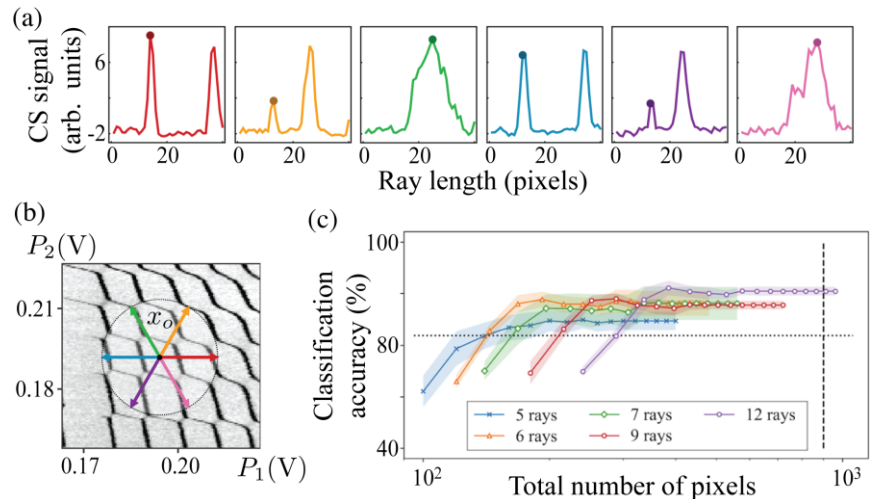
Figure 69. *Visualization of the ray-based fingerprinting framework. (a) The preprocessed charge sensor (CS) signal for six evenly distributed rays measured from a point in a double QD state shown in panel (b). (c) Classifier performance for different numbers of rays as a function of the total number of pixels measured for off-line experimental data (averaged over $N = 20$ models). The dotted black horizontal line represents the benchmark from [7], while the dashed vertical line represents the minimum data requirement for a CNN classifier as used in [7]. The connecting lines are a guide for the eye only. The error bands correspond to one standard deviation.*

for High-Dimensional Data. In *Third Workshop on Machine Learning and the Physical Sciences (ML4PS @ NeurIPS 2020)*, Vancouver, Canada, December 11, 2020.

[5] B. J. Weber, S. S. Kalantre, T. McJunkin, J. M. Taylor, and J. P. Zwolak. Theoretical Bounds on Data Requirements for the Ray-Based Classification. *SN Computer Science* **3** (2022), 57.

[6] J. P. Zwolak, T. McJunkin, S. S. Kalantre, S. F. Neyens, E. R. MacQuarrie, M. A. Eriksson, and J. M. Taylor. Ray-Based Framework for State Identification in Quantum Dot Devices. *PRX Quantum* **2** (2021), 020335.

[7] J. P. Zwolak, T. McJunkin, S. S. Kalantre, J. P. Dodson, E. R. MacQuarrie, D. E. Savage, M. G. Lagalla, S. N. Coppersmith, M. A. Eriksson, and J. M. Taylor. Autotuning of Double-Dot Devices In Situ with Machine Learning. *Physical Review Applied* **13** (2020), 034075.

Charge Field Decomposition and State Identification for Quantum Dot Data

Justyna P. Zwolak
Zachary J. Grey
Joshua Ziegler
Brian J. Weber (ShanghaiTech University)

Confining electrons in arrays of semiconductor nanostructures, called quantum dots (QDs), is a promising approach to quantum computing. Due to the ease of control of the relevant parameters, fast measurement of the spin and charge states, relatively long decoherence times, and their potential for scalability, QDs are gaining popularity as building blocks for solid-state quantum devices. In semiconductor quantum computing, devices now have tens of individual electrostatic and dynamical gate voltages that must be carefully set to isolate the system to the single electron regime and to realize good qubit performance. However, the relevant parameter space scales exponentially with QD number (dimensionality), making heuristic control unfeasible for even moderate scale applications.

There has been recent work using machine learning (ML) techniques as part of the automating process. However, training ML models require large amounts of labeled data indicating the true state of the device for a given voltage range. So far, ML efforts for QD rely on datasets that either come from simulations (and thus lack important features representing real-world noise and imperfections) or are labeled manually (and subject to qualitative or erroneous classification). Automatic labeling will streamline the creation of large, accurate datasets for purposes of ML training. Moreover, the ML approach currently lacks explainable and interpretable features for reliable diagnostics. Thus, it is desirable to have a simplified and theoretically motivated automated protocol for labelling experimental data.

The configuration space of a QD array supports a charge field, whose level-sets create an irregular polytope tiling of the space, with each polytope indicating a distinct quantum state (e.g., a left or right single dot for a double QD device). The polytope shapes provide information about electron behavior within the discrete states they represent. Polytopes with similar characteristics will cluster together, allowing the subdivision of configuration space into distinct *domains* where the system exhibits a consistent behavior. Our work is aimed at creating automatic procedures for domain decomposition, and if possible automatic characterization of individual polytopes within each domain. The project currently focuses on the case of double QD devices, where polygonal tilings are readily understandable.

Our team has focused its effort in three directions, which fit under the subheadings of gradient persistence methods, activity detectors (with a focus on Harris activity detectors), and the RBC method, all described below. Each method has shown substantial successes for the double QD data and has a greater or lesser potential for generalization to higher dimension. We are assessing our approaches on simulated and real-device datasets, to concretely evaluate their suitability and robustness.

Gradient Persistence Methods. When smoothing a noisy dataset, the convolutional scale must be large enough to eliminate noise yet small enough to retain essential features. Because datasets have no predetermined scale on which features are known to exist, making *a priori* choices of convolutional scale is difficult. The persistence method deals with this problem by convolving datasets essentially on all scales and observing which features most strongly persist across scales. Features that appear in only a small number of scales are considered noise. Persistent features are retained.

At each convolutional scale, the gradient of the charge-field is computed, and a density histogram of gradient vectors is created. Clustering within the gradient histogram is detected using a potential-differencing method. Clusters that persist across multiple scales are kept. Other clusters are discarded. Persistent clusters are used to divide the configuration space into domains by assigning to each cluster the dominant locations within the configuration space that the vectors of that cluster came from.

Experimentally, this method was implemented on several dozen simulated datasets, with consistent, accurate results; see Figure 70 for an example. Alongside these strengths, a weakness of this method is that second-order information—which is necessary for a complete domain decomposition—is substantially less persistent than first order information. In dimension 2 this has been addressed, but the problem compounds in higher dimensions as second-order information persists on ever fewer scales.
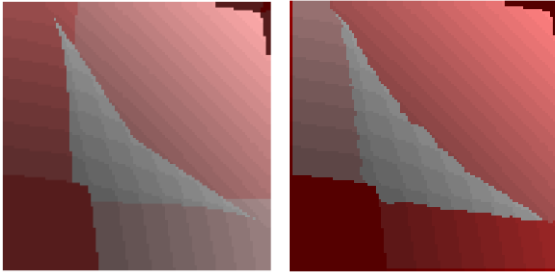
Figure 70. *Domain decomposition of simulated data using the gradient-persistence method. The theoretical decomposition is shown on the left and the decomposition from the gradient persistence method on the right.*
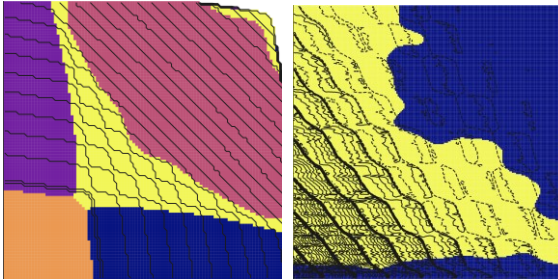
Figure 71. *K-mean clustering results over the four vectorized features for the simulated (left) and experimental (right) devices. K is set to five for the simulated device while K is set to two for the experimental device.*

Local Activity Detection and Harris Detectors. For practical applications, particularly relevant are the domains with hexagonal tilings. These are precisely where second derivatives of the charge field show the most activity, and where the persistence methods are weakest. What is needed is a specialized method for detection of second-order features that works independent of scale. Such methods appear in the image-recognition literature in the form of "corner detectors," and in particular the Harris scale-invariant corner detector.

To begin, given a function $f \colon \mathcal{X} \subset R^N \to R$, an active subspace analysis proceeds by eigendecomposition of the positive definite, N-by-N, matrix

$$ \boldsymbol{C} = \int_{\mathcal{X}} \nabla f(\boldsymbol{x}) \nabla f(\boldsymbol{x})^T d\mu(\boldsymbol{x}). $$

This global integral produces the *active subspaces* of f, defined as the span of the non-zero eigenvectors of $\boldsymbol{C}$. The sizes of the eigenvectors roughly indicate how much f changes on average throughout the domain—the *activity* of f in that eigendirection. This can be used as a proxy for second-order activity: a linear function will have one active direction, whereas quadratics can produce any desired eigenspace decompositions of $\boldsymbol{C}$. To detect local rather than global second order activity, we use the Harris operator, given as

$$ h(\boldsymbol{x}; \sigma_I, \sigma_d) = \det\left|\widetilde{\boldsymbol{C}}(\boldsymbol{x}; \sigma_I, \sigma_d)\right| - \kappa tr^2\left(\widetilde{\boldsymbol{C}}(\boldsymbol{x}; \sigma_I, \sigma_d)\right), $$

where

$$ \widetilde{\boldsymbol{C}}(\boldsymbol{x}; \sigma_I, \sigma_d) = \sigma_d{}^2 (g(\cdot\,; \sigma_I) * \nabla f \nabla f^T(\cdot\,; \sigma_d))(\boldsymbol{x}) $$

is the component-wise convolution of the approximated gradient outer product with kernel g. Using kernels of various sizes leads to the scale-invariant version of this operator and creates a *characteristic scale* or *Harris scale* at each point, for which the *cornerness* quantity of the charge-field is maximized at that point. Locations in the configuration space with high cornerness are regions with high second-order activity and are likely to support hexagonal tilings.

To obtain the domain decomposition, we retain four important features: the three distinct smooth products of partials from the outer product of the approximated gradient (i.e., three unique entries of the symmetric matrix $\widetilde{\boldsymbol{C}}(\boldsymbol{x}; \sigma_I, \sigma_d)$), and the Harris operator response. Subsequent K-means clustering is then applied to this four-dimensional feature space to create a decomposition. Figure 71 depicts a resulting domain decomposition for both a simulated (left) and an experimental (right) device.

The strength of this method is its natural ease with second-order charge field features, and its robust method of isolating the hexagonally tiled domain. It is not yet clear how robust to noise this method may be, but it can always be combined with other noise-reduction strategies (see below). Another open question is how to generalize the Harris detector to higher dimensions, where second-order activity and corner detection are both substantially more complicated. Generalizations will require not just traces and determinants of the local activity matrices, but also other symmetric invariant polynomials. Moreover, higher dimension considerations may benefit from the combined perspective involving global dimension reduction informed by eigenspaces of $\boldsymbol{C}$ and local activity $\widetilde{\boldsymbol{C}}(\cdot\,; \sigma_I, \sigma_d)$ defining decompositions over a reduced dimension subspace along which the decomposition is constant in the orthogonal directions. Alternatively, studying the classification over transformations inferred from non-linear (non-subspace based) dimension reduction may also enable higher dimension considerations.

RBC Method. We have successfully implemented a version of the ray-based classification (RBC) schema, which was developed in an earlier series of papers out of our group [1-2]. In this method, from any given "observation point" in configuration space, a sequence of evenly spaced rays is emitted, and the distance to polytope edge is recorded. In the original idea, the polytope is identified using only these distances, but in the present context we possess greater knowledge of the charge function and can record its gradient at polytope edges. Each ray records one gradient vector upon intersection with an edge, and on this collection of gradient vectors we use a clustering algorithm to find dominant gradient directions. In locations where only one gradient cluster
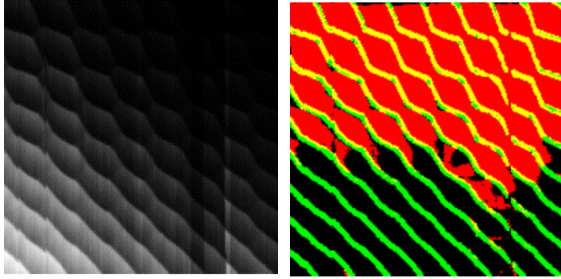
Figure 72. *RBC domain decomposition for a real device. Original charge field (left) has noise and spurious sensor artifacts. The RBC-based domain decomposition (right) occurs after a de-noising process.*

is detected, the observation point is in a linear region of the charge field. Where two or more gradient clusters are observed, we are in a polygonally tiled region, and we can observe the shape of the polygon.

This method has the advantage of being suited to identifying individual polygons, which is an even more valuable goal than creating a domain decomposition. Another advantage is vastly reduced run-time burden, due to few or no computationally expensive convolutions being required (run-time burdens grow exponentially with dimension). Even more broadly, this method is well-suited to higher dimensional generalization. There is no serious theoretical barrier to using RBC methods in higher dimensions, although substantial technical hurdles exist in the form of understanding details of high-dimensional geometry, such as ray placement. Our group has already done much work on this in the form of creating ray-placement estimates and rough polytope classifications [3].

A potential weakness is noise sensitivity. To obtain sharp polygon boundaries, little or no smoothing can be allowed to take place, which means noise will be a problem. We have developed noise-elimination techniques which work well on tested devices. Determining the robustness of these methods across various noise regimes and device architectures is an ongoing effort.

In Figure 72, even though the RBC method can be used to identify individual polygons, we used the RBC method to produce the less-precise domain decomposition. In that figure, every pixel is an observation point, and decides independently which domain it is in.

Summary. Initial computational experiments have proved promising in replicating known qualitative classifications of simulated and real QD devices. Our three domain decomposition methods have each demonstrated strong successes and show a great potential for generalization to higher dimensions. Our next step is to quantitatively assess strengths and weakness of each approach by studying performance of the methods across many datasets from both real-world and simulated devices and a variety of noise regimes. On completing these quantitative assessments, best strategies or combinations of strategies will be created with the aim of building a robust, general domain decomposition process for double QD data, which we believe is in sight.

Future work will focus on generalizing to higher dimensions, where additional challenges are expected, particularly related to high run-time expense and sparsity of available data. Understanding of higher-dimensional geometry will be a crucial component of the upcoming work. An additional avenue for exploration is the use of more sophisticated approaches to clustering salient features to mitigate boundary effects, which in some cases create serious problems. We will eventually integrate our systematic analytic choices as constraints or training sets in an ML framework. The ultimate hope is to combine this work with ray-based tracing for real-time classification in the laboratory.

[1] J. P. Zwolak, S. S. Kalantre, T. McJunkin, B. J. Weber, and J. M. Taylor. Ray-based classification framework for high-dimensional data. In *Third Workshop on Machine Learning and the Physical Sciences (ML4PS @ NeurIPS 2020)*, Vancouver, Canada, December 11, 2020.

[2] J. P. Zwolak, T. McJunkin, S. S. Kalantre, S. F. Neyens, E. R. MacQuarrie, M. A. Eriksson, and J. M. Taylor. Ray-based Framework for State Identification in Quantum Dot Devices. *PRX Quantum* **2**:2 (2021), 020335.

[3] B. J. Weber, S. S. Kalantre, T. McJunkin, J. M. Taylor, and J. P. Zwolak. Theoretical Bounds on Data Requirements for the Ray-Based Classification. *SN Computer Science* **3**:1 (2022), 57.

Noisy Quantum Dot Devices

Joshua Ziegler
Justyna P. Zwolak
Jacob M. Taylor (NIST PML)
Thomas McJunkin (NIST PML)
Sandesh Kalantre (University of Maryland)
E. S. Joseph (University of Wisconsin-Madison)
Benjamin Harpt (University of Wisconsin-Madison)
D. E. Savage (University of Wisconsin-Madison)
M. G. Lagally (University of Wisconsin-Madison)
Mark A. Eriksson (University of Wisconsin-Madison)

Gate-defined quantum dots (QDs), in which electrons in a semiconductor are trapped in electric potential wells, are a promising platform for quantum computing due to their small device footprint and the possibility of operation at temperatures of a few degrees Kelvin. However, due to minor inconsistencies inherent to the fabrication process, initialization of these devices with potentials that properly trap electrons cannot be achieved deterministically. Currently, initialization is performed mostly manually via heuristic tuning, although some recent progress has been made towards automating elements of this process.
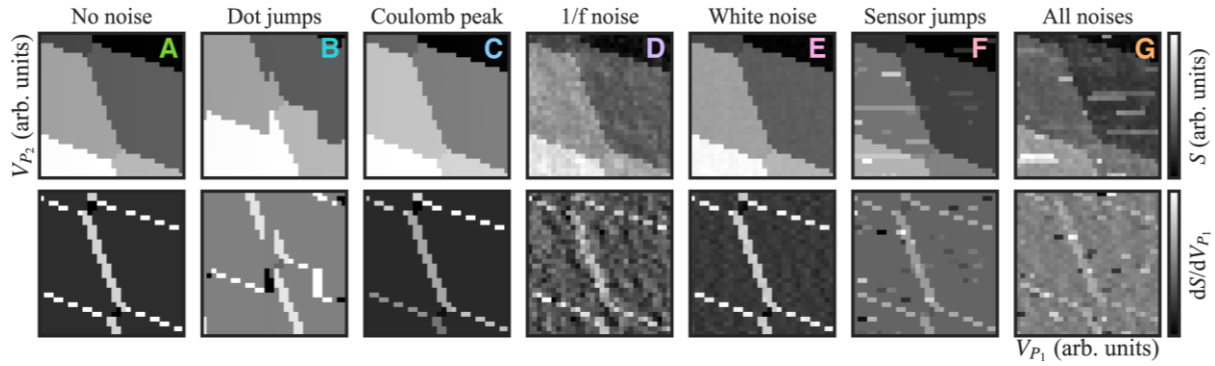
Figure 73. *Various types of noise added to a simulated charge stability diagram for a QD device in a "double QD" state, with electrons trapped in two potential wells. The bright lines in the bottom row indicate a change of electron occupation in the device. All noise magnitudes are the same as for the best accuracy quoted except for Coulomb peak and dot jumps which have increased effect for visibility.*

The best strategies for automating initialization rely on supervised machine learning (ML) methods to autonomously identify the state of the device as parameters are tuned and use an algorithmic or ML optimizer to find the desired state [1, 2]. However, training such ML models requires large amounts of data with labeled device states. There are two paradigms for preparing datasets for training, each having its own advantages and limitations. The first is to use unrealistically noiseless simulated data, which not only necessitates specialized pre-processing of the experimental data, but also makes the trained models less robust when implemented in the experiment. The other paradigm relies on manual labeling of experimental data. This is laborious, less reliable, and not scalable, especially as gate-defined QD devices change or grow in qubit numbers and complexity. Thus, approaches demonstrated to date force a choice between scalability and robustness. Furthermore, ML methods typically do not include assessment of the quality of the input data, leading to unknown prediction reliability.

An idealized noiseless charge stability diagram shows lines when an electron is added to the device's trapping potential landscape overlayed on a uniform background, as shown in Figure 73A. In order to overcome the issues described above, we have expanded the capabilities of our QD simulator [3] to include various types of realistic, physical noise. Depending on the type of added noise, the appearance of simulated data changes as shown in Figure 73B-F. Using a set of 542 manually labeled experimental scans for testing, we find an accuracy improvement of about 46 %, from 48.7(5.5) % when trained on noiseless simulated data to 95.0(9) % when trained on simulated data with all relevant noises added. We show how each added noise individually contributes to this performance in Figure 74B-F. Figure 74G shows the performance for the optimized noise combination. To get best performance sensor jumps, 1/f noise and white noise are varied together to yield a varied signal-to-noise ratio.

We then generate a similar dataset with a broad variation of noise and use the performance of the noise-trained ML state classifiers on it to develop an ML data quality control (DQC) module. We confirm the utility of the DQC module by showing a correlation between the accuracy of the state classifiers and the assessed quality. Specifically, for experimental data assessed as "high quality" the state classifier has 96.4(9) % accuracy, for data assessed as "moderate quality" it has 91.9(2.1) % accuracy, and for "low quality" data the state classifier has 69.3(5.6) % accuracy. This validates the ML DQC module as a useful tool to filter data that might lead to poor performance of an ML state classifier and, consequently, the autotuner failure. Currently, we are taking this approach a step further by training ML models to identify specific types of device and data imperfections.

With this approach, we expand the applicability of ML-based autotuning strategies to non-ideal devices by making the ML models more robust against potential failure modes. This is especially important when considering future, large-scale QD devices. More broadly, we show that making simulated data more physical can greatly improve the efficacy of ML models when deployed in a real lab environment, which may be a useful insight for other experiments combining ML and physics. An article on this work is currently in review [4].

[1] J. P. Zwolak, T. McJunkin, S. S. Kalantre, J. P. Dodson, E. R. MacQuarrie, D. E. Savage, M. G. Lagally, S. N. Coppersmith, M. A. Eriksson, and J. M. Taylor. Autotuning of Double-Dot Devices In Situ with Machine Learning. *Physical Review Applied* **13**:3 (2020), 034075.

[2] H. Moon, D. T. Lennon, J. Kirkpatrick, N. M. van Esbroeck, L. C. Camenzind, Liuqi Yu, F. Vigneau, D. M. Zumbühl, G. A. D. Briggs, M. A. Osborne, D. Sejdinovic, E. A. Laird, and N. Ares. Machine Learning Enables Completely Automatic Tuning of a Quantum Device Faster Than Human Experts. *Nature Communications* **11**:1, (2020), 4161.

[3] J. P. Zwolak, S. S. Kalantre, X. Wu, S. Ragole, J. M. Taylor. QFlow Lite Dataset: A Machine-learning Approach to the Charge States in Quantum Dot Experiments. *PLoS ONE* **13**:10 (2018), 1–17.
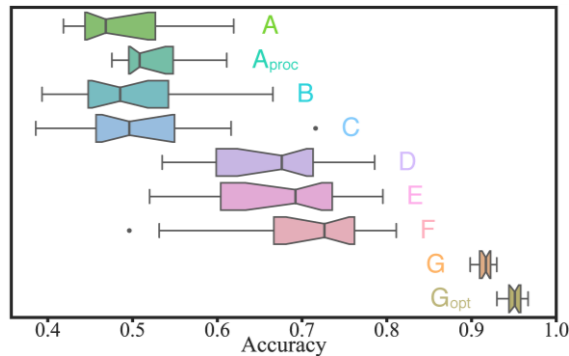
**Figure 74.** *Box plots depicting the accuracies of state classifier models trained on the data as indicated in Figure 73. Various types of noise added to a simulated charge stability diagram for a QD device in a "double QD" state, with electrons trapped in two potential wells. The bright lines in the bottom row indicate a change of electron occupation in the device. All noise magnitudes are the same as for the best accuracy quoted except for Coulomb peak and dot*

[4] J. Ziegler, T. McJunkin, E. S. Joseph, S. S. Kalantre, B. Harpt, D. E. Savage, M. G. Lagally, M. A. Eriksson, J. M. Taylor, J. P. Zwolak. Toward Robust Autotuning of Noisy Quantum Dot Devices. Preprint arXiv:2108.00043, 2021.

Physics-driven Tuning of Quantum Dot Charge States

Joshua Ziegler
Justyna P. Zwolak

Gate-defined quantum dots (QDs), formed by trapping electrons using finely calibrated electric gates, are a promising quantum computing platform. Unfortunately, working with single electrons means that the operating regime strongly depends on nanoscale details (e.g., material impurities and fabrication defects) that are difficult to control across devices. For QD devices to scale effectively, this operating regime must be identified using automated methods. There have been several recent demonstrations of automation of QD tuning, yet these approaches have only taken limited advantage of the designed effect of the electric gates on the device state. Currently, we are developing a modular autotuning procedure that navigates to the operating regime of a QD device by first coarsely tuning to a specific QD topology (i.e., a known number of QDs), then calibrating the real effect of the gates on individual QDs (the so-called "virtual gates"), followed by a two-step ray-based algorithm for tuning to a specific charge state.

In QD devices with the overlapping gates architecture (the type of device we have mostly been focusing on) there are two types of electric gates used to control the flow of electrons. The plunger gates serve to create potential wells that can hold electrons. The barrier gates

serve to prevent the flow of electrons between trapping potentials or into the leads. Together, these gates serve to create an electrical potential landscape in which electrons can be trapped and manipulated. Typically, in heuristic tuning procedures, coarse voltage values for barriers and plungers that cutoff electron flow are identified during the so-called "cold-start" stage. These voltages are then used as a starting point for scanning pairs of plungers over a large range of voltages to identify the target regime, that is, voltage ranges where single electrons are trapped in each potential well.

Previously, we have tackled tuning of pairs of gates using algorithms that seek to maximize a fitness function designed to quantify the difference between the target and the captured state of the device. However, none of the optimization algorithms we used took into account information about the pre-existing relationship between the states [1, 2] which led to repeated failures for tuning initialized far from the target area in voltage space. The several algorithms for navigating to specific charge states that have been demonstrated recently also show unsatisfactory performance [3-5].

To improve the effectiveness and efficiency of automatically setting the device topology, as well as tuning to a desired charge configuration, we propose a method that leverages the intended effect of each gate on the overall QD device state. In principle, changing voltages on a particular plunger gate should lead to a change of the electron occupation only in the corresponding potential well. We show that such targeted control (made possible due to the virtual gates) combined with the knowledge of the expected device topology (i.e., relative position of states in the plunger-plunger space) and an ML model trained to identify the captured state of the QD device, enable meaningful, efficient, and direct navigation to a target region in voltage space over a large range of voltages. Implementing this tuning strategy, which we call an "action-based tuning", we find that we can navigate to the target region (in our case double QD state) both within noise-augmented simulated scans (with 99.8(2) % success rate) and within a pre-measured experimental scan (with 99.2 % success rate). Figure 75 shows the performance of the state tuning algorithm, with panel (a) depicting the initialization points and panel (b) showing the state tuning result.

However, since local device measurements only contain information about changes in the electron occupation of each potential well, navigation to a specific charge occupation requires a more nuanced approach. We tackle this problem by following the typical procedure of first emptying a QD device of electrons and then loading a desired number back into each potential well. To achieve this most reliably, prior to the emptying-loading stage, we calibrate the capacitive coupling of each gate to the potential wells using an ML model trained for this purpose. This enables us to take one dimensional (1D) scans (the rays) that change only the
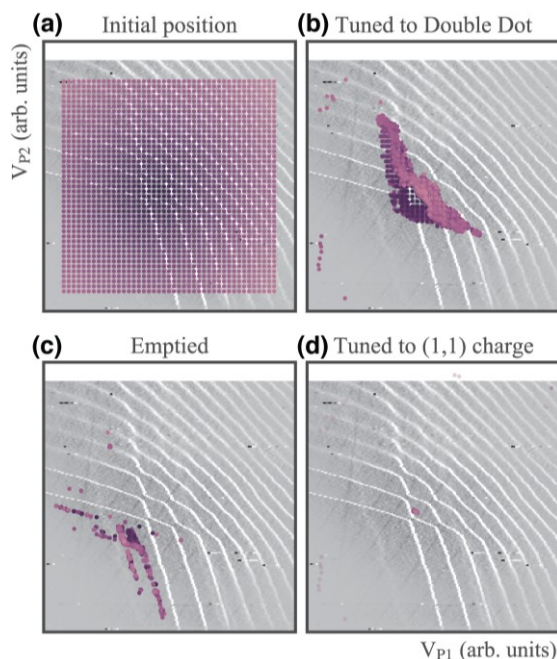
Figure 75. *Results of our tuning algorithm applied to this experimental scan. The color scale shows the initial distance from the center of the target charge region with darker being closer. (a) Initial positions. (b) Positions after tuning to the double QD region (isolated charge in two isolated potential wells). (c) Positions after tuning to the empty region. (d) Positions after tuning to the (1,1) charge region. Here the concentration of points near the center of the image are in the correct state. Note that the points in the last panel have some transparency to visualize density.*

electron occupation of a single potential well. We then use conventional algorithms to detect transitions in the 1D measurements as electrons are loaded off and back on each potential well. Initial tests of this emptying-loading algorithm on simulated data show a high success rate for navigating from any point within the double QD region (the outcome of the action-based tuning step) to the single-electron regime in moderately noisy simulated scans.

Figure 75 (c) shows the positions where the algorithm declares the device to be in the empty state when initiated for the points correctly tuned to the double QD regime in panel (b). The success rate for the emptying stage is 92.3 %, with additional 6.9 % points terminating in the correct region due to boundary effects. Finally, Figure 75(d) shows the final positions of the loading algorithm, with the target charge state set to (1,1). The success rate for the loading phase is 99.0 %, with nearly all final positions concentrated in the correct state. This gives an overall 98.2 % success rate for off-line tuning to a specific charge state for this sample experimental scan. Combining these algorithms yields a complete tool for navigation from a coarsely tuned device to a specific charge state, enabling effective, efficient, and fully automated tuning.

[1] J.P. Zwolak, T. McJunkin, S.S. Kalantre, J.P. Dodson, E.R. MacQuarrie, D.E. Savage, M.G. Lagally, S.N. Coppersmith, M.A. Eriksson, and J.M. Taylor. Autotuning of Double-Dot Devices In Situ with Machine Learning. *Physical Review Applied* **13**:3 (2020), 034075.

[2] J.P. Zwolak, T. McJunkin, S.S. Kalantre, S.F. Neyens, E.R. MacQuarrie, M.A. Eriksson, and J.M. Taylor. Ray-Based Framework for State Identification in Quantum Dot Devices. *PRX Quantum* **2**:2 (2021), 020335.

[3] R. Durrer, B. Kratochwil, J.V. Koski, A.J. Landig, C. Reichl, W. Wegscheider, T. Ihn, and E. Greplova. Automated Tuning of Double Quantum Dots into Specific Charge States Using Neural Networks. *Physical Review Applied* **13**:5 (2020), 054019.

[4] S. Czischek, V. Yon, M. Genest, M. Roux, S. Rochette, J. Camirand Lemyre, M. Moras, M. Pioro-Ladrière, D. Drouin, Y. Beilliard, R. Melko. Miniaturizing Neural Networks for Charge State Autotuning in Quantum Dots. *Machine Learning: Science and Technology* **3** (2021), 015001.

[5] M. Lapointe-Major, O. Germain, J. Camirand Lemyre, D. Lachance-Quirion, S. Rochette, F. Camirand Lemyre, and M. Pioro-Ladrière. Algorithm for Automated Tuning of a Quantum Dot into the Single-Electron Regime. *Physical Review B* **102**:8 (2020), 085301.

Transition to Open Source Visualization Software

William George
Terence Griffin
Sandy Ressler
Steven Satterfield
William Sherman
Simon Su
Judith Terrill
Cory Quammen (Kitware)
Scott Wittenburg (Kitware)

As part of our research on Virtual Measurements and Analysis, the ACMD High Performance Computing and Visualization Group (HPCVG) operates a fully immersive visualization environment (IVE) and has developed high end visualization (HEV) software to run it. We started developing this software for our IVE more than two decades ago. During this time, we have upgraded and rewritten this software as our understanding of scientific visualization in an IVE developed and as outside innovations in hardware appeared. However, there were many limitations to our software. For example, it could only run on one specially configured operating system, and it had not kept up with recent advances in hardware visualization capabilities. Internally this software uses a scene-graph-based framework. Our IVE is on a critical path for the success of collaborations with several NIST research groups and is used at every stage of these

collaborations. These projects are diverse and span applications from nanotechnology to medical to materials, and often contribute to standard reference data and materials. For example, the IVE was essential in the success of NIST's development of standard reference materials (SRMs) for the measurement of the flow of cement paste, mortar, and concrete. To take advantage of recent advances in visualization hardware, we are moving our IVE to ParaView, a fully open source software environment.

The ParaView software system is complex. Internally it uses a pipeline and proxy-based framework. The software consists of a Qt interface which uses over 2000 VTK C++ classes to produce visualizations. It runs in an IVE as well as on Windows, Mac OS X, Linux, SGI, IBM Blue Gene, Cray and various Unix workstations, clusters, and super-computers. It supports rendering shaders. It has a new real-time path-tracing back end using NVIDIA RTX technology. ParaView extends the environments that the HPCVG HEV can work in, as well as provides access to real time ray tracing and global illumination made possible with the new GPUs. Access to high-end GPU rendering will continue to grow as ParaView adopts the ANARI rendering standard from the Khronos Group.

This year we completed the SAVG (NIST internal format) reader plugin for ParaView, and a utility to convert the thousands of NIST SAVG format files to the to the subset defined for the SAVG plugin. In addition, we nearly completed an OpenXR option to the OpenVR plugin in ParaView to bring it into compliance with the Khronos standard. This creates a uniform environment across all Extended Reality (XR) so there is a common software environment across all our hardware.

In the coming year, we will add support for different coordinate systems to ParaView. This will enhance ParaView's scene graph capability. We will also add timeline file with command capability. NIST had developed the concept of a timeline file that contains a sequence of time deltas from the previous time, location, orientations, and scale that together create a path through a visualization. Interspersed throughout a SAVG file are commands. SAVG commands can alter the visualization by, for example, changing the visibility of parts of the visualization, or it might snap and save an image at a particular time, or it might save a movie of the entire timeline sequence or a subset. Including these features into ParaView will provide the capability to re-run a visualization session.

Finally, we developed a low-cost hardware methodology that enables the use of consumer VR tracking hardware to produce a CAVE-like experience in an office-environment enabling researchers to have a similar,
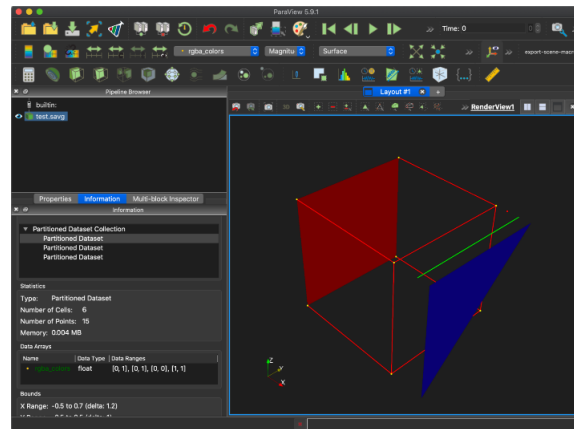
Figure 76. *A test image in SAVG format displayed in ParaView, showing lines, points, and polygons.*

if slightly reduced, experience to that of a full CAVE system such as the NIST CAVE. This is accomplished by using existing open-source tools along with 6-DOF position tracking systems such as the HTC/Valve Corp "Lighthouse" system. With these tools, collaborators working without the benefit of a full CAVE system can still work with the full user interface of a CAVE benefiting both development and deployment of the immersive visualization features.

[1] William Sherman. "Building an At-home CAVE-style VR Display to Broaden User Access." 6th Annual Conference on Campus Alliance for Advanced Visualization (CAAV 2021) West Lafayette IN, November 3, 2021.

Standards in Visualization

William Sherman
Sandy Ressler
Simon Su
Judith Terrill

Last year we joined the Khronos Group[15] to contribute to software and data standards that relate to high performance computing and standards for immersive visualization. Researchers from HPCVG are participating members of several software standards working groups, including OpenXR[16], ANARI[17], 3D Commerce[18], and glTF[19].

OpenXR. The membership of the Khronos OpenXR working group has continued to grow and includes all of the prominent companies working in the consumer extended reality (XR) arena. This past year new consumer software releases have reflected the growing adoption of

---
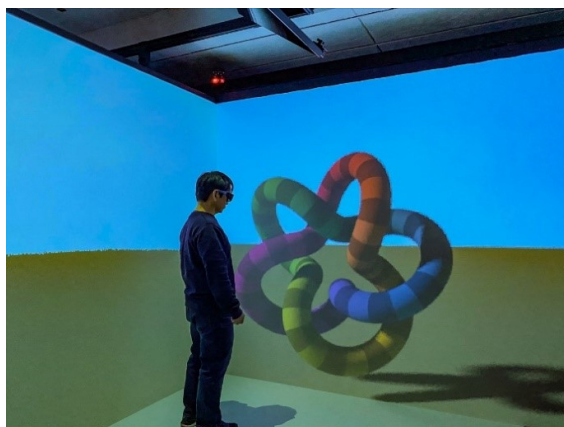
Figure 77. *Image generated with ANARI in the NIST CAVE.*

this standard, in a shift away from hardware-dependent interfaces. This is important as the number of available consumer products in the XR space continues to grow.

NIST, and the ACMD High Performance Computing and Visualization Group (HPCVG) in particular participating in OpenXR working group events as we continue to evaluate the standard for use of internal projects. Indeed, as part of the Transition to Open Source Visualization Software effort described on page 93, we are collaborating with Kitware Inc. to enable their ParaView scientific visualization tool to operate using the OpenXR API. Presently, the primary OpenXR focus has been toward head-mounted display (HMD/headset) systems, but as a goal of OpenXR is to be completely display agnostic, the HPCVG group will be bringing our expertise in large-format, CAVE-style displays such as the NIST Immersive Visualization Environment (IVE).

ANARI. The ANARI (Analytic Rendering Interface for Data Visualization) API is a Khronos Group led effort to provide a common library interface for the rendering of scientific data. In particular, ANARI provides a consistent interface for volumes, point clouds and polygonal mesh data that can be rendered using different methodologies, and tuned to the available graphics rendering hardware. This can be rasterized on a CPU-based system, or ray-traced renderings on an RTX-level GPU, or anywhere in between.

In the past year the ANARI standard has been released as a provisional 1.0 specification, with the anticipation that the final 1.0 API will be released in the Summer of 2022. One of ACMD's contributions to the ANARI effort has been to integrate it with the CAVE VR display system to ensure that the specification is amenable to needs of virtual reality rendering. The ParaView scientific visualization tool is also adopting the ANARI API for its advanced rendering output.

[1] William Sherman. "ANARI: A future for high-end visualization rendering." 6th Annual Conference on Campus Alliance for Advanced Visualization (CAAV 2021) West Lafayette IN, November 3, 2021.

WebXR Graphics

Sandy Ressler
Robert Thompson (NIST SPO)
Russel Sitka (US Holocaust Memorial Museum)
Jane Klinger (US Holocaust Memorial Museum)
Robert Ehrenreich (US Holocaust Memorial Museum)
Amanda Malinowski (NIST Library)

"WebXR" is a broad term that covers Web-based VR (virtual reality), AR (augmented reality), MR (mixed reality), all the "R"s. It is a coalescence of technologies that have much in common with each other. First, a device such as a headset, or soon a pair of glasses, displays images close to your eyeballs, in stereo. Second, your head is tracked so that as you move the images being displayed are modified, in real-time, to give the illusion that you are in the space of whatever you are being shown. Thus, you feel "immersed" in a scene of data or simulated objects. In the case of VR, you are cut off from the real world and are purely in a virtual world. In the case of AR, you remain in the real world because you still see the real world, with the graphics as an overlay. A current common method of using AR is via a mobile phone which uses its camera to capture the real environment and overlays graphics on it. A compelling example is holding your phone up to a sign in a foreign language, that is translated into your native language.

The Web part of WebXR comes from the ability to place these scenes inside of a Web browser. This is important from a technology dissemination point of view. We all have and use Web browsers. The Web has proven to be the most powerful information dissemination technology of the last few decades. The integration of immersive technologies with the Web leads us to standards and museums. Standards enable us to perform the integration and museums represent an "application" in great need of this integration and both benefit the public good, part of NIST's mission.

3D Object Input and Capture. Viewing a 3D object on a web site first requires the object itself. Historically we have obtained our objects either from CAD systems or simulation data. Missing from that set of objects are representations of "real" objects, artifacts that actually exist in real life. As hardware for capturing the shape and texture of real objects has plummeted, 3D object capture is becoming democratized. We recently published a blog article for the Khronos Group on just this topic [1].

There is a desire to scan items such as museum artifacts, or simply physical objects that are of interest. Allowing curators or the general public to view and manipulate objects that are too precious to handle, or inaccessible, in real life is an important benefit to the creation and access of objects in our virtual world. The goal of this aspect of our WebXR project is to codify
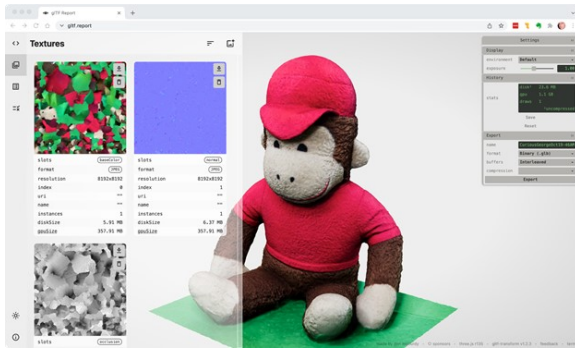
Figure 78. *Scanned object testing. A substitute for a real artifact displayed in gITF Report software [2, 3]*

standards, processes, and techniques to enable wider dissemination of priceless artifacts and to add information about these objects not easily seen via 2D representations. These types of capture processes fall into the domain of photogrammetry, and photometry. These are not new domains. Museums and forensic experts have been active in these domains for many years. We are beginning to assemble a working group of experts and end users to make NIST contributions in the context of web accessible objects and interactions with those objects. We've continued to hold meetings with staff of the United States Holocaust Memorial Museum as well as forensics experts at NIST.

Figure 78 illustrates a test artifact, standing in for an actual artifact, viewable on Sketchfab[2] a widely used 3D object distribution site. The use of practice artifacts is of great help as we refine our procedures and practices.

3D Standards Participation. Formal standards development is a long arduous process. We are involved with two particular efforts run by two different groups. First is the 3D Formats Group, part of the Khronos Group, which is responsible for the continued evolution of glTF, a 3D graphics format which is becoming the "jpeg of 3D." Second is the Immersive Web Group of the World Wide Web consortium which plans to "*...help bring high-performance Virtual Reality (VR) and Augmented Reality (AR) (collectively known as XR) to the open Web via APIs to interact with XR devices and sensors in browsers.*"

The 3D Formats Group is actively involved in a variety of extensions to glTF. Primarily concerned with achieving high quality and high performance, taking advantage of GPU hardware is a signature overarching goal of the group. glTF is becoming more and more popular as a representation of 3D objects. Our participation ensures that it is aligned to our, and the public's needs, and keeps us on the leading edge of 3D format developments. The Khronos Group is an industry-led consortium, and the participants are a highly skilled collection of technologists focused on real world practical problems. In particular, we assisted in helping the

Khronos Group move glTF into ISO and it is in the process of becoming an ISO standard by the fall of 2022.

[1] S. Ressler. "Will LiDAR Scanning break the 3D Asset Creation Bottleneck? September 15, 2021. URL: https://www.khronos.org/blog/will-lidar-break-the-3d-asset-creation-bottleneck

[2] S. Ressler. Curious George Vintage 1970s Stuffed Animal. URL: https://sketchfab.com/3d-models/curious-george-vintage-1970s-stuffed-animal-47fae6ca9ff64f1b823682799da54b82

[3] D. McCurdy. glTF.report Software. URL: https://gltf.report

Quantum Information

An emerging discipline at the intersection of physics and computer science, quantum information science is likely to revolutionize 21st century science and technology in the same way that lasers, electronics, and computers did in the 20th century. By encoding information into quantum states of matter, one can, in theory, enable phenomenal increases in information storage and processing capability. At the same time, such computers would threaten the public-key infrastructure that secures all of electronic commerce. Although many of the necessary physical manipulations of quantum states have been demonstrated experimentally, scaling these up to enable fully capable quantum computers remains a grand challenge. We engage in (a) theoretical studies to understand the power of quantum computing, (b) collaborative efforts with the multi-laboratory experimental quantum science program at NIST to characterize and benchmark specific physical realizations of quantum information processing, and (c) demonstration and assessment of technologies for quantum communication.

Quantum Information Science

Scott Glancy
Emanuel Knill
Sae Woo Nam (NIST PML)
Konrad Lehnert (NIST PML)
Tasshi Dennis (NIST PML)
Ezad Shojaee (University of Colorado)
Arik Agavayan (University of Colorado)
Mohammad Alhejji (University of Colorado)
Shawn Geller (University of Colorado)
Alex Kwiatkowski (University of Colorado)
Karl Mayer (University of Colorado)
Curtis Rau (University of Colorado)
Akira Kyle (University of Colorado)
Lynden K. Shalm (University of Colorado)
Yanbao Zhang (NTT Corporation)
Gerardo Ortiz (Indiana University)
James R. van Meter (Palo Alto)
Lucas Kocia (Sandia National Laboratory)

The last few years have seen a rapid expansion of public and commercial efforts to develop quantum computing and communication technology. A significant amount of work remains before real-life advantages of digital quantum technologies can become commonplace, as most such advantages cannot be realized without quantum resources orders of magnitude beyond what can be realized in the laboratory today. In the meantime, quantum phenomena are exploited in the design of many modern devices, for precision measurements, and in niche communication settings.

ACMD research covers a broad range of problems in quantum information science, from applications to foundations. Projects covered in this section include the use of entangled photon pairs to generate device-independent randomness, studies of transduction between microwave and optical qubits in quantum networks, statistical analysis to determine confidence intervals for quantum measurements with guaranteed coverage, and fundamental studies of field-amplification by perturbing curved space-time.

Randomness Generation. One form of device-independent randomness generation involves the use of untrusted, black-box devices controlled by deterministic digital timers and computers. It takes advantage of an initial public random seed to generated trusted, secret random bits. One type of untrusted device that can be used for this purpose is based on the entangled quantum systems employed in the loophole-free Bell tests such as the one first demonstrated by NIST researchers in 2015. Protocols that generate device-independent randomness can be inefficient and difficult to verify. ACMD researchers have developed the most data-efficient protocols [1] and helped implement them in the setting of the NIST loophole-free Bell tests operated by NIST PML researchers. This resulted in a first demonstration of true randomness expansion, where more random bits are produced than seed-bits are consumed [2]. NIST PML researchers are currently upgrading their equipment to increase the rate and efficiency of Bell tests and to support public randomness beacons. These setups will use streamlined protocols for generating randomness.

Quantum Networks. Networks for connecting quantum computers require conversion between stationary qubits of the computers and flying qubits of the network. Flying qubits are best supported by optical photons traveling in optical fiber or on direct, free-space links.

ACMD researchers are participating in projects that aim to accomplish this conversion between stationary qubits supported by ions and by superconducting elements. The latter requires converting microwave to optical photons coherently. ACMD researchers determined parameters under which it is possible to accomplish such conversion with entanglement swapping with Gaussian measurements, which are the simplest to implement. A promising conversion scheme uses electro-opto-mechanical devices (see Figure 79), and we engaged in a detailed study of the device performance required to generate the entanglement needed for
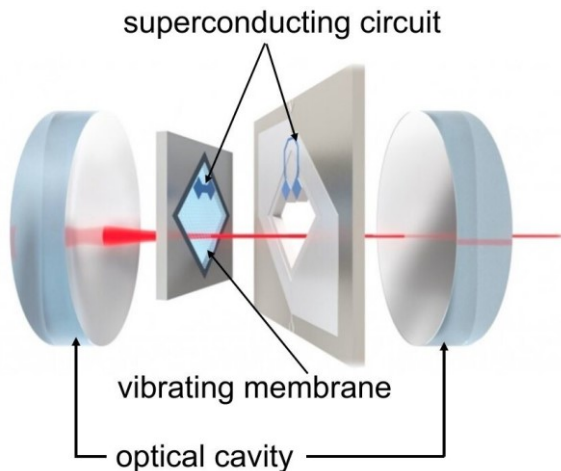
Figure 79. _Schematic of an opto-electro-mechanical device designed for transduction of quantum information residing in the superconducting circuit to optical quantum information carried by light (shown in red) via a vibrating membrane cooled to near its ground state of motion._

communication [3]. These studies are being extended to other methods for entanglement swapping mediated by optical, mechanical, or electrical resonators.

Confidence Intervals. Many experiments and protocols in quantum information science require verification of performance based on measured or otherwise observed quantities. Most applicable statistical tools cannot guarantee the claims sufficiently strongly, make independence assumptions that are observably not satisfied, or are inefficient in their use of data. ACMD researchers pioneered robust methods with guaranteed performance for the loophole-free Bell tests and for the randomness generation experiments. They are now extending their methods for estimating time-averages of performance quantities with confidence intervals at all confidence levels including confidence levels extremely close to 1. The methods can be applied to the situation where one needs to verify the quality of a large number of entangled pairs by measuring only a few of them.

Quantum Amplification on Curved Space-Time. Relativistic quantum information theory deals with the effects of special and general relativity on quantum measurements, information processing and communication. ACMD researchers are participating in investigations of the behavior and measurability of quantum fields on curved space-time backgrounds. To enable measurability of vacuum entanglement in causally separated regions they are generalizing homodyne measurement of quadratures to extremely broad-band measurements using calorimeters (which measure the total energy) instead of photodiodes (which measure the number of photons). Another question of interest is the ability of programmed space-time perturbations to amplify electromagnetic waves. Such amplification applied to vacuum can in principle yield usefully entangled

states. While there is no practical way to program such perturbations, the question is of fundamental interest and its answers may find applications in so-called "analog gravity" systems such as those based on Bose-Einstein condensates.

[1] E. Knill, Y. Zhang, and P. Bierhorst. Generation of Quantum Randomness by Probability Estimation with Classical Side Information. _Physical Review Research_ **2** (2020), 033465. DOI: 10.1103/PhysRevResearch.2.033465

[2] L. K. Shalm, Y. Zhang, J. C. Bienfang, C. Schlager, M. J. Stevens, M. D. Mazurek, C. Abellán, W. Amaya, M. W. Mitchell, M. A. Alhejji, H. Fu, J. Ornstein, R. P. Mirin, S. W. Nam, and E. Knill. Device-independent Randomness Expansion with Entangled Photons. _Nature Physics_ **17** (2021), 452-456. DOI: 10.1038/s41567-020-01153-4

[3] C. L. Rau, A. Kyle, A. Kwiatkowski, E. Shojaee, J. D. Teufel, K. W. Lehnert, and T. Dennis. Entanglement Thresholds of Doubly Parametric Quantum Transducers. arXiv preprint arxiv:2110.10235 (2021). URL: https://arxiv.org/abs/2110.10235

Approximating the Two-mode Two-photon Rabi Hamiltonian

Victor V. Albert
David Wu (Caltech)

The Rabi model is a simple quantum system modeling a ubiquitous interaction between light and matter that is behind many devices such as quantum switches and Josephson junctions. Therefore, understanding this model and its simplest extensions is useful for fruitful developments in these areas of physics.

We consider a two-mode two-photon variant of the Rabi model, which describes the interaction between a two-level system and two radiation modes. We derive approximate eigen-energies of this model using the symmetric generalized rotating wave approximation (S-GRWA). Designed with the symmetry of the system in mind, this method obtains simplified closed-form approximate eigen-energies of the system as functions of the coupling strength between the two photons. We verify our approximated eigen-energies against exact numerical eigen-energies of the system and find that our approximated eigen-energies are accurate in both the strong coupling and the ultra-strong coupling regime of the system. This work has now been published [1] and is the result of efforts by an undergraduate with minimal interference from his adviser.

[1] D. H. Wu and V. V. Albert. Approximating the Two-mode Two-photon Rabi Model. _Physics Letters A_ **422** (2022), 127779. DOI: 10.1016/j.physleta.2021.127779

Quantum Characterization Theory and Applications

Scott Glancy
Emanuel Knill
David T. C. Allcock (NIST PML)
José Aumentado (NIST PML)
Katarina Cicak (NIST PML)
Thomas Gerrits
Dietrich Leibfried (NIST PML)
Adriana E. Lita (NIST PML)
Richard P. Mirin (NIST PML)
Sae Woo Nam (NIST PML)
Raymond W. Simmonds (NIST PML)
Daniel Slichter (NIST PML)
John D. Teufel (NIST PML)
Andrew Wilson (NIST PML)
Arik Avagyan (University of Colorado)
Italo Bezerra (Federal University of Ceará)
Peter Bierhorst (University of New Orleans)
Shaun C. Burd (University of Colorado)
Daniel C. Cole (University of Colorado)
Stephen D. Erickson (University of Colorado)
Shawn Geller (University of Colorado)
Mingyuan Hong (Louisiana State University)
Pan-Yu Hou (University of Colorado)
Hannah Knaack (University of Colorado)
Steve Kolthammer (Imperial College London)
Shlomi Kotler (University of Colorado)
Alex Kwiatkowski (University of Colorado)
Florent Lecocq (University of Colorado)
Omar S. Magaña-Loaiza (Louisiana State University)
Gabriel A. Peterson (University of Colorado)
Ezad Shojaee (University of Colorado)
Raghavendra Srinivas (University of Colorado)
R. T. Sutherland (Lawrence Livermore National Lab)
Hilma Vasconcelos (Federal University of Ceará)
David J. Wineland (University of Oregon)
Jenny J. Wu (University of Colorado)
Chenglong You (Louisiana State University)

Many emerging technologies will exploit quantum mechanical effects to enhance metrology, computation, and communication. Developing these technologies requires improved methods to characterize the performance of quantum devices. This characterization requires solving statistical problems such as estimating an underlying quantum state, measurement, or process by using a collection of measurements made on the quantum system. Alternatively, one may also want to estimate figures-of-merit such as fidelity, error rates, and entanglement measures from that data. Accurate quantum characterization allows experimentalists to answer questions like "What is happening in my quantum experiment?" or "How well will my system perform some quantum information protocol?" and to characterize uncertainty in that answer.
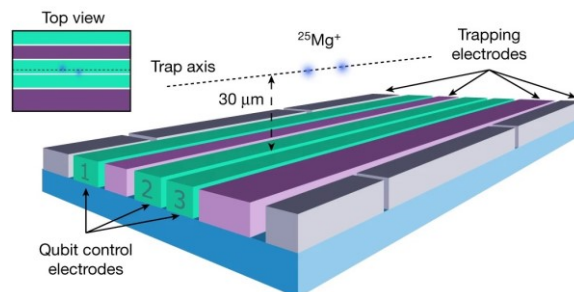
Figure 80. *3D perspective view of central region of surface electrode ion trap (not to scale). For this experiment, two $^{25}Mg^+$ ions are trapped 30 μm above the trap surface. MHz and GHz currents used to generate the entangling interaction are driven along the qubit control electrodes (green), while oscillating (purple) and static voltages (gray) are applied to the trapping electrodes to create the confining potential.*

NIST's Ion Storage Group has pioneered one of the world's most successful quantum computer development projects. To keep pace with their recent advances in qubit preparation, logical operation, and measurement fidelities, ACMD researchers are developing more advanced statistical techniques to characterize trapped-ion quantum computers. An ion qubit is measured by counting the number of photons the ion releases when illuminated by a fluorescent laser. A random number of photons will be produced, but the probability distribution depends on the ion's state. Distinguishing the $|0\rangle$ and $|1\rangle$ states requires differentiating between the two, slightly overlapping, probability distributions. Also, the ion might change its state during the measurement. ACMD researchers have developed a hidden Markov model of the fluorescent measurement process, a strategy for manipulating the ion during the measurement based on the incoming photon detections, and a Bayesian system for inferring the hidden Markov model's parameters and the ion's state at the start of the measurement process. These tools enable higher fidelity qubit measurement and a better understanding of measurement error [1]. ACMD researchers also contributed to an experiment demonstrating another technique to improve ion qubit measurements. In this experiment, a single beryllium qubit is measured repeatedly by using quantum logic gates to transfer its state to a magnesium ion, which can be measured without disturbing the state of the beryllium qubit. ACMD researchers developed a probabilistic model of the repeated measurements and determined that the measurement errors were in the 68 % confidence intervals $[0.6, 2.3] \times 10^{-4}$ and $[0, 1.9] \times 10^{-5}$ for the two qubit states [2].

One difficulty with scaling-up trapped-ion quantum computers is the management of lasers used to apply logic gates. To overcome this difficulty, the Ion Storage Group has developed techniques to apply all logic gates using only microwave fields (see Figure 80). These microwave logic gates were used in an experiment
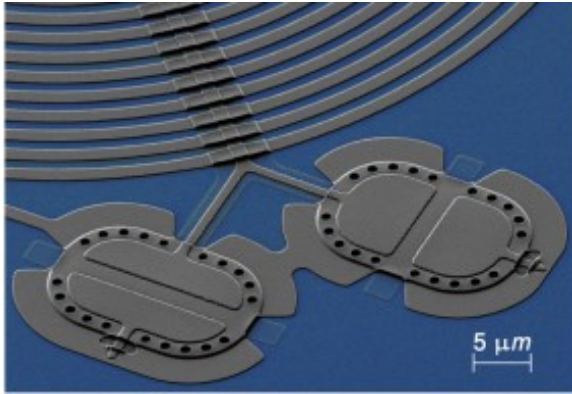
Figure 81. *Scanning electron micrograph of drum. Two aluminum drums are suspended above a sapphire substrate (false colored in blue) resulting in well-defined harmonic modes in the direction perpendicular to the substrate. Each drum forms the top plate of a capacitor, along with the bottom plate which is fixed to the substrate. The oscillations of the drum are entangled through their interactions with a microwave-frequency LC circuit.*

demonstrating universal quantum logic operations on two qubits. ACMD researchers designed a strategy for estimating the fidelity with which the microwave gates could entangle the qubits. That fidelity is in the interval [0.9983, 1] with 68 % confidence [3], which rivals the best fidelities achieved with lasers.

Inspired by the needs of the Metrology with Interacting Photons Innovations in Measurement Science program, ACMD researchers have developed a generalization of traditional homodyne detection. Traditional homodyne detection requires a strong reference beam and can only estimate the state in the mode matching the reference beam, but our generalization can use a weak reference beam (as required for some integrated circuit designs) and can learn about correlations between the signal mode and photons in nearby modes [4]. ACMD researchers partnered with PML to demonstrate this measurement strategy in an optical experiment. However, our analysis revealed subtle interference errors in the optics. We are now designing a new version of the experiment that will avoid those errors.

A basic resource state for continuous-variable quantum information processing is a low-temperature squeezed state. Such states are often prepared in optical modes or in the motion of trapped ions. ACMD researchers have developed a method to estimate the amount of squeezing, temperature, and other parameters of a multi-mode squeezed state from measurements of the number of photons or phonons in the state [5,6]. They are using this method to study the squeezing of ion motion, which was previously characterized assuming that its temperature was 0.

One important test of quantum theory is verification of entanglement between larger and more massive objects. Also, macroscopic oscillators may be useful as quantum memories or as transducers of quantum information. ACMD researchers contributed to a PML

experiment that demonstrated entanglement between two "drumhead" oscillators with masses of 70 pg each (see Figure 81). Their analysis showed that significant entanglement is present and directly observable, even without correcting for measurement inefficiency [7]. *Physics World* cited this accomplishment as the top physics breakthrough of 2021.

Entangled photons can be used to measure optical phases in an interferometer with higher precision per photon than can be achieved without entanglement. ACMD researchers contributed to an experiment demonstrating such quantum-enhanced phase measurement, which used NIST's high-efficiency photon number resolving detectors to achieve quantum enhancement of the measurement precision over a wide range of phases [8].

[1] S. Geller, S. Glancy, and E. Knill. Improving Quantum State Detection with Adaptive Sequential Observations. In preparation.

[2] S. D. Erickson, J. J. Wu, P.-Y. Hou, D. C. Cole, S. Geller, A. Kwiatkowski, S. Glancy, E. Knill, D. H. Slichter, A. C. Wilson, and D. Leibfried. High-Fidelity Indirect Readout of Trapped-Ion Hyperfine Qubits. Preprint arXiv:2112.06341 (2021).

[3] R. Srinivas, S. C. Burd, H. M. Knaack, R. T. Sutherland, Kwiatkowski, S. Glancy, E. Knill, D. J. Wineland, D. Leibfried, A. C. Wilson, D. T. C. Allcock, and D. H. Slichter. High-Fidelity Laser-Free Entanglement and Addressing of Trapped Ion Qubits. *Nature* **597** (2021), 209. DOI: /10.1038/s41586-021-03809-4.

[4] A. Avagyan, E. Knill, S. Glancy, and H. Vasconcelos. State Tomography with Photon Counting after a Beam Splitter. In preparation.

[5] I. Bezerra, H. Vasconcelos, and S. Glancy. Estimation of Squeezing and Temperature from Fock Number Measurements. Preprint arXiv:2111.04923, (2021).

[6] A. Avagyan, E. Knill, and S. Glancy. Multi-Mode Gaussian State Analysis with Photon Counting. In preparation.

[7] S. Kotler, G. A. Peterson, E. Shojaee, F. Lecocq, K. Cicak, A. Kwiatkowski, S. Geller, S. Glancy, E. Knill, R. W. Simmonds, J. Aumentado, and J. D. Teufel. Direct Observation of Deterministic Macroscopic Entanglement. *Science* **372** (2021), 622. DOI: 10.1126/science.abf2998

[8] C. You, M. Hong, P. Bierhorst, A. E. Lita, S. Glancy, S. Kolthammer, E. Knill, S. W. Nam, R. P. Mirin, O. S. Magana-Loaiza, T. Gerrits. Scalable Multiphoton Quantum Metrology with Neither Pre- nor Post-Selected Measurements *Applied Physics Reviews* **8** (2021), 041406. DOI: 10.1063/5.0063294

Correlated Noise in Quantum Devices

Yi-Kai Liu
Mohammad Hafezi (University of Maryland)
Kaixin Huang (University of Maryland)
Alireza Seif (University of Maryland)

Noise and decoherence are an important concern in the development of quantum devices for metrology and computation. While this topic has been studied for many years, it has recently gained new significance, as experimentalists start to build so-called NISQ (noisy intermediate-scale quantum) devices, which have tens to hundreds of qubits, with noise levels that are low but not insignificant. Two important aspects of this problem are the role of spatial correlations between qubits, which can become increasingly complicated when there are many qubits, and the role of non-Markovian memory effects, which are becoming more important as experimentalists build qubits with longer and longer coherence times.

We are developing various methods for modeling these kinds of effects and fitting these noise models to experimental data. First, we have developed a model of correlated dephasing noise on many qubits, where the correlations can be long-range, but are assumed to be sparse (in the sense that they are described by a sparse graph, which is not known *a priori*). This model is applicable to physical systems where two or more qubits can be coupled to the same environmental degree of freedom, and where these couplings are caused by defects or anomalies that occur at random, according to a heavy-tailed probability distribution. We have developed an improved method for fitting this model to experimentally accessible data, by performing measurements using entangled GHZ states on random subsets of the qubits and combining techniques from Ramsey spectroscopy and compressed sensing [1, 2].

Second, we are developing more efficient spectroscopy methods for characterizing non-Markovian noise. We have developed a general method for constructing random pulse sequences, with carefully controlled correlations between the pulses, which can be used to measure arbitrary linear functionals of the power spectral density of the noise. This method is relevant to physical systems such as superconducting qubits and nitrogen vacancy centers. It has several applications, such as characterizing noise spectra that consist of a few isolated peaks, estimating the parameters of noise spectra that have a Lorentzian shape, and learning effective representations of the environment (such as one-dimensional chain representations of a bosonic bath). During the past year, we have investigated additional applications, including continuous monitoring of the total noise strength, and real-time quantum control using spectator qubits [3].

[1] A. Seif, M. Hafezi and Y.-K. Liu. Compressed Sensing Measurement of Long-Range Correlated Noise. NIST Invention Disclosure, NIST Docket #21-020, February 19, 2021.

[2] A. Seif, M. Hafezi, and Y.-K. Liu. Compressed Sensing Measurement of Long-Range Correlated Noise. Preprint, arXiv:2105.12589. URL: https://arxiv.org/abs/2105.12589

[3] K. Huang, A. Seif, M. Hafezi and Y.-K. Liu. Random Pulse Sequences for Qubit Noise Spectroscopy. In preparation.

Time-Energy Uncertainty Relation for Noisy Quantum Metrology

Victor V. Albert
Philippe Faist (Freie Universität Berlin)
Mischa P. Woods (ETH Zurich)
Joseph M. Renes (ETH Zurich)
Jens Eisert (Freie Universität Berlin)
John Preskill (Caltech)

Quantum mechanics places fundamental limits on how well we can measure a physical quantity when using a quantum system as a probe. The most famous tradeoff is the Heisenberg uncertainty relation between the precision with which one can measure an object's position and momentum. The more precise the measurement of one of these "conjugate" variables, the less information one is able to extract about the other.

Another tradeoff often encountered in quantum error correction is the tradeoff between the effect of an observer and the uncontrolled environment on a quantum system. Typically, the less information the environment is able to extract from a quantum system, the more will be preserved for a later observer. One can think of the environment as a malevolent observer, whose clandestine probing of the system obscures the useful quantum information stored within the system. In this work, we introduce and study a fundamental tradeoff relating the amount that noise reduces the accuracy of a quantum clock to the amount of information about the energy of the clock that leaks to the environment.

Specifically, we consider an idealized scenario in which a party Alice prepares an initial pure state of the clock, allows the clock to evolve for a time that is not precisely known, and then transmits the clock through a noisy channel to a party Bob. Meanwhile, the environment (Eve) receives any information about the clock that is lost during transmission. We prove that Bob's loss of quantum Fisher information about the elapsed time is equal to Eve's gain of quantum Fisher information about a complementary energy parameter. We also prove a similar, but more general, trade-off that applies when Bob and Eve wish to estimate the values of parameters associated with two non-commuting observables. We

derive necessary and sufficient conditions for the accuracy of the clock to be unaffected by the noise, which form a subset of the Knill-Laflamme error-correction conditions.

A state and its local time-evolution direction, if they satisfy these conditions, form a metrological error-correcting code. We provide a scheme to construct metrological codes in the stabilizer formalism. We show that there are metrological codes that cannot be written as a quantum error-correcting code in which the Hamiltonian acts as a logical operator, potentially offering new schemes for constructing states that do not lose any sensitivity upon application of a noisy channel.

In Search of New Topological Defects for Robust Quantum Computation

Victor V. Albert
David Aasen (Microsoft and UC Santa Barbara)
Wenqing Xu (University of Maryland)
Wenjie Ji (UC Santa Barbara)
John Preskill (Caltech)
Jason Alicea (Caltech)

Topological quantum computation, in contrast to the many other available blueprints for a quantum computer, holds the promise of computation in an inherently robust fashion. A paradigmatic example of such physics in 1D systems is the Ising quantum spin chain. With its edge, spin, and fermionic contexts, the Ising model yields a web of connections that brings topological physics to a broad array of seemingly unrelated systems. When viewed as a fundamentally fermionic system (via the Jordan-Wigner transformation), the model admits a topological phase hosting unpaired Majorana zero modes [1] that underlie non-Abelian statistics and provide the foundation for topological qubits resilient to local noise. When viewed as a standalone effective theory for edges of the toric code [2], the Majorana zero modes correspond either to anyonic-excitation operators (a.k.a. ribbons) lying on the edge or "defects" between different edge Hamiltonians. Continuum versions of the model, the (free) boson and its fermionic equivalent, the quantum wire, have shined light on how to construct Majorana modes out of emergent degrees of freedom in electronic systems.

The most straightforward extension of the toric code (associated with spins taking values in the order-two group Z_2) is its Z_N version, which the Ising-like clock model has analogously provided connections between the toric code and various electronic systems that have witnessed encouraging recent experimental progress. Such connections, however, are less clear for the more intricate "nonabelian" generalizations of the toric

code, quantum double models [2], defined on lattices whose spins take values in a nonabelian finite group G. Such systems house excitations described by non-Abelian anyons, whose fusion can yield several outcomes (as opposed to just one in the abelian case). Information can be encoded in a robust fashion directly into superpositions of the fusion outcomes, and universal computation can be implemented with potentially less overhead than conventional schemes (e.g., without magic-state distillation). In contrast to the conventional toric code, a standalone 1D effective theory describing the edge of a general quantum double model and its continuum description have yet to be examined.

Recently, Munk, Rasmussen, and Burrello [3] introduced a G-site spin chain called the "flux ladder" and developed a corresponding generalization of the Jordan-Wigner mapping. However, direct connections of the flux ladder and the generalized mapping to quantum doubles remained unclear. In this work, we fill in this gap by deriving a standalone Ising-like model for a general quantum double edge and showing that this model is a lightly generalized flux ladder.

Armed with this model, we map out [4] a web of connections for quantum double models, generalizing the existing web for the Z_N toric code. In a mapping distinct from that of [1], we develop Jordan-Wigner-like mode operators that are directly related to anionic ribbon operators of the quantum double model. We obtain a continuum description of the general flux ladder and recast a particular case in terms of fermions with the help of non-Abelian bosonization. While our extensions arise from viewing the Z_N case, through the mathematical lens of representation theory the resulting physical connections to established systems should nevertheless brighten the prospects of eventual realization of quantum-double topological phases.

[1] A. Yu. Kitaev. Unpaired Majorana Fermions in Quantum wires. *Physics-Uspekhi* **44** (2001), 131. DOI:10.1070/1063-7869/44/10S/S29

[2] A. Yu. Kitaev. Fault-tolerant Quantum Computation by Anyons. *Annals of Physics* **303** (2003), 2. DOI:10.1016/S0003-4916(02)00018-0

[3] M. I. K. Munk, A. Rasmussen and M. Burrello. Dyonic Zero-energy Modes. *Physical Review B* **98** (2018), 245135. DOI:10.1103/PhysRevB.98.245135

[4] V. V. Albert, D. Aasen, W. Xu, W. Ji, J. Alicea, and J. Preskill. Spin Chains, Defects, and Quantum Wires for the Quantum-Double Edge. Preprint, 2021. URL: arXiv:2111.12096

Provable Accurate Machine Learning Algorithms for the Quantum Many-body Problem

Victor V. Albert
Hsin-Yuan Huang (Caltech)
Richard Kueng (Johannes Kepler University Linz)
Giacomo Torlai (Amazon AWS)
John Preskill (Caltech)

Solving quantum many-body problems, such as finding ground states of quantum systems, has far-reaching consequences for physics, materials science, and chemistry. Classical machine learning (ML) has emerged as a powerful approach to solving such problems. However, the advantages of ML over more traditional methods have not been firmly established, reflecting the relative paucity of rigorous theory in ML.

On the other hand, engineered quantum devices are believed to efficiently simulate quantum systems of interest, thereby helping solve many-body problems. Today's noisy intermediate-scale devices do not have the capabilities to simulate a system in perfect error-corrected fashion, but nevertheless should encode useful bits of data about properties of interesting many-body systems. However, since many-body states require exponential (in the number of qubits) memory to be stored exactly, it is unclear how to distill and utilize useful bits from noisy experimental data even at intermediate scale.

Our work [1] combines the best of both the ML and quantum worlds, distilling quantum experimental data and utilizing it with provably efficient ML algorithms. We prove that classical ML algorithms can efficiently reveal properties of quantum many-body states associated with physical systems. In order to circumvent the exponential memory requirement, we devise a way to feed memory-efficient classical snapshots [2] of many-body states — obtained either from experiment or from another classical device — into ML algorithms. We introduce and numerically test ML algorithms for classifying and extrapolating properties of many-body systems and derive information-theoretic bounds on their efficiency.

Viewed from a broader perspective, by illustrating how experimental data can be exploited to make accurate predictions about features of quantum systems that have never been studied directly, our work exemplifies a potentially powerful methodology for advancing the physical sciences. With further theoretical developments, perhaps we can learn how to use experimental data that is already routinely available to accelerate the discovery of new chemical compounds and materials with remarkable properties that could benefit humanity. This work was accepted as a plenary talk at QIP 2022, the most prestigious conference in quantum information science.

[1] H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill. Provably Efficient Machine Learning for Quantum Many-Body Problems. Preprint, 2021. URL: arXiv:2106.12627

[2] H.-Y. Huang, R. Kueng, and J. Preskill. Predicting Many Properties of a Quantum System from Very Few Measurements. *Nature Physics* **16** (2020), 1050. DOI:10.1038/s41567-020-0932-7

Provably Accurate Quantum Simulation of Gauge Theories and Bosonic Systems

Victor V. Albert
Yu Tong (Google and U. California at Berkeley)
Jarrod R. McClean (Google)
John Preskill (Caltech)
Yuan Su (Google)

Many quantum many-body systems of interest consist of lattices of particles that can occupy states of arbitrarily high energy. Their infinite-dimensional local Hilbert spaces must be truncated in order to perform simulations of real-time dynamics on classical or even quantum computers. To analyze errors resulting from such truncation, we develop [1] methods for bounding the rate of growth of local quantum numbers such as the occupation number of a mode at a lattice site, or the electric field at a lattice link. We show that, if these models are truncated by imposing an upper limit Λ on each local quantum number (analogue of energy), a truncation error no worse than ϵ can be achieved by choosing Λ to increase polylogarithmically with ϵ^{-1}, an exponential improvement over previous rigorous bounds.

Although formally the local Hilbert spaces are infinite-dimensional in the models we considered, our results show that, at least for some purposes, these models can be accurately approximated by models with finite-dimensional local Hilbert spaces of relatively modest size. Many fundamental results have been derived for quantum spin systems with finite-dimensional spins on each lattice site, and perhaps the tools we have developed can be exploited to extend some of these results to systems with infinite-dimensional local degrees of freedom. This work was accepted as a talk at QIP 2022, the most prestigious conference in quantum information science.

[1] Y. Tong, V. V. Albert, J. R. McClean, J. Preskill, and Y. Su. Provably Accurate Simulation of Gauge Theories and Bosonic Systems. Preprint, 2021. URL: arXiv:2110.06942

Standards for Characterizing Quantum Phases

Victor V. Albert

Today's quantum devices are capable of producing exotic quantum many-body states, but their verification is problematic due to several challenges:

1. the exponential memory requirement for perfect storage of a quantum state,

2. a paucity of local observables ("order parameters") whose expectation values characterize the zoo of quantum phases of matter, and

3. a lack of rigorous guarantees on the ability and efficiency of using existing order parameters to characterize said phases.

I have begun a multi-prong effort to resolve these three challenges. The first challenge can be tackled using efficient classical snapshots of a quantum state called *classical shadows* [1]. Armed with such classical descriptions, the goal now is to develop methods to extract from such descriptions the coarse-grained phase properties of their parent quantum states. We have made some progress along this direction with the help of machine learning algorithms [2], but an intuitive understanding of such processes is yet to be fleshed out.

My collaborators and I have made progress on the second issue, concocting an invariant that helps characterize properties of a subset of "topological" many-body states with robust edge excitations [3, 4]. The electronic quantum system exhibiting the quantum Hall effect is an example of such a phase, and its "topologically protected" edge currents have been used to determine values of fundamental physical constants. Robust edge excitations are characterized by a topological invariant called the chiral central charge, and, until our work, it was not known how to extract this quantity from a single copy of a physical many-body wavefunction. This work was accepted as a single-track talk at QIP 2022, the most prestigious conference in quantum information science.

I have also begun a journal club discussion at the University of Maryland College Park on the difficult third issue, bringing together experts from quantum computer science and topological quantum phases of matter.

[1] H.-Y. Huang, R. Kueng, and J. Preskill. Predicting Many Properties of a Quantum System from Very Few Measurements. *Nature Physics* **16** (2020), 1050. DOI:10.1038/s41567-020-0932-7

[2] H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill. Provably Efficient Machine Learning for Quantum Many-Body Problems. Preprint, 2021. URL: arXiv:2106.12627

[3] I. H. Kim, B. Shi, K. Kato, and V. V. Albert. Chiral Central Charge from a Single Bulk Wave Function. Preprint, 2021. URL: arXiv:2110.06932

[4] I. H. Kim, B. Shi, K. Kato, and V. V. Albert. Modular Commutator in Gapped Quantum Many-Body Systems. Preprint, 2021. URL: arXiv:2110.10400

Analog Quantum Algorithms

Lucas Brady
Timothy Mooney (George Mason University)
Alexey Gorshkov (NIST PML)
Aniruddha Bapat (University of Maryland)
Przemyslaw Bienias (University of Maryland)
Jacob Bringewatt (University of Maryland)
Yaroslav Kharkov (University of Maryland)
Lucas Kocia (Sandia National Labs)
Jeffrey Larson (Argonne National Labs)
Sven Leyffer (Argonne National Labs)
Xinyu Fei (University of Michigan)
Siqian Shen (University of Michigan)

As quantum computers and quantum technologies begin to be realized and become more powerful, it is important to develop algorithms that can exploit these small devices to best utilize any quantum advantage. Such devices are usually noisy and prone to errors, making it difficult to implement many circuit-based quantum algorithms that require a specific sequence of quantum operations. Another class of quantum algorithms that is thought to be more suitable to these small noisy devices is analog quantum algorithms. Rather than being described in terms of discrete quantum operations or gates, these algorithms are described in terms of how the system or Hamiltonian evolve with time.

Variational Quantum Algorithms. One of the most promising algorithmic approaches to utilizing small quantum devices involves applying a variational approach. In this approach, we use classical computation to pick some set of parameters for a quantum calculation or state preparation, using the small quantum computer to sample from the distribution of outputs that results from those parameters. This output is then used to update the classical algorithm, leading to a loop until a desired problem is solved or the state is prepared.

For instance, one of the most prominent quantum variational algorithms is the Quantum Approximate Optimization Algorithm (QAOA) [1]. In this algorithm, a system of quantum bits, or qubits, are exposed to two alternating Hamiltonians. These Hamiltonians are applied for p alternations, with the times for each pulse being fed to the classical computer as its variational parameters. The goal is to approximately prepare the ground state of one of the Hamiltonians (or some other desired state), and the output of sampling the quantum computer for a given parameter set is used in a classical loop to optimize the energy of the produced state (or another desired quantity).

Another important class of analog quantum algorithms is Quantum Adiabatic Optimization (QAO) [2] or its slightly more generalized form, Quantum Annealing (QA). In this algorithm, the system is initialized in the ground state of some initial Hamiltonian. If the Hamiltonian changes slowly enough, the quantum adiabatic theorem says that the system will stay in the ground state, allowing us to find the ground state of desired Hamiltonians that might encode some classical problem.

In previous work [3], we examined this class of algorithms through the lens of optimal control theory. QAOA and Quantum Annealing both use the exact same ingredients and can be compared on equal footing. In that work, we found that the optimal procedure had a hybrid bang-anneal-bang form, with pulsed structure at the beginning and end and a smooth annealing-like region in the middle. In the last year, we followed up this work with a longer paper [4] that examined where this structure comes from.

In [4] we found that the annealing region in a bang-anneal-bang curve has a set pattern as the time allocated to the procedure increases. Namely this region resembles a smooth monotonic curve that is independent of runtime along with a superposed oscillatory function which has an amplitude that decreases as runtime increases. We showed that the period of oscillations is related to maintaining adiabaticity and depend on the spectral gap of the underlying annealing problem in the long time limit. Furthermore, QAOA seeks to emulate this optimal protocol by matching up its bangs to the oscillations. The exact mechanisms of this matching are still under investigation.

This work with bang-anneal-bang schedules also led to the creation of a novel algorithm that seeks to create a good parameterization of the optimal schedule for use with a variational quantum computer. This parameterization captures many of the features of the optimal schedule and is empirically able to outperform QAOA and more naive annealing schedules. An invention disclosure was submitted for this, and a provisional patent application has been filed.

There is a large body of work on classical optimization strategies. The purpose of quantum optimization is to harness the power of quantum computing, but as we have described, many of the near-term quantum algorithms are themselves variational, requiring classical optimization. To this end, we have started a collaboration with some classical optimization specialists from Argonne National Lab to explore applications of these classical schemes to improving the performance of near-term quantum algorithms. As of the end of 2021, this work was in draft [5] with follow-up work ongoing.

Quantum-inspired Classical Algorithms. Quantum algorithms provide great speed-ups over existing classical algorithms, but a large part of ongoing research involved how to translate them over into better classical algorithms. This has led to a large field of quantum-inspired classical algorithms, foremost among these being Quantum Monte Carlo techniques. These involve simulation of quantum systems via a translation of the quantum system into a classical system and then Monte Carlo simulation of that classical system. The problem is that this translation often leads to quasi-probability distributions that are positive and negative and highly oscillatory, the so-called sign-problem. The sign-problem is hard to solve in general, but there are several partial mitigation techniques, among them complexifying the quasi-probability distribution and deforming the distribution into the complex plane to sit on the hyperplane of stationary phase, the Lefschetz Thimble. This technique had been performed before for high energy physics problems with continuous variables, and our contribution was translating those ideas over to discrete variable spin systems. Our methods worked and produced results which significantly reduced the severity of the sign problem for small test problems [6]

There is a class of Hamiltonians (quantum energy operators, represented by Hermitian matrices) that do not have the sign-problem, known as stoquastic Hamiltonians. In matrix form, these stoquastic Hamiltonians are real and have non-positive off-diagonal entries, which means that this is a basis dependent property. Every Hamiltonian can be stoquasticized by a change of basis (e.g., diagonalization), but finding such a basis change is an NP-hard problem. We asked the simple question of whether it is always possible to find a basis in which two Hamiltonians are simultaneously stoquastic, a question relevant to the simulation of quantum annealing which involves superpositions of two Hamiltonians. This is largely a mathematical question, and the answer turns out to be "no." There exist pairs of Hamiltonians that cannot be made simultaneously stoquastic, and we have developed a simple test which can positively identify some of these pairs (but a failed test does not guarantee simultaneous stoquasticizability). Work is ongoing to find a version of the test which works fully in both directions.

[1] E. Farhi, J. Goldstone, and S. Gutmann. A Quantum Approximate Optimization Algorithm. arXiv:1411.4028.

[2] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. Quantum Computation by Adiabatic Evolution. arXiv:0001106 .

[3] L. T. Brady, C. L. Baldwin, A. Bapat, Y. Kharkov, and A. V. Gorshkov, Optimal Protocols in Quantum Annealing and QAOA Problems. *Physical Review Letters* **126** (2021), 070505.

[4] L. T. Brady, L. Kocia, P. Bienias, A. Bapat, Y. Kharkov, and A. V. Gorshkov Behavior of Analog Quantum Algorithms. Preprint arXiv:2107.01218.

[5] X. Fei, L. T. Brady, J. Larson, S. Leyffer, and S. Shen. Binary Control Pulse Optimization for Quantum Systems. In preparation.

[6] T. C. Mooney, J. Bringewatt, and L. T. Brady. Lefschetz Thimble Quantum Monte Carlo for Spin Systems. Preprint arXiv:2110.10699.

Approximating Output Probabilities of Shallow Quantum Circuits

Matthew Coudron
Nolan J. Coble (University of Maryland)
Suchetan Dontha (U. of Illinois Urbana-Champaign)
Shi Jie Samuel Tan (Haverford College)
Stephen Smith (University of South Carolina)
Sangheon Choi (Rose-Hulman Institute of Technology)

It is believed that quantum computers will provide an exponential computational speed-up over classical machines for certain computational tasks. Exactly which computational tasks have this potential, and how far the near-term quantum hardware can go towards this goal, are two of the biggest open questions about quantum computation today. In this research project we study the classical complexity of simulating certain near-term quantum computational models. We do not use a "brute-force" simulation technique requiring exponential time, but rather, make use of non-trivial structure within the near-term quantum circuits to achieve provable simulation results in sub-exponential time. These results can serve as a litmus test for determining which near-term quantum circuit architectures have the potential for an exponential quantum advantage. If an architecture can be classically simulated efficiently, then it does not have the potential to exhibit an exponential quantum advantage. On the other hand, these novel simulation results may provide some intuition for the conditions under which an exponential advantage might be achieved by quantum circuits.

As our central technical contribution, we (Noble and Coudron) have developed a classical algorithm that, for any 3D geometrically local, constant-depth quantum circuit C, and any bit string $x \in \{0,1\}^n$, can compute the quantity

$$\left| < x \,|\, C \,|\, 0^{\otimes n} > \right|^2$$

to within any inverse-polynomial additive error in quasi-polynomial time. It is known that it is #P-hard to compute this same quantity to within 2^{-n^2} additive error [1]. The previous best-known algorithm for this problem used $O(2^{n^{\frac{1}{3}}} \left(\frac{1}{\epsilon}\right)^2)$ time to compute probabilities to within additive error ϵ [2]. Notably, the [2] paper included an elegant polynomial time algorithm for the same estimation task with 2D circuits, which makes a novel use of 1D matrix product states (MPS) carefully tailored to the 2D geometry of the circuit in question. Surprisingly, it is not clear that it is possible to extend
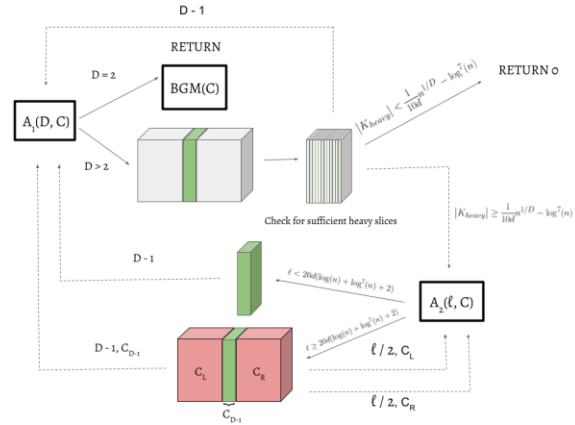
Figure 82. *This diagram illustrates our recursive decomposition procedure for D-dimensional circuits. This was invented for D=3 in [4] as described in this project description and extended to any constant D in the next project description.*

this use of MPS to address the case of 3D circuits in polynomial time. This raises a natural question as to whether the computational complexity of the 3D problem might be drastically higher than that of the 2D problem. In this work we address this question by exhibiting a quasi-polynomial time algorithm for the 3D case. We believe that our algorithm extends naturally to any fixed dimension D by induction on the dimension, but we focus on the 3D case, as the simplest unresolved case, for concreteness. Furthermore, we show that, under a natural, polynomial-time-checkable condition on the circuit C, our algorithm runs in polynomial time. This highlights the possibility that the super-polynomial worst-case time bound on our algorithm might be due to limitations of our analysis. In order to surpass the technical barriers encountered by previously known techniques we are forced to pursue a novel approach: Constructing a recursive sub-division of the given 3D circuit using carefully designed block-encodings.

Our algorithm has a divide-and-conquer structure, demonstrating how to approximate the desired quantity via several instantiations of the same problem type, each involving 3D-local circuits on at most half the number of qubits as the original. This division step is then applied recursively, expressing the original quantity as a weighted sum of smaller and smaller 3D-local quantum circuits. A central technical challenge is to control correlations arising from the entanglement that may exist between the different circuit "pieces" produced this way. We believe that the division step, which makes a novel use of block-encodings [3], together with an inclusion-exclusion style argument to reduce error, may be of interest for future research on low depth quantum circuits.

This original result, done last year, is the foundation for the subsequent results described below. It appeared as a presentation at the Annual Conference on Quantum Information Processing (QIP 2021) and was an accepted

paper at Annual ACM SIGACT Symposium on Theory of Computing (FOCS 2021) [4].

In [5] we extend the results on 3D circuits from [4] to any fixed constant dimension D. This is notable because, historically, the simulation of low-depth and geometrically local quantum circuits has required a new mathematical innovation every time the dimension of the geometric locality is increased. The $D=1$ case is solved using the famous technique of matrix product states (MPS), which is fundamental to the field and has been known for decades. However, the algorithm for estimating output amplitudes in the case $D=2$ was not invented until 2019, and required a novel insight, first appearing in [2]. Finding an algorithm for the $D=3$ case, [4], required a completely different approach, this time departing from the paradigm of MPS altogether, and requiring ≈ 50 pages of mathematics to formalize a divide-and-conquer algorithm. Our new result [5] is significant because it shows that this trend of requiring completely new techniques to extend to the $D+1$ case need not continue. We are able to show that the divide-and-conquer paradigm extends, by an appropriate induction argument on the dimension, to an algorithm that works in quasi-polynomial time for any fixed constant dimension D.

Currently the runtime of our algorithm is quasi-polynomial in the number of qubits, n, but triply exponential in the dimension D. Decreasing the time dependence on D is one of the interesting open problems from our project, showing yet another route to understanding how the computational power of low-depth quantum circuits may depend on their dimensionality. Furthermore, it is a fact that any constant-depth quantum circuit acting on n qubits can be written as a circuit which is geometrically local in dimension $D = n$. Therefore, the more efficient our simulation algorithm becomes as a function of D, the more we learn about the complexity of simulating quantum circuits which have constant depth but are *not* geometrically local. This is a major open problem in the field which would be resolved if the runtime of our algorithm scaled quasi-polynomially in D as well as n. So, there is still a large gap in D-dependence between our result and a resolution to simulating low-depth circuits which are not geometrically local. However, we believe that even improving the dependence to singly exponential in D, which may be within reach, would have complexity-theoretic consequences.

The main observation in our extension from $D=3$ to arbitrary D is that the divide-and-conquer algorithm from [4] relies on the dimension $D=3$ only in the base-case of the recursion, and not at each division step. This, in and of itself, is a sense in which the algorithm of [4] differs significantly from the technique of MPS or the $D=2$ solution of [2]. This observation leads naturally to the idea of extending to arbitrary dimension D via induction on the dimension. That is, given the $D=3$ algorithm of [4], we can define a $D=4$ algorithm using

the very same style of recursive decomposition on the 4D circuit, and then, once the recursively divided sub-problems have been reduced to very low width in the 4th dimension, rendering them "effectively 3D", we can use the existing $D=3$ algorithm to solve those base-cases of the recursion.

This idea takes some work to formalize, particularly in making rigorous how a 4D circuit with small width in the 4th dimension can be regarded as "effectively 3D" (by combining the qubits on a row in the 4th dimension into one large qudit of Hilbert space size which is exponential in the circuit-width in the 4th dimension). Once a 4D algorithm has been established, the same approach can be used to extend, inductively, to a 5D algorithm, and so on. We have formalized this inductive proof, and it gives direct intuition for why the divide-and-conquer algorithmic paradigm extends to any fixed dimension D. This inductive process is depicted in Figure 82.

This extension to D dimensions [5] is joint work with summer student participants in the UMD Research Experience for Undergraduates (REU) on Combinatorics, Algorithms, and AI for Real Problems, and the NSF-Mathematical Sciences Graduate Internship program. A poster describing this work has been accepted for presentation at the 2022 Annual Conference on Quantum Information Processing (QIP 2022).

[1] R. Movassagh. Quantum Supremacy and Random Circuits. Preprint arXiv:1909.06210, 2019.

[2] S. Bravyi, D. Gosset, and R. Movassagh. Classical Algorithms for Quantum Mean Values. *Nature Physics* **17** (2021), 337–341. DOI: 10.1038/s41567-020-01109-8

[3] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe. Quantum Singular Value Transformation and Beyond: Exponential Improvements for Quantum Matrix Arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (STOC 2019), 193–204. DOI: 10.1145/3313276.3316366

[4] N. J. Coble and M. Coudron. Quasi-polynomial Time Approximation of Output Probabilities of Constant-depth, Geometrically-local Quantum Circuits. In *Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science* (FOCS 2021), Online, February 7-10, 2022.

[5] S. Dontha, S. Jie S. Tan, S. Smith, S. Choi, and M. Coudron. Approximating Output Probabilities of Shallow Quantum Circuits which are Geometrically-local in any Fixed Dimension. In review.

Post-Quantum Cryptography

Yi-Kai Liu
Gorjan Alagic (NIST ITL)
Daniel Apon (NIST ITL)
Lily Chen (NIST ITL))
David Cooper (NIST ITL)
Quynh Dang (NIST ITL)
John Kelsey (NIST ITL)
Carl Miller (NIST ITL)
Dustin Moody (NIST ITL)
Rene Peralta (NIST ITL)
Ray Perlner (NIST ITL)
Angela Robinson (NIST ITL)
Daniel Smith-Tone (NIST ITL)

Since 2016, NIST has been engaged in a formal process to develop standards for post-quantum cryptography (PQC). The goal is to standardize one or more cryptosystems that could replace those currently used schemes that are vulnerable to attack by quantum computers. These include RSA, Diffie-Hellman, and elliptic curve cryptosystems, which play a crucial role in electronic commerce and cybersecurity. While large quantum computers have not yet been built, NIST believes it is prudent to begin preparing for that possibility.

NIST is focusing on three main functionalities: public-key encryption, key establishment, and digital signatures. These are fundamental cryptographic primitives that enable a variety of applications, including secure web browsing, digital certificates, and secure software updates.

There are a number of candidate cryptosystems that are believed to be quantum secure. These are based on a variety of mathematical techniques, including high-dimensional lattices, coding theory, systems of multivariate polynomial equations, elliptic curve isogenies, hash-based signatures, and many others. However, further research is needed in order to increase confidence in the security of these schemes, and to improve their practical performance.

In July 2020, NIST began the "third round" of its process for analyzing candidate cryptosystems for standardization. During the third round, NIST is focusing its attention on seven "finalists" (four schemes for public-key encryption and key establishment, and three schemes for digital signatures) that will be considered for standardization at the end of the third round; and eight "alternates" (five schemes for public key encryption and key establishment, and three schemes for digital signatures) that will still be considered for possible standardization at a later date.

During the third round, NIST and the broader cryptography community have made progress on a number of issues, including cryptanalysis of multivariate signature schemes (leading to significant degradation of their security); getting better estimates of the security strength (i.e., the number of "bits of security") of all of the cryptosystems; understanding the possible impact of third-party patents on deployment and adoption of the different cryptosystems; and understanding how the migration to post-quantum cryptography will affect other internet protocols and use cases (such as the TLS protocol, certificate chains, DNS security, and smart cards).

In June 2021, NIST hosted its Third PQC Standardization Conference, to highlight recent research on topics related to the candidate cryptosystems, gather feedback from stakeholders, and facilitate discussions with the broader cryptography community. In addition, in October 2021, the NIST National Cybersecurity Center of Excellence (NCCoE) began a project on Migration to Post-Quantum Cryptography and is currently recruiting industry partners.

In early 2022, NIST expects to end the third round of the evaluation process and select a set of cryptosystems for standardization. NIST may also select some cryptosystems for a fourth round of evaluation. Finally, NIST expects to issue a new "call for proposals" for post-quantum signature schemes only (not public key encryption or key establishment). This is motivated by recent advances in the cryptanalysis of multivariate signature schemes, and the needs of future applications and use cases (such as vehicle-to-vehicle communications).

Joint Center for Quantum Information and Computer Science

Victor Albert
Matthew Coudron
Yi-Kai Liu
Carl Miller (NIST ITL)
Andrew Childs (University of Maryland)

http://quics.umd.edu/

Established in October 2014, the Joint Center for Quantum Information and Computer Science (QuICS) is a cooperative venture of NIST and the University of Maryland (UMD) to promote basic research in understanding how quantum systems can be effectively used to store, transport and process information. QuICS brings together researchers from the University of Maryland Institute for Advanced Computer Studies (UMIACS) and the UMD Department of Physics and Computer Science with NIST's Information Technology and Physical Measurement Laboratories, together with postdocs, students, and a host of visiting scientists.

QuICS has quickly established itself as a premier center for research in quantum information science. Fourteen Fellows, 3 Affiliate Fellows, 21 postdocs, and 59 graduate students are currently associated with the

Figure 83. *Two new Fellows joined QuICS this year. Top: Daniel Gottesman of UMD. Bottom: Nicole Yunger-Halpern of NIST PML. (Photos courtesy of UMD)*

center. Yi-Kai Liu of ACMD is Co-Director of the Center along with Andrew Childs of the UMD Computer Science Department.

Two new Fellows joined the Center in FY 2021. The first, Daniel Gottesman, holds the Brin Family Endowed Professorship in Theoretical Computer Science at UMD. Gottesman has made seminal contributions in several areas, including quantum error correction, fault-tolerant quantum computation, quantum cryptography, and quantum complexity theory. He is best known for developing the stabilizer formalism for quantum codes, and for developing techniques for performing quantum gates using quantum teleportation. The second is Nicole Yunger-Halpern of NIST PML, whose research investigates quantum thermodynamics and its applications in other areas.

The Center is very productive. In CY 2021 more than 125 research papers were produced by those associated with the center. More than 65 seminars by both internal and external speakers were held.

In FY 2021 a formal review of QuICS and its sister center, the Joint Quantum Institute, was undertaken by a distinguished group of external reviewers. Based upon that review, the NIST Cooperative Agreement that supports QuICS was renewed for an additional five years.

Of note this year, QuICS researchers participated in a proposal to the National Science Foundation that resulted in the awarding of an NSF Quantum Leap Challenge Institute for Robust Quantum Simulation. This new institute is led by the University of Maryland together with four partner universities: Duke, Princeton, North Carolina State, and Yale. Nine QuICS Fellows are involved in this effort, which is being led by QuICS co-director Andrew Childs.

Quantum Communications and Networking R&D

Oliver Slattery
Lijun Ma
Thomas Gerrits
Xiao Tang
Anouar Rahmouni
Sumit Bhushan
Paulina Kuo
Ya-Shian Li-Baboud (NIST ITL)
Dhananjay Anand (NIST ITL)
Abdella Battou (NIST CTL)
Alan Migdall (NIST PML)
Sergey Polyakov (NIST PML)
Neil Zimmerman (NIST PML)
Josh Bienfang (NIST PML)
Ivan Burenkov (NIST PML)
Hala (NIST PML)
Ping-Shine Shaw (NIST PML)
Joe Rice (NIST PML)
Rob Densock (NIST OISM)
Qing Li (Carnegie Mellon University)
Leah Ding (American University)

www.nist.gov/quantum

The Quantum Communication and Networks Project develops and studies quantum components, systems, and networks. In addition to continuing our research and development of single photon sources, detectors, interfaces, and quantum memories (outlined in previous reports), we have during this year:

- implemented novel prototypes for light-fiber coupling and light trap detection; explored silicon-carbide (SiC) as a promising photonics platform,

- designed and built tools that will be deployed into quantum networks for characterization of quantum network components and quantum network links,

- begun to build and synchronize nodes on a campus-wide quantum network testbed in collaboration with colleagues from ITL, PML and CTL,

- developed a plan for a regional quantum network testbed (called the DC Quantum Network testbed or DC-QNet) in collaboration with other NIST colleagues and in coordination with a number of local federal agencies, and

- explored emerging quantum application areas such as quantum edge computing.

Components and prototype development. We have developed a nearly polarization-independent, low-cost light *trap detector* targeting 1 % uncertainty optical power measurements from 1100 nm to 1600 nm to simplify and greatly reduce the cost of calibrating single

photon sources and detectors needed for quantum communications and networks (Figure 84). While single-photon sources and detectors exist, their calibration and accurate characterization remains difficult and costly. A common method of calibrating single-photon detectors (SPD) is to use an attenuated laser source whose laser power is accurately known at nanowatt levels. Typically, the accuracy of such

Figure 84. *(a) The prototype trap detector design. (b) 3D design illustration with metallic part removed where the InGaAs photodiode (PDs) are shown in yellow, and the flat mirror is on the underside. Light enters the trap from the FC optical fiber connector. (c) Photo of the prototype trap detector with cover, the FC-connector is visible in the front.*

lasers is established by using a calibrated optical power meter traceable to a primary standard such as a cryogenic radiometer. However, this approach is costly and time-consuming and therefore not available to many users. Our trap detector eliminates the need for a frequent recurring calibration service and can be used for reliable low-cost calibration of quantum systems. Once the detector is successfully calibrated, it can subsequently be used to calibrate compatible sources (see 2020 report).

Micro-resonators have been used for various photonic and quantum processes including single photon sources. In non-integrated micro-resonator devices, coupling the light into and out of the micro-resonator (or micro-ring) has been an ongoing challenge. Typically, the evanescence waves of a tapered (or stretched) fiber are used but this process is very fragile, and the tapered fiber easily breaks. We have designed a whole new type of optical coupler based on a _bended optical fiber_ that is polished to expose its core and while supported by a specifically designed assembly can couple light to and from the micro-resonators in a way that is much more robust, easy to use, precise and reproducible – which will be important for mass deployment in quantum networks. Detailed modeling has shown the feasibility of the approach and we are continuing to experiment with various polishing procedures, coupling configurations and support assembly designs to implement and improve direct fiber to micro-resonator coupling.

Integrated photonics have become sufficiently developed to herald high-performance miniaturization of quantum systems and will be essential for scaling and high-volume production and deployment of quantum systems into the future. Silicon-carbide (SiC) has recently been identified as a particularly promising material for integrated photonic devices due to its uniquely strong second *and* third-order non-linear coefficients and its capability to perform (including as a bright single photon source) over a wide variety of frequencies that are appealing to quantum applications. In addition, SiC device fabrication is compatible with existing CMOS foundry processes and such devices can be integrated on a chip with a small footprint. We have established a new research program exploring _integrated SiC devices_ for various quantum processes including single photon sources and interfaces.

Quantum memory is an important device for quantum communication and networks, and an essential part of quantum repeaters and quantum network nodes. Previously we studied and demonstrated cesium-based quantum memory based on electromagnetically induced transparency (EIT) using an atomic vapor. Currently we are implementing a _quantum memory system based on a 2-dimensional (2D) magneto-optical trap (MOT)_, which is expected to have higher efficiency and longer storage time. During this year, we advanced our design and modeling of the 2D MOT and the EIT scheme based on the 2D MOT.

Quantum Network and Network Component Metrology. Building upon our recently developed calibration chain for single-photon detector efficiency, we developed a baseline tool that can be used to characterize single-photon detector performance in terms of efficiency, dark counts, afterpulsing, etc. [1]. We have expanded this capability to cover the spectral range from 700 nm to 1750 nm, useful for characterization of future quantum networks and related components and have worked with external partners to characterize single-photon detectors. In addition, we developed a rack-mounted single-photon source based on spontaneous parametric downconversion (SPDC) that can be used to characterize single-photon detectors. Further, this source can also serve for distributing polarization entanglement with the flip of a switch. Two polarization entanglement analyzers (receivers) were built to quantify and measure the polarization entanglement. Also, these receivers are rack mountable. Because of the by-design shoe-box size integration into a 48 cm (19 in) rack, we can deploy these sources and receivers anywhere within a quantum network. This source and the receivers will serve as benchmarking setups for the implementation of metrology protocols in quantum network testbeds.

While quantum network technology is still being researched, the associated _quantum network metrology_ will need to mature hand-in-hand with the core quantum communication technologies. Quantum networks will require methods for measuring the performance of components and for efficient entanglement routing. In current optical communication systems, development of new telemetry and control protocols have been essential
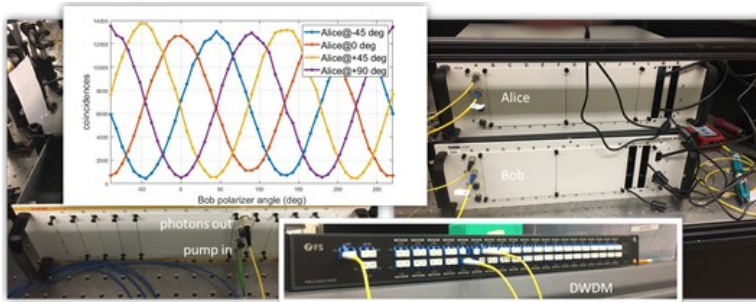
Figure 85. *First measurement of polarization entanglement (graph) from rack-mount source (bottom left) using the rack-mount receivers (right). Polarization entanglement is achieved by generation of spectrally entangled type-II SPDC and subsequent ITU channel separation by use of a DWDM (bottom right).*

to the success of each new generation of commercial deployments. Quantum systems have even more challenging performance requirements, and advanced metrology must play a central role in any future standards and deployments. In collaboration with our colleagues in CTL and PML, we are developing tools and protocols based on classical and quantum architectures to allow for entanglement routing and its management in future quantum networks. We are currently developing the infrastructure to perform entanglement distribution, routing to different users and quantum interference on ≈1 km length scale with below 10 ps timing uncertainties. Ultimately, this work should lead to the development of a NIST measurement suite (a NIST metrology node), a convenient rack-based tool that provides telemetry data on entanglement fidelities through network links, noise contamination, timing information, losses, etc.

In collaboration with colleagues in ITL and PML, we have been working on terminology for measurement metrics related to single-photon sources and single-photon detectors. These *single-photon source and detector dictionaries* may eventually be used as a basis to develop standards. It is very clear from community input that such dictionaries will aid the development of technology, as a common language will foster common understanding about metrics and measurement protocols, i.e., a quantum components' spec sheet that has commonly accepted metrics and protocols applied to determine these metrics.

Quantum Network Testbeds. During this year, our group in collaboration with colleagues from CTL and PML, and with the support of OISM, have designed and begun to build infrastructure for a *NIST campus dedicated quantum network* that includes extensive fiber connectivity to multiple strategic laboratory locations across the NIST Gaithersburg campus, optically transparent switches and dense/coarse wavelength division multiplexers (D/CWDM) to route single photon signals for medium distance (≤ 1 km) quantum level experiments. We will deploy various commercial and NIST built quantum systems such as entangled photon

sources, quantum interfaces, quantum memories and single photon detectors to evaluate their performances within this real-life quantum network setting. The component and network link metrology tools developed by our group will be used to study the transmitting photonic qubit states as they are transmitted through noisy, lossy, and fluctuating conditions. In a project being led by CTL, this test bed will be used to develop, deploy, and test a practical quantum network manager.

The Hong-Ou-Mandel (HOM) interference effect lies at the heart of future quantum networking protocols such as entanglement swapping and teleportation. For these protocols to entangle two distant quantum network nodes, a photon from each node must arrive at a third node within a fraction of their coherence time. Both nodes, which may be kilometers apart, therefore need to be *synchronized* such that their path length differences to the measurement plane are known and compensated for. Photon coherence times vary depending on the physical implementation of the qubit system from several nanoseconds (ns) in certain quantum memories to picoseconds (ps) or even femtoseconds (fs) using spontaneous parametric down-conversion (SPDC) sources. We showed how the Ethernet-based time transfer protocol, White Rabbit (WR), can be used to synchronize two distant quantum-networked nodes to below 4 ps, sufficient to observe HOM interference with greater than 90 % visibility using SPDC photons with 17.6 ps FWHM coherence times. In collaboration with researchers in PML, we also showed and demonstrated that, in principle, low-noise quantum channels exist such that the optical WR signal can coexist with the weak quantum signals for synchronization of the actual fiber path lengths to the measurement location. Further studies are being conducted to send and interfere entangled photon states over kilometers of fiber along with the classical synchronization signals.

The *DC-QNet* is a new initiative spearheaded by NIST that brings together multiple capital area federal agencies to form a regional quantum network. The goal of this long-term project is to establish quantum "node" infrastructure at each agency and implement dedicated dark fiber links throughout for various quantum level experiments. In coordination with colleagues from ITL, PML and CTL, and with the support of the NIST quantum network initiative, we have begun to build a facility to house the NIST node of the DC-QNet. This facility will connect directly to the DC-QNet and will be integrated into the campus quantum network (described above). The node will initially have an operational space for early single-photon level experiments such as fiber

link characterization, photonic quantum state tomography, multi-km distance timing synchronization (see above) and will expand to host a user facility for more advanced quantum experiments such as entanglement distribution, deploying undistinguishable photon sources and long-distance quantum communications.

Emerging quantum application areas. _Edge computing_ has emerged to support the computational demand of delay-sensitive applications in which substantial computing and storage processes are performed at the network edge close to data sources. The vision of a hybrid classical-quantum edge is to provide a fundamentally new computing paradigm by expanding the computing capabilities and security of edge computing. The distributed nature of edge computing networks will also enable new scalable quantum networking schemes and applications. We are studying this new hybrid computing paradigm, from the fundamental concept, basic architecture, and its system simulation and emulation [2].

[1] T. Gerrits, A. Migdall, J. C. Bienfang, J. Lehman, S. W. Nam, J. Splett, I. Vayshenker and J. Wang. Calibration of Free-Space and Fiber-Coupled Single-Photon Detectors. _Metrologia_ **57**:1 (2019), 015002.

[2] L. Ding and L. Ma, Hybrid Quantum-Edge Computing: A New Computing Paradigm. _IEEE International Conference on Communications_, 2022, in review.

Designing Fabrication-Imperfection-Tolerant Thin-film Lithium Niobate Frequency Converters

Paulina Kuo

Thin-film lithium niobate (TFLN) is an exciting, emerging platform for integrated optics and quantum information processing. It offers multiple functionalities including waveguiding, routing, modulation, and frequency conversion. Owing to tight optical confinement,
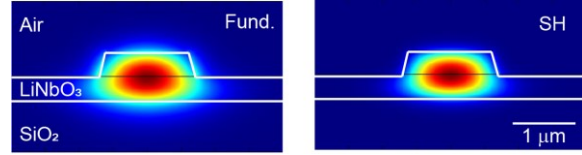
Figure 86. _Calculated fundamental and SH modes._

TFLM waveguides can achieve very high conversion efficiencies, more than 10 times higher than that of state-of-the-art diffused LN waveguides.

The performance of the TFLN frequency converters depends on the fabrication quality. However, one can design devices for minimum sensitivity to fabrication imperfections. Frequency conversion is controlled by the phase mismatch, $\Delta\beta$. For second-harmonic generation (SHG), $\Delta\beta = (\omega_2 n_2 - \omega_1 n_1)/c$ where ω_1 and ω_2 are the optical frequencies at the second harmonic (SH) and fundamental, n_i are the respective effective refractive indices, and c is the speed of light. n_i depends on the waveguide parameters, such as width, thickness, etc. If f_m represents a geometrical parameter with variations δf_m df$_m$ from fabrication errors, then we can write

$$\Delta\beta(f_m) = \Delta\beta(f_{m,0}) + (\partial\Delta\beta/\partial f_m)\,\delta f_m + O(\delta f_m^2).$$

To minimize sensitivity to fabrication imperfections, we design waveguides where $\partial\Delta\beta/\partial f_m = 0$ (or $|\partial\Delta\beta/\partial f_m|$ _is smallest_).

We used optical simulation software to model x-cut TFLN waveguides and calculate , $\Delta\beta$ for SHG (1570 nm → 785 nm). We generated maps of , $\Delta\beta$ as a function of waveguide width, total thickness and etch depth. Figure 86 shows optical mode distributions in an example x-cut waveguide. Figure 87a plots the dependence of , $\Delta\beta$ on geometrical parameters of the waveguide, and Figure 87b plots the slope $\partial\Delta\beta/\partial t$, where t is the waveguide thickness for a 1200 nm wide waveguide. We are interested in the dependence on t because thickness has the largest variation due to fabrication. We see that $|\partial\Delta\beta/\partial t|$ never reaches zero but is smallest at larger thicknesses.
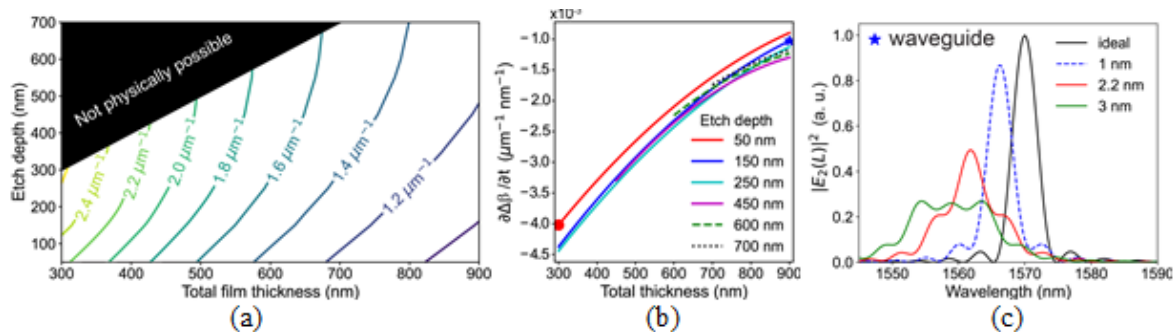
Figure 87. _(a) Map of_ $\Delta\boldsymbol{\beta}$ _for a 1200 nm wide x-cut waveguide. (b) Calculated slopes of_ $\partial\Delta\boldsymbol{\beta}/\partial\boldsymbol{t}$. _(c) Impact of linear thickness variation,_ $\Delta\boldsymbol{t}$ _on SH tuning curves for the waveguide marked by the blue star in (b)._

We investigated the effects of non-zero $|\partial\Delta\beta/\partial t|$ by deriving a model of SHG in the presence of linearly varying thickness (see details in [1]). Figure 87c plots SH tuning curves for 5 mm long TFLN waveguides with $\partial\Delta\beta/\partial t = -1\,\mu\text{m}^{-1}\,\text{nm}^{-1}$. Each curve represents different amounts of end-to-end thickness variation, Δt. The peak SH intensity is reduced by half of its ideal value for $\Delta t = 2.2$ nm. These studies are particularly important for high-efficiency frequency converters such as those needed for quantum frequency conversion.

[1] P. S. Kuo, Noncritical Phasematching Behavior in Thin-Film Lithium Niobate Frequency Converters. *Optics Letters* **47**:1 (2022) 54. DOI: 10.1364/OL.444846

Foundations of Measurement Science for Information Systems

ITL assumes primary responsibility within NIST for the development of measurement science infrastructure and related standards for IT and its applications. ACMD develops the mathematical foundations for such work. This can be very challenging. For example, many large-scale information-centric systems can be characterized as an interconnection of many independently operating components (e.g., software systems, communication networks, the power grid, transportation systems, financial systems). A looming new example of importance to NIST is the Internet of Things. Exactly how the structure of such large-scale interconnected systems and the local dynamics of its components leads to system-level behavior is only weakly understood. This inability to predict the systemic risk inherent in system design leaves us open to unrealized potential to improve systems or to avoid potentially devastating failures. Characterizing complex systems and their security and reliability properties remains a challenging measurement science problem.

A Software Framework for Simulation Research

Brian Cloteaux

Many research projects involve repeatedly running a created simulation to obtain data on the research domain [1, 2]. These simulations tend to have two important similarities: (a) the workflow that is required, and (b) a need for supporting publication-quality simulations. To support these needs, a framework to support this type of research has been under construction. This framework started with several requirements:

- *Ease of programming.* Because of the large number of different types of simulations needed, an important goal for the framework is to maximize code reuse. The chosen approach was to create a complete general-purpose simulation application where new simulation code is added as a library and accessed through polymorphism. Information about a user simulation is defined using a JSON file which describes the project including the parameters needed in the simulation. This file is parsed, using a Python script, to produce the files needed to interface with the framework. This also allows us to easily change parameters in a simulation. We simply modify the parameters in the project description file and then reparse it.

- *Recording of simulation metadata.* For the simulation framework, all results are stored in a JSON file format. This allows for metadata of the run to be stored with the results. Thus, a simulation run automatically records all the parameters used in the run. In addition, it also stores information about the run environment, including date of execution, length of runtime, and the operating system it ran on. This allows us to have all these descriptive values if we want to use the results in a publication.

- *Reproducibility of results.* For any run in the framework, the simulation reads in all its parameters through a JSON formatted parameter file. For all
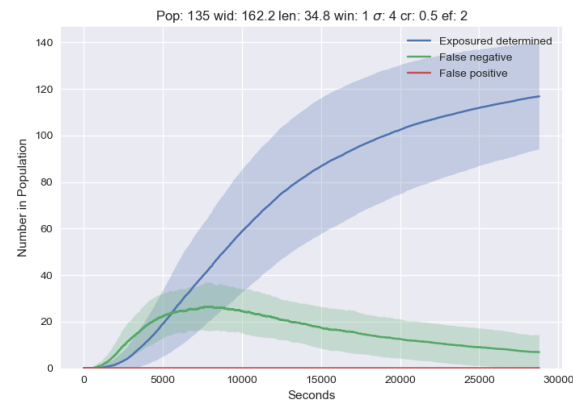
Figure 88. *Example of the results from an application created in the simulation framework. The confidence intervals represent one standard deviation.*

simulations, there is an implicit parameter: the seed for the random number generator. If a seed is given as a parameter, then the simulation uses it as the seed, else the simulation will generate a new random seed. When the results are records in the output file, all the parameters used for the run, including the random seed, are recorded in the file. This allows for a mechanism to reproduce a simulation run. The file for the result of a simulation can be reused as the parameter file to reproduce the same result. This ability is important when (a) trying to explain unexpected results, and (b) having confidence in reproducing published values.

- *Support for confidence intervals.* Since all simulation results coming from NIST are required to have confidence intervals, the framework has automatic support for these in the results files. A simple viewer is also set up to show these intervals on the results (see Figure 88).

The overall design of the framework has allowed for a couple of other advantages. The first is allowing for ease in GUI (graphical user interface) development. The choice of parameterizing all the runs using a JSON file and then also outputting the results in a JSON file

allows for strict isolation of the simulator. This allows for ease in overlaying a GUI on any simulator using a model/view architecture. A simple Python script can take the description file for a project and generate a light GUI.

A second advantage has been allowing for simpler multiple platform development. Because of the pandemic, development and run environments are often different. This causes problems because of differences in operating systems, compilers, libraries, etc. The simulation framework reduces some of these issues by integrating tools such as cmake and git into the workflow through the project setup script.

This is not a finished work. Currently the framework is being expanded in two ways. The first is to integrate it with GitHub. This integration will allow recording version numbers of software libraries and simulation code directly into the metadata of the results.

A second area for development is support for checkpoint / restart. A number of simulations have extended run times. Allowing for a generalized checkpoint / restart mechanism to the framework will allow the framework to be used for investigations over larger domains.

When this framework is completed, the plan is to release it on GitHub. This will allow simulation results to be released through NIST's Midas system and allow a reproducibility mechanism for published results by allowing the results to be regenerated on released code.

[1] B. Cloteaux and V. Marbukh. SIS Contagion Avoidance on a Network Growing by Preferential Attachment. In *Joint International Workshop on Graph Data Management Experiences & Systems (GRADES) and Network Data Analytics (NDA)*, Article 11, pages 1-4, 2019.

[2] K. Sayrafian, B. Cloteaux, V. Marbukh, and C. Emiyah. Evaluation of the Bluetooth-based Proximity Estimation for Automatic Exposure Determination. In *IEEE Consumer Communications & Networking Conference (IEEE CCNC)*, Online, January 8-11, 2022.

Algorithms for Identifying Important Network Nodes for Communication and Spread

Fern Y. Hunt
Roldan Pozo

The identification of nodes in a network that enable the fastest spread of information is an important, if not fundamental, problem in network control and design. It is applicable to the optimal placement of sensors, the design of secure networks, and the problem of control when network resources are limited. Our approach to this problem has its origins in models of opinion dynamics and the spread of innovation in social networks. The mode of communication between nodes is described by simple models of random or deterministic propagation of information from a node to its neighbors. During the past few years, we have made progress in understanding the structural requirements for sets of nodes enabling effective spread in networks and have developed scalable algorithms for constructing these sets in real world networks.

We consider a discrete time model of information spread (represented by a variable assigned to each node) in a network with a set of nodes V and a subset $A \subseteq V$ of k nodes representing leaders or stubborn agents that are initially assigned a single value. Propagation occurs by iterated averaging or diffusion defined by a stochastic matrix P. All node values will eventually converge to the single value at a speed determined by the sub-stochastic matrix $P|_{A'}$, the matrix P restricted to the complement of A. An effective spreader is then a set of nodes for which convergence to this single value is fastest, i.e., the set A for which the Perron-Frobenius eigenvalue of $P|_{A'}$ is largest. Using a classical result of Markov chain theory, the problem can be recast in terms of finding the set A of cardinality k that minimizes the mean first hitting time, i.e., the expected time a random walker reaches the target set A for the first time.

We first proposed a polynomial time algorithm for finding an approximation to the optimal set [1]. It is an extension of the classic greedy algorithm, and it begins with a class of optimal and near optimal starter sets of smaller cardinalities rather than the conventional choice of a best singleton set. An optimal spreader in our setting is defined in terms of a set function F where for a subset $A \subset V$, $F(A)$ is the sum of mean first arrival times to A by random walkers that start at nodes outside of A. Direct comparison of the algorithm results with the actual optimal solution and lower bounds on the performance ratio can be obtained because F is a supermodular set function [2]. However, for large complex networks commonly encountered in applications, another approach is needed.

We then developed a set of fast heuristics that work well on graphs with large hubs, a common feature of complex networks. When the desired set cardinality is k, subsets of hub vertices are rapidly screened to produce candidate sets. Each set consists of k nodes whose first (or higher order) neighborhoods have minimal overlap. After further screening, the offered approximation is selected by ranking the results of a Monte Carlo calculation of the optimal set F for each candidate. This process allows us to find near optimal and optimal spreaders in networks with millions of nodes and dozens of millions of edges in less than a few seconds on a typical laptop. After conducting tests on real world graphs from diverse application areas including molecular biology, traffic control, and social networks, we hypothesize that the method is most effective in terms of speed and

quality of offered solutions when it is used on graphs with a large ratio of maximal degree to average degree.

Understanding that the resulting offered set was an approximate solution of a discrete stochastic optimization problem, we established sufficient conditions that imply that it is also an approximate solution of the original problem. The first step was to establish the accuracy of the Monte Carlo calculation of F. The fact that the first hitting time to a set A has a distribution with exponential tails means that a sample average of simulated hitting times produces a consistent estimate of F in the limit of large sample size i.e., number of simulations.

Establishing the degree of optimality of any offered solution is very difficult since supermodularity cannot be used and the size of the graphs are so large. However, the methods we use make it possible to rapidly sample the distribution of possible F values. We suppose the screening and ranking procedures produce candidate sets with F values that rank in the highest percentile of a distribution of such values over all subsets of fixed cardinality. Independent repetition of the heuristic calculation enables us to produce an estimate of a fixed percentile value along with a confidence interval for that estimate. The latter follows from an application of Chebyshev's inequality. Note that the resulting interval contains both the offered solution value and the optimal value of the original problem. Even in the case of a large number of repetitions, this approach is promising because it takes very little time to perform a single execution. The results of our work are reported in [3].

Recently, we have focused on related models of information spread, such as the broadcast model (*k*-median problem) which seeks the minimal sum of distances, and epidemiological models (Susceptible-Infected-Recovered) that better represent the spread of infections on network topologies. We have studied and compared the efficacy of several heuristic algorithms used in the literature and have developed a different approach (minimizing overlapping neighborhoods) which aims to produce quality solutions at a fraction of the computational cost, making it appropriate for use on large networks. For example, we are able to process graphs of several million vertices with competitive solutions that run one to two orders of magnitude faster than conventional methods.

[1] F. Hunt. Using First Hitting Times to Maximize the Rate of Convergence to Consensus. Preprint, 2018. arXiv:1812.08881

[2] F. Hunt. An Algorithm for Identifying Optimal Spreaders in a Random Walk Model of Network Communication. *Journal of Research of the NIST* **121** (2016), 121008. DOI: 10.6028/jres.121.008

[3] F. Hunt and R. Pozo, Fast Methods for Identifying Effective Spreaders in Real Network. *Journal of Research of the NIST* **125** (2020), 125036. DOI: 10.6028/jres.125.036

Towards Actionable Cybersecurity Risk Metrics in a Hyperconnected World

Vladimir Marbukh
Brian Cloteaux
Debasis Mitra

The goal of this work is to study the applicability of models and metrics from economics and finance to the measurement of risk in today's networked cybersystems. Conventional security risk measures are based on point estimates of the likelihoods of potential multi-step attacks that combine multiple vulnerabilities. Drawbacks of these measures are due to disregard for the tail risk, inherent inaccuracy of estimates of low probabilities, and reliance on the specific attacker(s) model. A recently proposed measure of cybersecurity risk, Cyber security Value at Risk (CyVaR), which is based on the VaR measure of financial risk, accounts for the tail risk. However, CyVaR still suffers from reliance on the specific attack model, and moreover has its own problems, e.g., is not a coherent risk measure. Following the recent trend of replacing VaR with robust and coherent Entropic VaR (EVaR), in [1] we proposed and discussed a Cyber security Entropic Value at Risk measure (CyEVaR). Figure 89 depicts CyEVaR as a function of the attack severity measured by the expected number of exploited vulnerabilities. As this number increases, system CyEVaR grows from the average to the maximum loss.

The highly aggregated macroeconomic Gordon-Loeb (G-L) optimization model for cyber security investments assumes a known system-level return on the security investment, measured by the macroeconomic utility. A missing link in application of the G-L model to a specific system is the derivation of this macroeconomic utility from the system-specific microeconomic model for cybersecurity investments. This microeconomic model should identify the optimal mixture of investments in elimination/mitigation of specific vulnerabilities, given the aggregate security investment. In [2] we report initial results on developing a microeconomic optimization model for security investments for system, which produces the system-specific macroeconomic utility to be used in the G-L model.

Our work on optimal cybersecurity investments combines two research threads: first, the Gordon-Loeb model for optimal investments in which a given function defines the tradeoff between the likelihood of a system breach and investments to mitigate vulnerabilities; second, interactions between the system Defender and Attacker [3]. In our model the Attacker (i) makes repeated attempts to breach defenses, while incurring costs with each attempt, (ii) breaches the system if even one attempt succeeds, and (iii) gains a monetary benefit
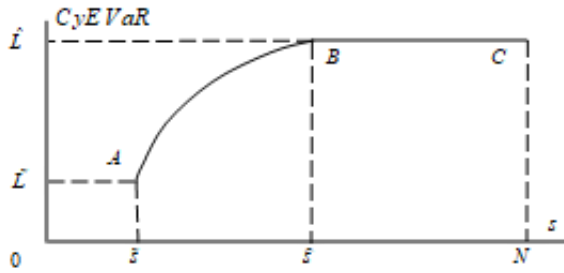
Figure 89. *Entropic Value at Risk vs. attack severity.*

if the system is breached. The Defender optimizes investments to minimize the expected monetary loss if the system is breached, given knowledge of the Attacker's decision-making policy and a function, as in Gordon-Loeb, which determines the tradeoff between investments and reduced susceptibility of system breach. Our formulation is of Min-Max type, or Stackelberg leader-follower game. We have succeeded in characterizing the Defender's optimal investment policy even though the problem is not convex, and therefore generally considered not tractable.

Economics drives two major evolutionary trends in networked system design/operation: towards the boundary of the system capacity region, and a trend towards an increase in the system interconnectivity allowing for enlargement of the capacity region through dynamic resource sharing. However, numerous recent systemic failures in various networked systems, e.g., power grids, cloud, and financial systems, have demonstrated that the economic benefits of interconnectivity often heighten risk of cascading failures/overloads. These empirical observations pose problem of systemic risk management while maintaining economic viability. Our preliminary results [4] indicate that a system operating point should be close to the triple point on the system phase diagram, where the following three regions converge: two regions where the normal (failed/overloaded) equilibrium state is globally stable and a region where these equilibrium states coexist as locally stable, i.e., metastable.

[1] V. Marbukh. Towards Robust Security Risk Metrics for Networked Systems: Work in Progress. In *Proceedings of the IFIP/IEEE International Symposium on Integrated Network Management*, Online, May 17-21, 2021.

[2] V. Marbukh. Towards Economically Efficient Security Risk Reduction. In *Proceedings of the 31st European Safety and Reliability Conference* (ESREL'21). Online, Sept. 19-23, 2021.

[3] A. Ebel and D. Mitra. "Attacker-Defender Investment Strategies in Cybersecurity." Data Day, Data Science Institute, Columbia University. March 2021.

[4] V. Marbukh. Towards Economically Efficient Mitigation of Systemic Risk of Undesirable Contagion. In *Proceedings of the 31st European Safety and Reliability Conference* (ESREL'21). Online, September 19-23, 2021, Virtual.

Measurements of Cyber Risks in Complex Systems and Optimal Cybersecurity Investments

Richard J. La
Van Sy Mai (NIST CTL)
Abdella Battou (NIST CTL)

Modern engineered systems, such as information and communication systems and power systems, are actually comprised of many (sub-)systems. In order to provide their services, the comprising systems must work together and support each other. However, due to this interdependence among comprising systems, it is possible for a local failure or infection of a system by malware to spread to other systems, potentially compromising the integrity of the overall system. For instance, an outage in one part of a power grid can trigger cascading failures and cause a large-scale blackout (e.g., the Northeast blackout of 2003).

Clearly, the structure of underlying interdependence among systems, which is modeled using a directed interdependence network, has considerable impact on the propagation dynamics of failures or malware infections. This suggests that any sound investments in cybersecurity of complex systems or the control of epidemics should consider interdependence in the systems.

With this in mind, we studied the problem of measuring the cyber risks to nodes (e.g., subsystems) and finding optimal cybersecurity investments for hardening vulnerable nodes in large systems, which will minimize the long-run average costs of a system operator. Our model accounts for both cybersecurity investments and recovery/repair costs ensuing infections or failures, which we call infection costs. We first considered a scenario in which malicious actors launch external attacks under the assumption that the interdependence network is strongly connected [1, 2]. In the second case, we assumed that there are no external attacks and examined the steady state, starting with an initial state where some nodes are infected [2]. Even in the absence of external attacks, infections can persist due to secondary attacks by infected systems on other systems and, as a result, may not be able to attain an infection-free steady state.

We formulated the problem of determining the optimal cybersecurity investments that minimize the average costs as an optimization problem. Unfortunately, this optimization problem is nonconvex and cannot be solved easily. In order to gauge the quality of a feasible solution, we obtained both a lower bound and an upper bound on the optimal value of our problem. A lower bound can be acquired using convex relaxations of the original problem. For these convex relaxations, we also derived a *sufficient condition* under which the solution of a convex relaxation solves the original nonconvex optimization problem. An upper bound on the

optimal value can be obtained using an algorithm that finds a local minimizer. To this end, we proposed two methods: a reduced gradient method (RGM) and sequential convex programming, both of which produce a local minimizer. Together, our approach offers a bound on the optimality gap.

Numerical studies show that the computational requirements for the proposed methods are light to modest even for large systems, except for one method that requires an inverse matrix [2]. They suggest that, in almost all cases that we considered, the gap between the lower bound on the optimal value and the cost achieved by our solutions is small. In addition, when the infection costs are large, which are more important scenarios, the sufficient condition for the convex relaxations to be exact holds, and we obtain an optimal point by solving the convex relaxations. Finally, the RGM is computationally most efficient (with the computational time being less than two seconds in all considered cases and less than 0.1 seconds in most cases) and the quality of solutions is on par with that of other methods. This suggests that the RGM may offer a good practical solution.

As mentioned earlier, these findings require that the underlying interdependence network be strongly connected. We recently investigated scenarios where this assumption is relaxed, and the interdependence network is only weakly connected. A key challenge here is that we can no longer guarantee the existence of a unique equilibrium of the underlying dynamics at steady state. Despite this challenge, by perturbing the original system and leveraging our results for the strongly connected cases, we were able to design a computationally efficient algorithm for finding nearly optimal solution.

[1] V.-S. Mai, R.J. La, and A. Battou. Optimal Cybersecurity Investments for SIS Model. In *IEEE Global Communications Conference*, December 2020.

[2] V.-S. Mai, R.J. La, and A. Battou. Optimal Cybersecurity Investments in Large Networks Using SIS Model: Algorithm Design. *IEEE/ACM Transactions on Networking* **29**:6 (2021), 2453-2466.

Detection of Hidden Infections and Parameter Estimation in Epidemic Processes

Richard J. La
Van Sy Mai (NIST CTL)

Epidemic processes are often used to study the spread of malware and viruses in computer networks and information systems, contagious diseases, rumors, or new product information. Many existing studies assume that the infected nodes (e.g., subsystems infected by malware in complex systems or individuals who contracted

a disease) and, in some cases, the sequence of infections are known so that key parameters of the epidemic processes, such as reproduction numbers or transmission rates, can be estimated from available data.

Unfortunately, in practice such information is not always available or may not be easily acquired in some cases. For example, when devices on a network, such as servers in an enterprise system, with inadequate security measures are infected, users and network managers may not discover that their devices are compromised for some time. This concern is growing with the emergence of artificial intelligence-powered malware [1, 2]. Similarly, an infectious disease can cause widely varying symptoms and some infected individuals may exhibit no obvious outward symptoms that can be easily identified.

Developing appropriate and effective cybersecurity measures or public health policies requires accurate information. This task would become more challenging when it is hard to obtain a clear picture of how aggressively or quickly computer malware or infectious diseases are spreading in a network or society; missing a large number of infections in reported cases at an early stage of an outbreak of a computer virus or a deadly pandemic would likely lead to a potentially costly underestimation of the severity of the problem, which could have been mitigated by timely countermeasures and policies. For this reason, it is crucial to ascertain in an early stage of an epidemic whether the numbers of reported cases accurately represent all infection cases, or they are missing many undetected cases, so that cybersecurity or healthcare experts can introduce suitable countermeasures and remedies.

To this end, we proposed a new approach to determining whether the reported numbers of infection cases are accurate or there are many unreported cases [3]. Our approach is based on a simple observation that infected devices or individuals that go undetected spread their infection differently than those that are detected shortly after suffering infection. Making use of this observation, we formulated our problem as one of identifying the model that best describes or fits the available data, and our proposed solution leverages insights from (statistical) model selection [4].

[1] J. Menn. New Genre of Artificial Intelligence Programs Take Computer Hacking into Another Level. *Reuters*, August 8, 2018.

[2] B. Dickson. How Hackers Can Use AI to Hide Their Malware and Target You. *The Daily Dot*, August 27, 2018.

[3] V.-S. Mai and R. J. La. Story of Two Populations in Epidemics: Is Every Infection Counted? In *Complex Networks and Their Applications X. COMPLEX NETWORKS 2021*. Studies in Computational Intelligence **1015**, Springer 2022.

[4] H. Linhart and W. Zucchini. *Model Selection*. John Wiley & Sons, 1986.

Statistical Change Detection for Network Anomalies

Assane Gueye
Anthony Kearsley
Pulkit Grover (Carnegie Mellon University)
Rosaline Su (Carnegie Mellon University)

Around the world, network traffic ebbs and flows, producing enormous amounts of observable information (e.g., data flow, queue times). There are instances where networks are affected by some outside interference and this, in turn, leads to changes in network performance and thus a change in observable operational statistics of the network. In this project we seek to study and characterize these changes. We hope to answer the following questions. Are these changes statistically significant and can they be used to indicate outside interference? If so, with what kind of certainty can we say interference has taken place and how quickly can we say it? Finally, can this lead to characterizing and classification of detected interferences?

Our approach is to examine time series representations of network statistics like queuing times, routers' load, and other time-dependent network characteristics. Initial data will be a straight-forward time-series histogram of a collection of statistical network data. Given a metric, a regression of this data can then be performed. However, it is clear that this regression will vary substantially as a function of selected metrics. Our current plans include a variety of metrics including the l_1, l_2, l_∞ vector norms, the X_2 metric and the Wasserstien metric, sometimes called the Earth Mover Distance (EMD). A stochastic regression of the time series can then decompose our time series into advection and diffusion where, presumably, the ambient noise present in every network should be identifiable.

A stochastic regression of our network observations yields a stochastic differential equation (SDE) comprised of a deterministic and stochastic component. If successful, this decomposition will isolate ambient noise in the network and allow more scrutiny of the underlying network behavior signal. Large variations in these two quantities can then be used to identify changes indicative of outside stimulation. Our hope is that if statistical quantities are well selected, malicious interference will become apparent from our regression. Roughly speaking, the regression should provide an estimate and statistical characteristics of an ordinarily running network and large variations from that baseline can be easily identified using this technique.

We hope to answer the following questions. Have the statistics changed in a way that indicates outside interference? If so, with what kind of certainty can we say this interference has taken place and when? Finally, is there a way to characterize the type of interference?

This project is with collaboration with the main Carnegie Mellon University campus in Pittsburgh and CMU-Africa (Kigali Campus).

[1] J. Fan, J. Jiang, C. Zhang, and Z. Zhou. Time Dependent Diffusion Models for Term Structure Dynamics. *Statistica Sinica* **13** (2003), 965–992.

[2] C. Manikopoulos and S. Papavassiliou. Network Intrusion and Fault Detection: A Statistical Anomaly Approach. *IEEE Communications Magazine* **40**:10 (2002), 76-82.

[3] V. Kotu and B. Deshpande. *Data Science, Concepts and Practice*. Second Edition, Morgan Kaufmann, 2019, Chapter 13, 447-465.

[4] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, Heidelberg, 1992.

[5] Y. Rubner, C. Tomasi, and L. J. Guibas. The Earth Mover's Distance as a Metric for Image Retrieval. *Int. Journal of Computer Vision* **40** (2000), 99-121.

Measurement-Based End-to-End Quality-of-Service Assurance in 5G/6G Telecommunication Systems

Richard J. La
Van Sy Mai (NIST CTL)
Tao Zhang (NIST CTL),
Abdella Battou (NIST CTL)

Providing differentiated services to meet the unique requirements of different use cases is a major goal of the fifth generation (5G) telecommunication networks and will be even more critical for future 6G systems. Fulfilling this goal requires the ability to assure end-to-end (E2E) quality-of-service (QoS), which remains a major challenge. A key factor that makes E2E QoS assurance difficult in a telecommunication system is that access networks (ANs) and core networks (CNs) manage their resources autonomously. To this day, few results are available that can ensure E2E QoS over autonomously managed ANs and CNs.

Existing techniques rely predominately on each subsystem (i.e., an AN or a CN) to meet static local QoS budgets with no recourse in case any subsystem fails to meet its local budget, which leads to a violation of E2E QoS. Moreover, most existing distributed optimization techniques that can be applied to assure E2E QoS over autonomous subsystems require the subsystems to exchange sensitive information, such as their local decision variables that can reveal how they manage their networks.

To address this challenge, we developed a new framework for ensuring E2E QoS with autonomous subsystems, in which the required E2E QoS is cast as global
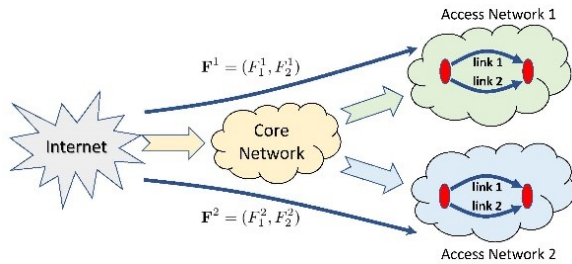
Figure 90. *Example network with one CN and two ANs.*
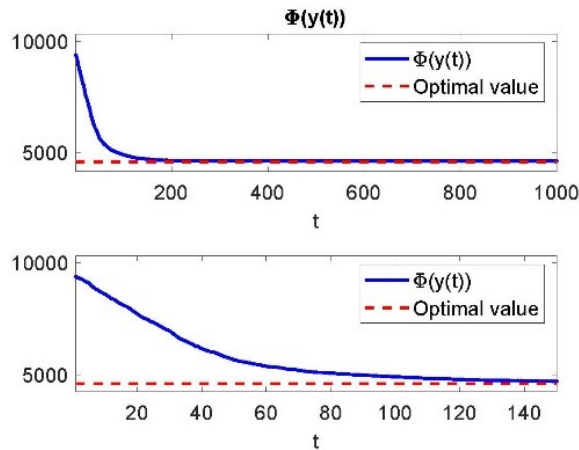
Figure 91. *Plot of the global objective function.*
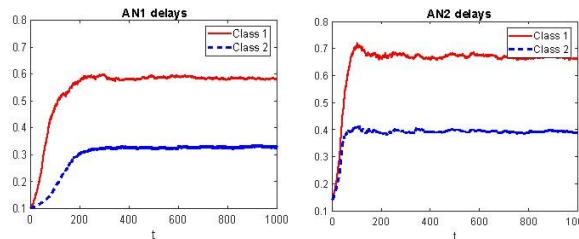
Figure 92. *Plot of the E2E delays experienced by traffic destined for AN 1 and AN 2.*

constraints that ANs and CNs need to satisfy collectively [1]. We then proposed a distributed algorithm that can enable ANs and CNs to autonomously "cooperate" with each other to dynamically negotiate their local QoS budgets by sharing only their estimates of the global constraint functions, without disclosing their local decision variables [2]. We proved that the distributed algorithm, which requires only noisy measurements, guarantees almost sure convergence to an optimal solution under mild technical conditions. Furthermore, our numerical results demonstrate that the convergence to an optimal solution occurs quickly even with significant measurement noise.

A simple example network with one CN and two ANs, which support two classes of traffic, is shown in Figure 90 [1]. The E2E delay experienced by class 1 (resp. class 2) traffic should not exceed 0.7 (resp. 0.4).

We only consider downstream traffic that traverses through the CN and is destined for one of the two ANs.

We plot in Figure 91 the global objective function that accounts for both the weighted delays experienced by the traffic and the costs of provisioning bandwidth for two classes of traffic at each network. The figure illustrates quick convergence of the objective function to the optimal value.

Figure 92 shows the E2E delay experienced by traffic destined for AN 1 and AN 2. It is clear that the delays remain below the upper bounds of 0.7 and 0.4 for class 1 and 2 traffic, respectively, for both networks.

[1] V.-S. Mai, R. J. La, T. Zhang, and A. Battou. End-to-End Quality-of-Service Assurance with Autonomous Systems: 5G/6G Case Study. In *IEEE Consumer Communications and Networking Conference*, January 2022.

[2] V.-S. Mai, R.J. La, T. Zhang, and A. Battou. Distributed Optimization with Global Constraints Using Noisy Measurements. In review.

Bi-Criteria Radio Spectrum Sharing with Subspace-Based Pareto Tracing

Zachary J. Grey
Andrew Dienstfrey
Susanna Mosleh (NIST CTL)
Jacob Rezac (NIST CTL)
Yao Ma (NIST CTL)
Jason Coder (NIST CTL)

Motivated by demands for more efficient use of limited radio spectrum, next-generation wireless communications devices are likely to share spectrum between multiple radio access technologies. While spectrum sharing holds promise, balancing operational requirements of new network paradigms against those of incumbent technologies presents significant technical barriers. To meet these challenges, CTL researchers, in collaboration with ACMD, are developing measurement and simulation techniques to verify configurations of wireless devices which allow for coexistence while minimizing degradation.

In one such example of managing spectrum scarcity, providers are beginning to operate Long-Term Evolution License-Assisted Access (LAA) in unlicensed bands. Operating LAA in unlicensed bands improves spectral-usage efficiency. However, LAA may or may not impede Wi-Fi operation depending on details of operational and environmental variables. Understanding and addressing these challenges calls for a deep dive into the operations and parameter selection of both networks in the medium access control (MAC) and physical layers (PHY).
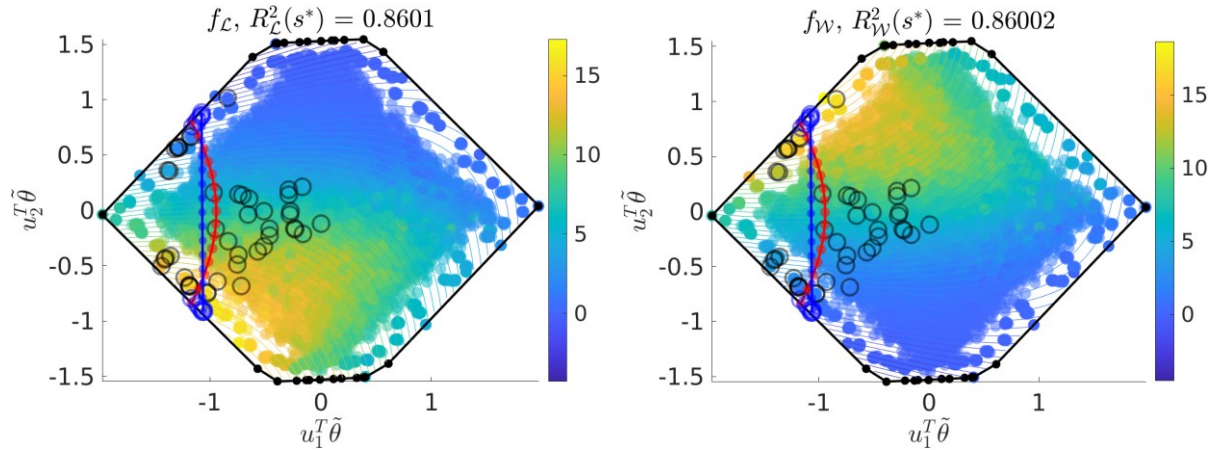
Figure 93. *Pareto trace of quadratic ridge profiles. The quadratic Pareto trace (red curve and dots) is overlaid on a shadow plot over the mixed coordinates (colored scatter) with the projected bounds and vertices of the domain (black dots and lines). Convex quadratic approximations over the mixed subspace coordinates (colored contours) are computed with a quadratic program. Also depicted is the projection of the non-dominated domain values from 10 000 random samples (black circles). The trace begins at $\tau = 0$ with near maximum quadratic Wi-Fi throughput and we move (smoothly) along the red curve to $\tau = 1$ obtaining near maximum quadratic LAA throughput---maintaining an approximately best trade-off over the entire curve restricted to a projection of the origin. This is contrast with a linear interpolation of $\tau = 0$ and $\tau = 1$ solutions approximated by interior point method (blue curve) as well as a brute force approach consisting of 15 successive approximations by interior point method according to discrete values of $\tau \in [0, 1]$ (blue circles). Note: original parameter scales were transformed for numerical considerations (among other reasons) thus resulting in the accented (tilde) vector.*

Mathematically, we formulate spectrum-sharing analysis as a multi-criteria optimization problem in which both Wi-Fi and LAA key-performance-indicators (KPIs), such as network throughputs on unlicensed bands, are simultaneously maximized with respect to their PHY and MAC layer parameters. This set of maximizing arguments quantifies the inherent trade-off between LAA and Wi-Fi data throughputs. While this is a desirable goal, it is complicated by the high dimensionality of the problem. Furthermore, a systematic parameterization the Pareto front of non-dominated solutions is needed.

Our model uses 17 MAC and PHY variables to characterize the coexistence performance. However, previous experience suggests that not all variables (or combinations of variables) are equally important in determining qualities of KPIs. We used active subspace analysis to provide a systematic dimension reduction strategy, providing a rigorous foundation for these observations.

Our published work has facilitated an approach for simultaneously optimizing network KPIs sharing limited unlicensed spectrum resources. An exploratory analysis utilizing an example of LAA coexistence with Wi-Fi network identified a common subspace-based dimension reduction of a basic model of network behavior. This enabled visualizations and low-dimensional approximations which led to a *continuous* approximation of the Pareto frontier for the multi-criteria problem of maximizing all convex combinations of network throughputs over MAC and PHY parameters. Such a result simplifies the search for parameters which enable high quality performance of both networks, particularly compared to approaches which do not operate on a reduced parameter space. Analysis of the LAA-Wi-Fi example revealed an *explainable and interpretable* solution to an otherwise challenging problem—*devoid of any known convexity until subsequent exploration*. These results were presented at the 2021 IEEE International Conference on Communications [1].

A follow-on paper [2] incorporates alternative low-dimensional approximations to emphasize that a convex quadratic surrogate-based approach admits unstable Pareto trace approximations when posed with the full dimensionality. We also summarize active subspace approximation diagnostics, comparisons to alternative subspace approximations, a numerical study of stability, a space filling scheme over the reduced dimension subspace (called stretch sampling), and an explicit parametrization of a predominantly flat manifold of near-Pareto optimal solutions.

Future approaches will enable spectrum sharing in unlicensed bands by simplifying the design of wireless network operation and architecture—ultimately quantifying MAC and PHY parameter combinations with reduced intrinsic dimension giving near-optimal KPI trade-offs. In recent work CTL researchers studied how changing Wi-Fi transmit power impacted KPIs of the networks such as percentage error rate and transmission throughput. They introduced a new concept, the *region of wireless coexistence* (RWC), representing a set of transmit powers which lead to acceptable values of KPIs. This technique reduces required measurements by around 40% compared to a baseline uniform measurement. However, scaling to larger problems requires identification the RWC in higher dimension wireless coexistence scenarios. These scenarios will have tens or hundreds of independent variables, leading to a dramatic
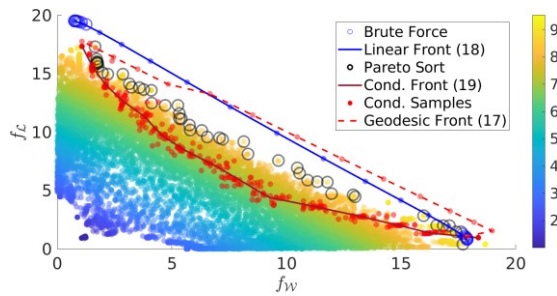
Figure 94. *Approximation of the Pareto front resulting from the quadratic trace. The approximate Pareto front (red curve) is shown with the non-dominated throughput values (black circles) and scatter of 10 000 random responses colored according to the averaged throughputs. The dashed-red curve is the image of the continuous geodesic trace through parameter space (visualized as the red curve in Figure 93 representing a near Pareto optimal set of solutions). The solid-red curve represents an approximation of the conditional front by taking conditional averages over inactive parameters. The results in blue are the corresponding throughput responses according to the description in Figure 93.*

increase in the volume of the candidate space of RWC sets. Existing RWC surrogate models (Gaussian process) are unlikely to be effective for such high

dimensional problems. Thus, we aim to use the dimensionality-reduction techniques described above to combat issues of computational complexity in larger problems. We hope to accelerate the RWC algorithm by experimenting with how to efficiently take measurements over a manifold of reduced intrinsic dimension—similar to the one identified in this work. We anticipate that the corresponding measurement requirements for estimating RWC will be orders-of-magnitude fewer than a technique which does not use adaptive sampling in conjunction with dimension reduction.

[1]   Z. Grey, S. Mosleh, J. Rezac, Y. Ma, A. Dienstfrey, and J. Coder. Optimizing Unlicensed Band Spectrum Sharing with Subspace-Based Pareto Tracing. Proceedings of the IEEE International Conference on Communications, June 2021. URL: https://arxiv.org/pdf/2102.09047.pdf

[2]   Z. J. Grey, S. Mosleh, J. D. Rezac, Y. Ma, J. B. Coder, and A. M. Dienstfrey. Multi-Criteria Radio Spectrum Sharing with Subspace-Based Pareto Tracing. *IEEE Transactions on Communications*, to appear. URL: https://arxiv.org/pdf/2108.10414.pdf.

Mathematical Knowledge Management

We work with researchers in academia and industry to develop technologies, tools, and standards for representation, exchange, and use of mathematical data. Of particular concern are semantic-based representations which can provide the basis for interoperability of mathematical information processing systems. We apply these representations to the development and dissemination of reference data for applied mathematics. The centerpiece of this effort is the Digital Library of Mathematical Functions, a freely available interactive and richly linked online resource, providing essential information on the properties of the special functions of applied mathematics, the foundation of mathematical modeling in all of science and engineering.

Digital Library of Mathematical Functions

Barry I. Schneider
Bruce R. Miller
Bonita V. Saunders
Howard S. Cohl
Marjorie A. McClain
Daniel W. Lozier
Ronald F. Boisvert
Charles W. Clark (NIST PML)
Adri B. Olde Daalhuis (University of Edinburgh)
Gergő Nemes (Alfréd Rényi Inst. of Mathematics)
Wolter Groenvelt (Delft University of Technology)
Tom Koornwinder (University of Amsterdam)
Yuan Xu (University of Oregon)
Bill Reinhardt (University of Washington)

http://dlmf.nist.gov/

Progress in science has often been catalyzed by advances in mathematics. More recently, developments in the physical sciences, such as investigations into string theory, have influenced pure mathematics. This symbiotic relationship has been extremely beneficial to both fields. Mathematical developments have found numerous applications in practical problem-solving in all fields of science and engineering, while cutting-edge science has been a major driver of mathematical research**.** Often the mathematical objects at the intersection of mathematics and physical science are mathematical functions. Effective use of these tools requires ready access to their many properties, a need that was capably satisfied for more than 50 years by the *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, which was published by the National Bureau of Standards (NBS) in 1964 [1].

The 21st century successor to the NBS Handbook, the freely accessible online Digital Library of Mathematical Functions (DLMF) together with the accompanying book, the *NIST Handbook of Mathematical Functions* [2], published by Cambridge University Press in 2010, are collectively referred to as the DLMF.

Figure 95. *A visual history of the DLMF from its roots in the 1964 NBS Handbook to the graphical contents of the present DLMF.*

The DLMF continues to serve as the gold standard reference for the properties of the special functions of applied mathematics.

The DLMF has considerably extended the scope of the original handbook as well as improving accessibility to the worldwide community of scientists and mathematicians. To cite a few examples, the new handbook contains more than twice as many formulas as the old one, coverage of more functions, in more detail, and an up-to-date list of references. The website covers everything in the handbook and much more: additional formulas and graphics, math-aware search, interactive zooming and rotation of 3D graphs, internal links to symbol definitions and cross-references, and external links to online references and sources of software.

While the original Handbook still receives an enormous number of citations, citations to the DLMF are steadily growing in relation to the original handbook. Google Scholar now reports more than 7262 citations to

the DLMF, a roughly 16 % increase from 2020. The number of DLMF website pages served up and the number of unique visitors to the website each increased by 7 % over calendar 2020.

Today's DLMF is the product of many years of effort by more than 50 contributors. Its initial release in 2010, however, was not the end of the project. Corrections to errors, clarifications, bibliographic updates, and addition of new material all need to be made on a continuing basis. And new chapters covering emerging subject areas need to be added to assure the continued vitality of the DLMF deep into the 21$^{\text{st}}$ century. Since December of 2020, there were five DLMF releases, 1.1.0 (2020-12-15), 1.1.1 (2021-03-15), 1.1.2 (2021-06-15), 1.1.3 (2021-09-15), and 1.1.4 (2022-01-15) which kept us on our quarterly release schedule. Release 1.1.0 was a significant revision from the previous year's release (1.0.28). Most notable in this release was the development and implementation for the possibility of new chapters, sections, subsections, and equations being introduced with a decimal numbering scheme using "_" to delimit intermediate numbers for sections, equations, etc. Since that release, thirty-four new equations and two new sections have been introduced into the DLMF. See Figure 96 for some examples.

The updating of various DLMF chapters and the development of new ones continues. These include a new chapter on Several Variable Orthogonal Polynomials (SVOP) and substantial updates to the chapters on Orthogonal Polynomials (OP), Algebraic Methods (AM), Painlevé Transcendents (PT) and Zeta and Related Functions (ZE). Four authors and two validators were identified to carry out the work. Drafts are now available for three of the chapters and are being internally reviewed. External validation of the chapters is following in much the same manner as the original DLMF. The ZE chapter revision with full validation was completed in release 1.1.4. The OP validation is underway it is expected that the full revision of the OP and AM chapters, will be released in 2022.

One of the design goals for the DLMF was that each formula would be connected to a proof in the literature. This data, visible as annotations on the website, provides either a proof for the formula, a reference to the proof for the formula or, for definitions, a reference which gives that definition. Unfortunately, this information has not previously been provided in all cases. Our work to systematically verify the completeness and traceability

- (14.30.8_5). Herglotz generating function for spherical harmonics in §14.30. Spherical and Spheroidal Harmonics, $\mathbf{a} = \left(\frac{1}{2\lambda} - \frac{\lambda}{2}, -\frac{i}{2\lambda} - \frac{i\lambda}{2}, 1\right)$, $\mathbf{x} = \left(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta\right)$:

$$e^{t\mathbf{a}\cdot\mathbf{x}} = \sqrt{4\pi}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{t^n r^n \lambda^m Y_{n,m}(\theta,\phi)}{\sqrt{(2n+1)(n+m)!(n-m)!}}.$$

- (16.4.2_5). Lerch Sum:

$$_3F_2\left(\begin{matrix}-n,a,1\\-n,c\end{matrix};1\right) = \sum_{k=0}^{n}\frac{(a)_k}{(c)_k} = \frac{c-1}{c-a-1}\left(1 - \frac{(a)_{n+1}}{(c-1)_{n+1}}\right),$$

with limiting form $a(\psi(a+n+1)-\psi(a)) = \frac{a}{(a)_{n+1}}\frac{\mathrm{d}}{\mathrm{d}a}(a)_{n+1}$, in the case $c = a+1$.

- (15.5.16_5). Contiguous relation for Gauss hypergeometric function:

$$F(a,b;c;z) - F(a-1,b;c;z) - \frac{bz}{c}F(a,b+1;c+1;z) = 0.$$

- (17.6.4_5). Formula related to Andrews–Askey Sum ($|cq^3| < |b^2|$):

$$_2\phi_1\left(\begin{matrix}b^2,b^2/c\\cq^2\end{matrix};q^2,cq^3/b^2\right) = \frac{1}{2b}\frac{(b^2,q;q^2)_\infty}{(cq^2,cq/b^2;q^2)_\infty}\left(\frac{(cq/b;q)_\infty}{(b;q)_\infty} - \frac{(-cq/b;q)_\infty}{(-b;q)_\infty}\right).$$

- (17.8.8). Sum related to Andrews–Askey Sum ($|cq^2| < |b^2|$):

$$_2\psi_2\left(\begin{matrix}b^2,b^2/c\\q,cq\end{matrix};q^2,cq^2/b^2\right) = \frac{1}{2}\frac{(q^2,qb^2,q/b^2,cq/b^2;q^2)_\infty}{(cq,cq^2/b^2,q^2/b^2,c/b^2;q^2)_\infty}\left(\frac{(c\sqrt{q}/b;q)_\infty}{(b\sqrt{q};q)_\infty} + \frac{(-c\sqrt{q}/b;q)_\infty}{(-b\sqrt{q};q)_\infty}\right).$$

- (17.9.3_5). $_2\phi_1$ expressed as sum of two $_3\phi_2$'s with a vanishing denominator parameter:

$$_2\phi_1\left(\begin{matrix}a,b\\c\end{matrix};q,z\right) = \frac{(c/a,c/b;q)_\infty}{(c,c/(ab);q)_\infty}{}_3\phi_2\left(\begin{matrix}a,b,abz/c\\qab/c,0\end{matrix};q,q\right) + \frac{(a,b,abz/c;q)_\infty}{(c,ab/c,z;q)_\infty}{}_3\phi_2\left(\begin{matrix}c/a,c/b,z\\qc/(ab),0\end{matrix};q,q\right).$$

- (15.4.34). Gauss hypergeometric sum:

$$F\left(3a,a;2a;e^{i\pi/3}\right) = \sqrt{\pi}\,e^{i\pi a/2}\frac{2^{2a}\Gamma\left(\frac{1}{2}+a\right)}{3^{(3a+1)/2}}\left(\frac{1}{\Gamma\left(\frac{1}{3}+a\right)\Gamma\left(\frac{2}{3}\right)} + \frac{1}{\Gamma\left(\frac{2}{3}+a\right)\Gamma\left(\frac{1}{3}\right)}\right),$$

where the limit interpretation (15.2.6) needs to be taken when $a = 0, -1, -2, \ldots$.

- (19.5.4_1), (19.5.4_2), (19.5.4_3). Legendre's incomplete elliptic integral of the 1$^{\text{st}}$, 2$^{\text{nd}}$ and 3$^{\text{rd}}$ kinds given expressed in terms of the Appell function of the first kind:

$$F(\phi,k) = \sum_{m=0}^{\infty}\frac{\left(\frac{1}{2}\right)_m\sin^{2m+1}\phi}{(2m+1)m!}{}_2F_1\left(\begin{matrix}m+\frac{1}{2},\frac{1}{2}\\m+\frac{3}{2}\end{matrix};\sin^2\phi\right)k^{2m} = \sin\phi\,F_1\left(\frac{1}{2};\frac{1}{2},\frac{1}{2};\frac{3}{2};\sin^2\phi,k^2\sin^2\phi\right),$$

$$E(\phi,k) = \sum_{m=0}^{\infty}\frac{\left(-\frac{1}{2}\right)_m\sin^{2m+1}\phi}{(2m+1)m!}{}_2F_1\left(\begin{matrix}m+\frac{1}{2},\frac{1}{2}\\m+\frac{3}{2}\end{matrix};\sin^2\phi\right)k^{2m} = \sin\phi\,F_1\left(\frac{1}{2};\frac{1}{2},-\frac{1}{2};\frac{3}{2};\sin^2\phi,k^2\sin^2\phi\right),$$

$$\Pi(\phi,\alpha^2,k) = \sum_{m=0}^{\infty}\frac{\left(\frac{1}{2}\right)_m\sin^{2m+1}\phi}{(2m+1)m!}F_1\left(m+\frac{1}{2};\frac{1}{2},1;m+\frac{3}{2};\sin^2\phi,\alpha^2\sin^2\phi\right)k^{2m}.$$

- (5.15.9). Asymptotic expansion of the polygamma function:

$$\psi^{(n)}(z) \sim (-1)^{n-1}\left(\frac{(n-1)!}{z^n} + \frac{n!}{2z^{n+1}} + \sum_{k=1}^{\infty}\frac{(2k+n-1)!}{(2k)!}\frac{B_{2k}}{z^{2k+n}}\right),$$

as $z \to \infty$ in $|\mathrm{ph}\,z| \le \pi - \delta$, and B_{2k} are the Bernoulli numbers.

Figure 96. *Examples of equations that have been added to the DLMF since its original release.*

to published proofs for DLMF formulae at the equation level is well underway. This audit has been completed for Chapter 9 (Airy and Related Functions) and Chapter 25 (Zeta and Related Functions, with validation provided by Gergő Nemes) and is actively continuing for Chapters 1-5 and 22-30. Furthermore, inherited metadata at the subsection and section levels has been fully deployed.

As LaTeXML has been for some time, the DLMF is now fully embedded in the regular use of GitHub's capabilities for ongoing maintenance, as well as tracking changes and enhancements. In particular, changes to the DLMF are now implemented via GitHub issues and targeted pull requests of GitHub branches, each of which are reviewed by other project members before merging with the master branch in full adoption. All changes are reviewed and discussed by the DLMF team of at weekly DLMF meetings of the editorial staff prior to their appearance which occurs in quarterly DLMF revisions.

There have been notable additional advances during the current reporting period.

- When the DLMF is viewed using a MathML accessible browser, one may now use the mouse to "hover" over particular objects and an English description of the object is revealed.

- A significant number of mathematical formulas, errata and new mathematical information have been provided, many of which originated from the DLMF readership, validation staff, and contributors. Furthermore, mathematical constraints and symbols associated with equations and in the text, have been improved, clarified, fixed, or disambiguated.

- Proof sketches in Chapters 9, 25, and elsewhere are now carefully differentiated at the equation level providing useful metadata for the formula origins.

- Improved notations and updated citations have been introduced.

[1] M. Abramowitz and I. Stegun, eds. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Applied Mathematics Series 55, National Bureau of Standards, Washington, DC 1964.

[2] F. Olver, D. Lozier, R. Boisvert and C. Clark, eds. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.

Visualization of Complex Functions Data

Bonita Saunders
Bruce Miller
Brian Antonishek (NIST EL)
Sandy Ressler
Qiming Wang (NIST retired)

The definitions, recurrence relations, differential equations, integrals, and asymptotic expansions of the DLMF[20] provide crucial information for understanding complex mathematical functions arising in application areas of the mathematical and physical sciences, but interactive visualizations can also provide additional clarity as indicated in Figure 97.

Since the DLMF provides the impetus for this project, our first priority is to ensure that the quality and accessibility of its visualizations are maintained. Recent updates of JavaScript and X3DOM[21] graphics libraries associated with the generation of our visualizations introduced anomalies in the color maps. The first question was whether the new maps were actually more accurate representations, but this was not supported by preliminary examinations of the data files, plots of similar data on other websites, or plots in Jahnke, Emde, and Losch [4]. By backing out of several updates we determined the earliest versions of X3DOM that produced the altered color maps. Ideally, we want to use the latest update available. Therefore, we will look at the X3DOM changelog available through GitHub for a possible "culprit", and most likely, contact X3DOM developers for further clarification.

Another priority is addressing changes suggested by DLMF users, other members of the DLMF editorial staff, and chapter authors. Based on feedback received, modifications and new figures related to the Lambert *W*-Function in Chapter 4 on Elementary Functions will soon be added. New figures that will be part of a substantial update to the chapter on Orthogonal Polynomials will appear later.

We also continue to look for opportunities to generalize our work to benefit the larger research community. Our current work on adaptive meshes to improve our underlying computational grids supports our design process, but it may also interest other researchers in the fields of mesh generation, optimization, approximation theory, or any area related to the design of curves and surfaces for mathematical or physical applications.

A paper on our mesh generation and 3D Web work was published in the 50th anniversary proceedings for the National Association of Mathematicians [1]. Also, two talks discussing the context of the graphics work in the DLMF's history were presented. The talk [2], given to a history of mathematics class, inspired a student to do a class project on the Airy function. The second talk was the 2021 Etta Z. Falconer Lecture, an honor sponsored by the Association for Women in Mathematics (AWM) and the Mathematical Association of America (MAA) at MathFest [3]

As time permits, we will explore opportunities to increase user visibility. Some DLMF chapters display chapter related thumbnail images on the title page. Creating a gallery of images for all chapters would add visual interest to the DLMF. Each image would link to a short descriptive sidebar that could include links to the chapter's application section or related function visualizations. New and updated DLMF chapters currently under development can provide a test bed for this work.

[1] B. Saunders. Complex Variables, Mesh Generation, and 3D Web Graphics: Research and Technology Behind the Visualizations in the NIST Digital Library of Mathematical Functions. In *Proceedings of the Golden Anniversary Celebration of the National Association of Mathematicians,* AMS Contemporary Mathematics Series **759**, (2020). DOI: 10.1090/conm/759/15272

[2] B. Saunders, "From Abramowitz and Stegun to the NIST Digital Library of Mathematical Functions and Beyond." History of Mathematics Class, Prof. Hortensia Soto, Colorado State University, Fort Collins, CO, Online, November 10, 2020.

---

[20] https://dlmf.nist.gov/

[21] https://www.x3dom.org/

Figure 25.3.3: Modulus of the Riemann zeta function $|\zeta(x+iy)|$, $-4 \le x \le 4$, $-10 \le y \le 40$.
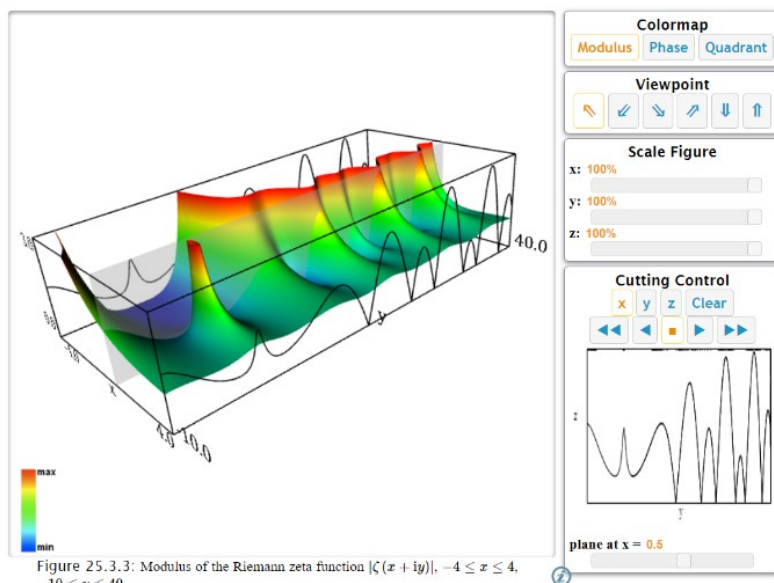
Figure 97. *DLMF visualizations provide several options for exploring mathematical function surfaces. The Riemann zeta function is shown with cutting plane and pop-up control window illustrating nontrivial zeros. Other options allow users to change the color map, examine different viewpoints and scale the figure in various coordinate directions.*

[3]   B. Saunders, "Complex Functions, Mesh Generation, and Hidden Figures in the NIST Digital Library of Mathematical Functions." AWM-MAA Etta Zuber Falconer Lecture, MAA MathFest 2021, Online, August 6, 2021.

[4]   E. Jahnke, F. Emde, and F. Losch. *Tables of Higher Functions*. 6th Edition, B. G. Teubner Verlagsgesellschaft mbH, Stuttgart, 1960 (Bilingual).

NIST Digital Repository of Mathematical Formulae

Howard S. Cohl
Marjorie McClain
Bonita V. Saunders
Abdou Youssef
Moritz Schubotz
Lisa Ritter
Andre Greiner-Petter (University of Wuppertal)
Avi Trost (Brown University)
Rajen Dey (University of California Berkeley)

The NIST Digital Repository of Mathematical Formulae (DRMF) is an online compendium of formulae for orthogonal polynomials and special functions (OPSF) designed to (a) facilitate interaction among a community of mathematicians and scientists interested in OPSF; (b) be expandable, allowing the input of new formulae from the literature; (c) provide information for related linked open data projects; (d) represent the context-free full semantic information concerning individual formulas; (e) have a user friendly, consistent, and hyperlinkable viewpoint and authoring perspective; (f) contain easily searchable mathematics; and (g) take advantage of modern MathML tools for easy-to-read, scalably rendered content-driven mathematics. The DRMF has been summarized in a series of papers [1-3].

Our DRMF implementation, previously built using MediaWiki (the wiki software used by Wikipedia), is currently in migration to a different software platform, namely the platform used by the NIST Digital Library of Mathematical Functions (DLMF). NIST NRC postdoc Lisa Ritter will be collaborating on this transition. See Figure 98 for the current draft of the DRMF home page, and Figure 99 for a sample DRMF formula page.

A key asset in the development of DRMF context free semantic content is the utilization of a set of LaTeX macros and macro call functionality created by Bruce Miller (ACMD) to achieve the encapsulation of semantic information within the NIST Digital Library of Mathematical Functions (DLMF) [4]. These macros give us the capability to tie LaTeX commands in a mostly unambiguous way to mathematical functions defined in an OPSF context. There are currently 540 DLMF LaTeX macros, as well as an additional 156 which have been created specifically for the DRMF. Most DLMF macros have at least one DLMF web page associated with them. One goal is to have definition pages for all additional DRMF macros. The use of DLMF and DRMF macros guarantees mathematical and structural consistency throughout the DRMF. We refer to LaTeX source with incorporated DLMF and DRMF macros as semantic LaTeX.

DRMF formula seeding is currently focused on:

a. Koekoek, Lesky, and Swarttouw (KLS) chapters 1 (Definitions and Miscellaneous Formulas), 9 (Hypergeometric Orthogonal Polynomials), and 14 (Basic Hypergeometric Orthogonal Polynomials) [5],

b. Koornwinder KLS addendum LaTeX data [5],

c. Wolfram Computational Knowledge of Continued Fractions Project (eCF) [3],

d. Continued Fractions for Special Function (CFSF) Maple dataset hosted by the University of Antwerp [3,7],

e. Bateman Manuscript Project (BMP) books [8], and

f. Magnus, Oberhettinger, and Soni (MOS) books [3, 9].

For these seed projects, we have developed Python and Java software to incorporate DLMF and DRMF macros into the corresponding LaTeX source. Our coding efforts have also focused on extracting formula data from LaTeX source as well as generating DRMF semantic LaTeX. We have developed Java software for the seeding of the eCF and CFSF projects which involve conversion from Mathematica and Maple to DLMF and DRMF macro incorporated semantic LaTeX [3].

In 2014 the DRMF Project obtained permission and license to use BMP material as seed content for the DRMF from Adam Cochran, Associate General Counsel of Caltech. Caltech has loaned us copies of the BMP. In February 2018, we received permission and license to use the KLS and MOS material as seed content for the DRMF from Springer Nature. With soon to be NIST NRC postdoc Camilo Montoya, we plan on implementing the BMP and MOS datasets using mathematical optical character recognition software to obtain LaTeX source using software developed with MathType.

Current and future DRMF MediaWiki development projects include the production of formula output representations (such as semantic LaTeX, MathML, Mathematica, Maple, and Sage); incorporation of sophisticated DLMF and DRMF macro related formula search; and the development of capabilities for user community formula input. In this vein, A. Youssef has written a grammar-based mathematical language processor (MLP) that uses JavaCC to parse mathematical LaTeX expressions [10]. Based on the MLP, A. Greiner-Petter has developed a Java tool referred to as LaCASt to convert mathematical LaTeX expressions, which contain DLMF and DRMF macros, to a given computer algebra system source format. This Java tool provides further information of the conversion about possible ambiguities and differences in definitions, domains and branch cuts between the semantic LaTeX source and the CAS source. Furthermore, it is designed to be easily extendable to other computer algebra systems and currently supports Maple and Mathematica input sources.

Some of our additional activities of the past year include the following:

a. NIST ACMD summer intern Avi Trost worked on the project "Semantic Enhancement for Wronskians and Prime Derivative Notation."

b. NIST ACMD summer intern Rajen Dey worked on the project "Semantic Enhancement for Sums and Products."
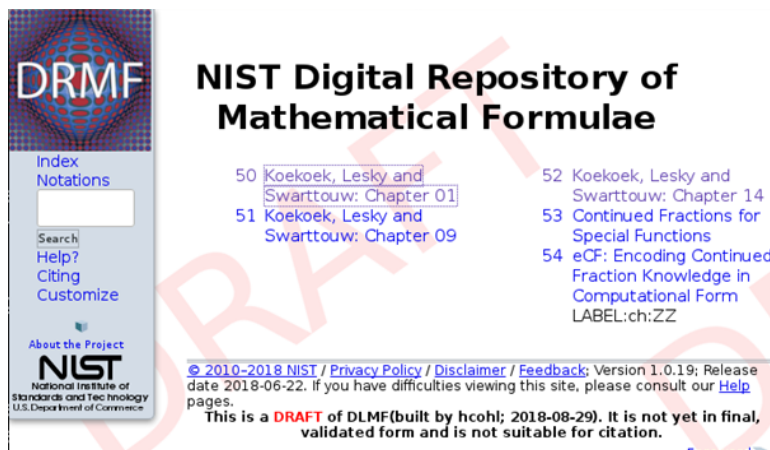
Figure 98. *Draft of the DRMF home page displaying the current table of contents.*

c. Cohl was a member of the program committee for the 14th Conference on Intelligent Computer Mathematics (CICM 2021), Timisoara, Romania, held virtually in July 2021.

d. In [11] which focuses on LaCASt, we present a first comprehensive approach to verify a digital mathematical and two computer algebra systems with one another by converting mathematical expressions from one system to the other. This is accomplished by our development of LaCASt, which translates formulae from the NIST Digital Library of Mathematical Functions to the computer algebra systems Maple and Mathematica. This tool will be actively used in DRMF.

e. In [12] we reported on the recently launched zbMATH Links API and discussed its potential based on its initial link partner, the NIST DLMF.

f. The KLS datasets have been uploaded to our DLMF platform as well as the CFSF and eCF datasets.

By working with Andrea Fisher-Scherer, Rights Administrator, Artists Rights Society, New York, NY, we have received permission from Foundation Vasarely, to use an image of one of Victor Vasarely's paintings as the DRMF logo; see Figure 98.

[1] H. S. Cohl, M. A. McClain, B. V. Saunders, M. Schubotz and J. C. Williams. Digital Repository of Mathematical Formulae. *Lecture Notes in Artificial Intelligence* **8543** (2014), Proceedings of the Conferences on Intelligent Computer Mathematics 2014, Coimbra, Portugal, July 7-11, 2014, (S. M. Watt, J. H. Davenport, A. P. Sexton, P. Sojka, and J. Urban, eds.), Springer, 419-422.

[2] H. S. Cohl, M. Schubotz, M. A. McClain, B. V. Saunders, Cherry Y. Zou, Azeem S. Mohammed, and Alex A. Danoff. Growing the Digital Repository of Mathematical Formulae with Generic LaTeX Sources. *Lecture Notes in Artificial Intelligence* **9150** (2015), Proceedings of the Conference on Intelligent Computer Mathematics 2015, Washington DC, USA, July 13-17, 2015, (M. Kerber, J.

Carette, C. Kaliszyk, F. Rabe, and V. Sorge, eds.), Springer, 280-287.

[3] H. S. Cohl, M. Schubotz, A. Youssef, A. Greiner-Petter, J. Gerhard, B. V. Saunders, M. A. McClain, J. Bang, and K. Chen. Semantic Preserving Bijective Mappings of Mathematical Formulae between Word Processors and Computer Algebra Systems. *Lecture Notes in Computer Science* **10383** (2017), Proceedings of the Conference on Intelligent Computer Mathematics 2017, Edinburgh, Scotland, U.K., July 17-21, 2017, (H. Geuvers, M. England, O. Hasan, F. Rabe, O. Teschke, eds.), Springer, 115-131.

[4] B. Miller. "Drafting DLMF Content Dictionaries." Open-Math Workshop, 9th Conference on Intelligent Computer Mathematics, CICM 2016, Bialystok, Poland.

[5] R. Koekoek, P. A. Lesky, and R. F. Swarttouw. *Hypergeometric Orthogonal Polynomials and their q-Analogues*. Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2010.

[6] T. H. Koornwinder. Additions to the Formula Lists in Hypergeometric Orthogonal Polynomials and their q-analogues by Koekoek, Lesky and Swarttouw. arXiv:1401.0815, June 2015.

[7] A. Cuyt, V. Petersen, H. Waadeland, W. B. Jones, F. Backeljauw, C. Bonan-Hamada, and S. Becuwe. *Handbook of Continued Fractions for Special Functions*. Springer, New York, 2008.

[8] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi. *Higher Transcendental Functions*. Vols. I, II, III, Robert E. Krieger Publishing Co., Melbourne, FL, 1981.

[9] H. S. Cohl, A. Greiner-Petter, and M. Schubotz. Automated Symbolic and Numerical Testing of DLMF Formulae using Computer Algebra Systems. *Lecture Notes in Computer Science* **11006** (2018), Proceedings of the Conference on Intelligent Computer Mathematics 2018, Hagenberg, Austria, August 13-17, 2018, (F. Rabe, W. Farmer, G.O. Passmore, A. Youssef, eds.), Springer, 39-52.

[10] A. Youssef. Part-of-Math Tagging and Applications. *Lecture Notes in Computer Science* **10383** (2017), Proceedings of the Conference on Intelligent Computer Mathematics 2017, Edinburgh, Scotland, U.K., July 17-21, 2017, (H. Geuvers, M. England, O. Hasan, F. Rabe, O. Teschke, eds.), Springer, 356-374.

[11] A. Greiner-Petter, M. Schubotz, H. S. Cohl, M. Schubotz, A. Aizawa, B. Gipp, A. Trost, and R. Dey. Comparative Verification of the Digital Library of Mathematical Functions and Computer Algebra Systems. International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2022), to appear.

[12] H. S. Cohl, O. Teschke, and M. Schubotz. Connecting Islands: Bridging zbMATH and DLMF with Scholix, A Blueprint for Connecting Expert Knowledge Systems, European Mathematical Society Magazine **120** (2021), 66-67.

Figure 99. *Sample DRMF page, taken from the KLS Chapter 1 dataset.*

DLMF Standard Reference Tables on Demand

Bonita Saunders
Bruce Miller
Marjorie McClain
Annie Cuyt (University of Antwerp)
Stefan Becuwe (University of Antwerp)
Franky Backeljauw (University of Antwerp)
Sean Brooks (Coppin State University)
Ron Buckmire (Occidental College)
Rachel Vincent-Finley (Southern U. and A&M College)
Christopher Schanzle

http://dlmftables.uantwerpen.be/

Although reliable computing machines, computer algebra systems, and multiple precision computational packages have diminished the need for tables of reference values for computing function values by interpolation, mathematical and physical scientists, numerical analysts, and software developers still need a way to test software for computing mathematical functions. DLMF Standard Reference Tables on Demand

Figure 100. *DLMF Tables generates tables of special function values at user specified precision. Users input real values and parameters where the function is to be evaluated. The user may request output in interval mode, where the output is shown as a table of intervals that bound the true results or may request output in one of several rounding modes. Users may also choose to compare their own table of values to the reference values generated by the system.*

(DLMF Tables) project is a collaboration between ACMD and the University of Antwerp Computational Mathematics Research Group (CMA) [1-5] to address this problem. The goal is to develop an online system where users can generate tables of special function values at user-specified precision with an error certification to test their own algorithms or confirm the accuracy of results from a commercial or publicly available package.

The DLMF Tables team has developed a beta site at the University of Antwerp, based on CMA's MpIeee, a multiple precision IEEE 754/854 compliant C++ floating point arithmetic library. Ultimately, the goal is a permanent NIST location accessible from the NIST Digital Library of Mathematical Functions (DLMF), but this will require a significant ramping up of ACMD capabilities in the field of validated or reliable computation of special functions. A needed boost to this effort occurred in 2020 when B. Saunders was offered the opportunity to lead a project at the inaugural African Diaspora Joint Mathematics Workshop[22] (ADJOINT 2020) sponsored by the Mathematical Sciences Research Institute (MSRI) in Berkeley, California [6]. The aim of the two-week workshops is to encourage collaborative work in

the mathematical sciences between established and less experienced Ph.D. researchers, especially US citizens from the African Diaspora, that is, descendants of people from Sub-Saharan African countries. For the 2020 workshop, held virtually due to the pandemic, Saunders chose a project that introduced her three member ADJOINT group to Validated Numerical Computations of Mathematical Functions, a field that includes the error analysis and code development needed to compute certifiably accurate function values. The group members, from various universities, were supported by MSRI through NSF, NSA, and Sloan grants[23] for more than a year. Post-workshop contact has continued with online meetings to discuss and verify applicable results from the literature, including Higham [7] and CMA [1-5]. The ADJOINT group's hard work led to three invited talks at the 2021 Joint Mathematics Meetings [8-10] and a featured MSRI colloquium talk [11]. Two more talks are scheduled for minisymposia at the 2022 Joint Mathematics Meetings, and a paper co-authored by the group with members of the ACMD and U. Antwerp groups is in review [12].

ACMD has primarily been responsible for the front-end interface for DLMF Tables, while CMA designed and maintains the back-end computational engine built around the error analysis and MpIeee library. The work with the ADJOINT group marks the first step toward merging these capabilities so that CMA's function computation codes can be maintained and eventually, further developed here at NIST.

[1] F. Backeljauw, S. Becuwe, A. Cuyt, J. Van Deun, and D. Lozier. Validated Evaluation of Special Mathematical Functions. *Science of Computer Programming* **10** (2013), 1016.

[2] M. Colman, A. Cuyt, and J. Van Deun, Validated Computation of Certain Hypergeometric Functions. *ACM Transactions on Mathematical Software* **38**:2 (January 2012), 11.

[3] F. Backeljauw. A Library for Radix-independent Multiprecision IEEE-compliant Floating-point Arithmetic. Technical Report 2009-01, Department of Mathematics and Computer Science, Universiteit Antwerpen, 2009.

[4] A. Cuyt, V. B. Petersen, B. Verdonk, H. Waadeland, and W. B. Jones. *Handbook of Continued Fractions for Special Functions*. Springer, New York, 2008.

[5] A. Cuyt, B. Verdonk, and H. Waadeland. Efficient and Reliable Multiprecision Implementation of Elementary and Special Functions. *SIAM Journal of Scientific Computing* **28** (2006), 1437-1462.

[6] B. Saunders, S. Brooks, R. Buckmire, and R. Vincent-Finley. "Validated Numerical Computation of Mathematical Functions." African Diaspora Joint Mathematics Workshop (ADJOINT) 2020, Online, June 26, 2020.

[7]   N. Higham. *Accuracy and Stability of Numerical Algorithms.* Second edition. Society for Industrial and Applied Mathematics, Philadelphia, 2002.

[8]   B. Saunders, S. Brooks, R. Buckmire, and R. Vincent-Finley. "Validated Computation of Special Functions I: Overview." ADJOINT Research Showcase, 2021 Joint Mathematics Meetings, Online, January 9, 2021.

[9]   R. Buckmire, S. Brooks, R. Vincent-Finley, and B. Saunders. "Validated Computation of Special Functions II: Error Analysis." ADJOINT Research Showcase, 2021 Joint Mathematics Meetings, Online, January 9, 2021.

[10] R. Vincent-Finley, S. Brooks, R. Buckmire, and B. Saunders. "Validated Computation of Special Functions III: DLMF Tables." ADJOINT Research Showcase, 2021 Joint Mathematics Meetings, Online, January 9, 2021.

[11] R. Vincent-Finley, S. Brooks, R. Buckmire, and B. Saunders. "Validated Computation of Special Mathematical Functions." MSRI ADJOINT Colloquium, Online, April 27, 2021.

[12] B. Saunders, S. Brooks, R. Vincent-Finley, R. Buckmire, F. Backeljauw, S. Becuwe, B. Miller, M. McClain, A. Cuyt. Validated Computation of Special Mathematical Functions. In review.

Scientific Document Corpora for Natural and Mathematical Language Research

Bruce Miller
Deyan Ginev (Chakra Consulting)
Tom Wiesing (University of Erlangen, Germany)

Machine learning is currently a very active research area, both in its theoretical underpinnings and technology, as well as its application to find, understand and reuse information. We would like to see the application of these methods to scientific documents, unique in their style of natural language used, as well as the extensive use of mathematical notation. For this to happen, large collections of scientific documents are needed, first for training and then for mining. To that end, we have been applying our LaTeXML tool to the massive corpus at arXiv.org[24]. To demonstrate the utility of this dataset, we have carried out initial experiments on statement classification using that data set.

LaTeXML[25] was originally developed for use in converting the LaTeX sources of the NIST Digital Library of Mathematical Functions (DLMF) into web format, namely HTML and MathML. Most of arXiv.org is also in LaTeX format, albeit using a significantly less disciplined markup style and a wide variety of support packages, including uncommon or very complex ones. We have continued to develop LaTeXML, making it more robust and more accurately simulating the TeX engine. The internal infrastructure for mathematical semantics has also been enhanced, allowing us to process material using the rich siunitx and physics packages, preserving both the displayed formulae as well as their mathematical intent. Moreover, we have refined the encoding of mathematics into the textual "lexemes" needed by machine learning methods to reflect each symbol's role more accurately in the text. We have further increased the coverage of LaTeX packages, as well as improving the appearance of the generated HTML to better reflect the author's intentions. Given the increasing variety of markup used on arXiv, as well as its fast growth, effort is required just to stay even. This year was a very busy one on arXiv, with the number of documents increasing to 1.8M from 1.6M last year. We are in the midst of reprocessing the entire current corpus into HTML+MathML and anticipating maintaining our 70 % success rate (i.e., documents producing, at worst, warnings).

Although the markup used in LaTeX documents found in the wild seldom emphasizes semantics, there are nevertheless a number of macros and environments which can be taken to reflect the author's intent. Such markup may indicate theorems, proofs, definitions as well as introductions and acknowledgments and so on. We were able to extract some 10M such annotated paragraphs in 50 categories from our conversion of arXiv [1]. Using 80 % of the documents as a training set, we discovered that many categories were too similar. For example, theorems, lemmas, and propositions share language patterns. Combining these "confusion nests" yielded 13 clear cut categories into which the paragraphs could be reliably categorized with a 0.91 F1 score.

[1]   D. Ginev and B. R. Miller. Scientific Statement Classification over arXiv.org. Preprint arXiv:1908.10993, 2019.

Towards a Machine-Readable Digital Library of Mathematical Functions

Bruce Miller
Deyan Ginev (Chakra Consulting)
Tom Wiesing (University of Erlangen, Germany)

The principal initial goal of the Digital Library of Mathematical Functions (DLMF) project was to present useful definitions, relationships, properties, and other information about special functions for use by scientists and engineers; that is, directly to people. In the medium term, we have striven to develop procedures for DLMF enhancements, maintenance, and corrections. While we

---

[24] https://arxiv.org/

[25] https://math.nist.gov/~BMiller/LaTeXML/

feel we have been successful so far, the longer-term goal of putting all data into a fully machine-readable form has been slower in coming, given the scale and complexity of the material. This form would improve the findability, accessibility, reuse, and, in particular, computability of the mathematics. We are particularly motivated by the efforts of the Math Working Group of the World Wide Web Consortium to develop annotations for Presentation MathML that will assist in making Math on the Web accessible to those with a variety of disabilities. We are using our past experience with mathematical semantics to assist in that effort.

While one can imagine a sufficiently rich set of semantic LaTeX macros allowing the author to specify not only the printed form of the formula, but completely specify the exact semantic meaning of every symbol and expression as well, this approach is so cumbersome as to be impractical. Conversely, one might imagine eventual artificial intelligence able to infer the meaning from any given TeX markup along with the document's full text and context, but that is not yet feasible. We seek the sweet spot between those extremes where authors supply just enough clarity via semantic macros so that follow-up analysis can resolve remaining ambiguities.

Fully implementing this strategy is still a work in progress, particularly the latter inference aspect, even scaled down. Progress is being made in several fundamental components, however, and there are many benefits to even partly annotated and disambiguated data in terms of web services and information discovery.

Short of fully explicit author markup, the mathematical formula must be parsed at least to determine the expression structure. We are cataloging the range of mathematical notations both within and outside of the DLMF and have continued to enhance the mathematical grammar to increase coverage. Additionally, we are exploring parsing frameworks which provide multiple parses to cover cases where the notation has several possible interpretations. Techniques for pruning these possibilities according to consistency or partial knowledge are being developed.

We have developed and made use internally of a set of semantic macros for DLMF's special functions and the more notationally complex concepts. Focusing here resolves some of the most glaring ambiguities, both syntactic and semantic. (There are so many different f's!) This even provides a convenience to authors to simplify and regularize the typing required. As the benefits of machine-readable mathematical data on the web become realized and more apparent, a much broader audience of authors is likely to find these tools useful. For our macro set to be more palatable to that audience, it needs to be appropriately simplified and yet generalized. We are working towards this with the goal of publishing and making available our semantic macro set.

A common basis is also necessary for effective interoperability. Are the functions treated in DLMF actually the same as those in Mathematica, or Maple or the NAG libraries? As it turns out, usually they are, but sometimes they are not, and the differences can be subtle. Establishing these correspondences is exactly the purpose of the Special Function Concordance activity of the International Mathematical Knowledge Trust (IMKT) established by the Global Digital Mathematical Library (GDML) working group of the International Mathematical Union. We are participating in this effort to establish this concordance between the various special functions, in all their flavors, as used within various handbooks, such as the DLMF, and software systems, such as Mathematica, Maple and NAG, to assure interoperability.

For our part, we are developing a catalog of the special functions as defined in DLMF. Obviously, we start by listing those formulas which we consider to be the defining ones for each function. However, the key is to focus on those aspects, those choices, that potentially distinguish our version of a particular function from those of other systems. Choices of argument conventions (e.g., elliptic functions $\operatorname{sn}(u,k)$ vs. $\operatorname{sn}(u,m)$) are significant. Patterns of singularities and type signatures help distinguish different extensions and generalizations of functions. A trickier class of difference are the choices made for the location of branch-cuts or which value is taken on the cuts, or indeed the choice made to avoid them entirely as multi-valued functions. Given this set of distinguishing features, we are collecting and encoding that information for each of the DLMF's functions in the form of OpenMath Content Dictionaries.

Fundamental Solutions and Formulas for Special Functions and Orthogonal Polynomials

Howard S. Cohl
Lisa Ritter
Roberto S. Costas-Santos (University of Alcala)
Hans Volkmer (University of Wisconsin-Milwaukee)
Gestur Olafsson (Louisiana State University)
Mourad E. H. Ismail (University of Central Florida)
Tom H. Koornwinder (University of Amsterdam)
James Lawrence (George Mason University)
Jessica E. Hirtenstein (University of California Davis)
Philbert R. Hwang (University of Maryland)
Linus Ge (University of Rochester)
Justin Park (Massachusetts Institute of Technology)
Jason Zhao (University of California Los Angeles)

The concept of a function expresses the idea that one quantity (the input) completely determines another quantity (the output). Our research concerns special functions and orthogonal polynomials. A special function is a function that has appeared in the mathematical

sciences so often that it has been given a name. Green's functions (named after the British mathematician George Green, who first developed the concept in the 1830s) describe the influence of linear natural phenomena such as electromagnetism, gravity, heat, and waves. For example, in electrostatics, a Green's function describes the influence of a point charge, called the source, over all of space. The inputs for fundamental solutions (Green's functions) are all of space (apart from a singular region), and the output is the "force" exerted from the point throughout space. Green's functions are fundamental to the study of inhomogeneous partial differential equations and are powerful in that they provide a mechanism for obtaining their solutions.

We investigate fundamental solutions of linear partial differential equations on highly symmetric Riemannian manifolds (harmonic, rank-one symmetric spaces) such as real, complex, quarternionic, and octonionic Euclidean, hyperbolic, and projective spaces. Our recent focus has been on applications of fundamental solutions for linear elliptic partial differential operators on spaces of constant curvature. With Olafsson and soon to be NIST/NRC postdoc, Camilo Montoya, we will complete the study of fundamental solutions for the Laplace-Beltrami operator on rank one symmetric spaces of compact and noncompact type [1]. Cohl, working with NRC postdoc Lisa Ritter have prepared a paper with previous SURF student Jessica Hirtenstein and Jim Lawrence on deriving Gegenbauer expansions and addition theorems for binomial and logarithmic fundamental solutions of the polyharmonic operator in even-dimensional Euclidean space in Vilenkin polyspherical coordinates for powers of the Laplacian greater than or equal to the dimension divided by two [2].

In the following works, we expand on the properties of terminating and nonterminating generalized and basic hypergeometric functions to study themselves and their respective generalized and basic hypergeometric orthogonal polynomials (hereafter orthogonal polynomials) in the Askey and q-Askey schemes. By utilizing connection relations, we have computed generalizations of generalized and basic hypergeometric orthogonal polynomial generating functions as well as corresponding definite integrals using orthogonality [3].

We are also interested in the fundamental transformation, representation (symmetry) properties of the special functions and orthogonal polynomials which one often encounters in applied mathematics and mathematical physics. In this regard, in conjunction with Park and Volkmer, we recently published a paper which computes all Gauss hypergeometric representations of the Ferrers function of the second kind by starting with the full list of 18 Gauss hypergeometric representations for the associated Legendre function of the second kind [4].

Working with Costas-Santos, we have developed a series of papers which describe transformation and representation theory of symmetric basic hypergeometric

orthogonal polynomials, namely the Askey–Wilson polynomials (4 symmetric free parameters (sfp)) and its symmetric subfamilies, the continuous dual q-Hahn polynomials (3 sfp), the Al-Salam–Chihara polynomials (2 sfp), the continuous big q-Hermite polynomials and the continuous q-Hermite polynomials, and their q-inverse analogues. We have also been studying continuous q-Jacobi and continuous q-ultraspherical / Rogers polynomials. With Costas-Santos and Ge, we derived all terminating basic hypergeometric representation and transformation properties of the Askey–Wilson polynomials [5].

In a follow-up paper with Costas-Santos, we examine the transformation and representation properties of the Askey–Wilson polynomials with a focus on the terminating very-well poised $_8W_7$ representations. We examine the representation relation to the order of the well-known symmetry group of the terminating $_4f_3$ Askey–Wilson representations given by the symmetric group S_6 with respect to their and transformation inversion symmetries [6].

With Costas-Santos and Ge we are also finishing up an analogous analysis on the symmetric and q-inverse symmetric sub-families of the Askey–Wilson polynomials [7]. We have also investigated the use of integral representations for nonterminating basic hypergeometric functions and orthogonal polynomials [8].

With Costas-Santos we studied the relation between the Ferrers function of the first kind and Gegenbauer polynomials to derive a collection of new formulas [9].

With Ismail and Ritter, we derive 5-term contiguous relations for the linearization coefficients of generalized and basic hypergeometric orthogonal polynomials such as Laguerre, Gegenbauer, Hermite, Jacobi, continuous q-ultraspherical/Rogers, and continuous q-Jacobi polynomials [10].

With Ritter and Bodenshatz, we study representations of symmetric elliptic integrals in terms of incomplete Legendre elliptic integrals of the first, second and third kinds [11].

In an ongoing project which performs generalized hypergeometric and basic hypergeometric analysis of generating functions and linearization formulae for orthogonal polynomials, we have several ongoing projects. With Costas-Santos, Hwang and Wakhare, our series-rearrangement technique is extended to generalizations of other generating functions for basic hypergeometric orthogonal polynomials in [3]. Here, we derive generalizations of generating functions for Askey–Wilson, q-ultraspherical/Rogers, q-Laguerre, and little q-Laguerre/Wall polynomials.

With Costas-Santos and Zhao, we derive generalized linearization formulas for generalized and basic hypergeometric orthogonal polynomials by applying connection relations to them in [12]. Here, we generalize linearization formulae for continuous q-ultraspherical/Rogers, Jacobi, and continuous q-Jacobi

polynomials, Gegenbauer and Laguerre polynomials. With Costas-Santos and Koornwinder, we are investigating dual addition theorems, product formulas and other related formulas for the continuous *q*-Jacobi polynomials [13].

With Costas-Santos, generalizing some classical formulae for these functions, we derived multi-integral representations for the associated Legendre and Ferrers functions of the first and second kind [14]. We are also continuing this project with Jacobi functions of the first and second kind.

Cohl served on the Scientific and Local Organizing committees for the Orthogonal Polynomials and Special Functions Summer School 6 (OPSF-S6) workshop which was held on June 17-23, 2016, at the Norbert Wiener Center for Harmonic Analysis and Applications at the University of Maryland. The OPSF-S6 lecture notes were published by Cambridge University Press, edited by Cohl and Ismail, including lecture notes by Duran, Ismail, Koelink, Rosengren, and Zeng [15].

Richard A. Askey, one of the giants in the field of OPSF, passed away on October 9, 2019. Cohl, Ismail, and Wu have edited a memorial article for Dick Askey which was published in *The Notices of the American Mathematical Society* [16]. With 63 colleagues, Cohl and Ismail prepared a *Liber Amicorum for Dick Askey* which was presented to him and his family at an event held on September 4, 2019, at his hometown in Madison, Wisconsin. Cohl and Ismail have submitted an extended version of the Liber Amicorum composed of the remembrances of 83 of his colleagues which will be published online with *Celebratio Mathematica* [17].

Cohl remains editor or co-editor for a special issue on symmetry in special functions and orthogonal polynomials in the journal *Symmetry*; special volume dedicated to the legacy of Dick Askey for *The Ramanujan Journal*, OP-SF NET, SIAM Activity Group on Orthogonal Polynomials and Special Functions; and *The Ramanujan Journal*.

[1] H. S. Cohl, C. Montoya and G. Olafsson and. Fundamental Solutions for the Laplace-Beltrami Operator on the Rank One Symmetric Spaces. In preparation.

[2] H. S. Cohl, J. E. Hirtenstein, J. Lawrence and L. Ritter. Binomial and Logarithmic Gegenbauer Expansions for the Even-dimensional Polyharmonic Equation. In review.

[3] H. S. Cohl, R. S. Costas-Santos, P. R. Hwang, and T. V. Wakhare. Generalizations of Generating Functions for Basic Hypergeometric Orthogonal Polynomials. In review.

[4] H. S. Cohl, J. Park, and H. Volkmer. Gauss Hypergeometric Representations of Ferrers Functions of the Second Kind. *Symmetry, Integrability and Geometry: Methods and Applications* **17** (2021), 053.

[5] H. S. Cohl, R. S. Costas-Santos and L. Ge, Terminating Basic Hypergeometric Representations and Transformations for the Askey–Wilson Polynomials. *Symmetry* **12**:8 (2020), 1290.

[6] H. S. Cohl and R. S. Costas-Santos. Symmetry of Terminating Basic Hypergeometric Representations of the Askey–Wilson Polynomials. *The Ramanujan Journal*, Special Volume dedicated to Dick Askey, to appear.

[7] H. S. Cohl, R. S. Costas-Santos and L. Ge. Basic Hypergeometric Transformations from Symmetric and *q*-inverse Sub-families of the Askey-Wilson Polynomials in the *q*-Askey-scheme. In preparation.

[8] H. S. Cohl and R. S. Costas-Santos. Utility of Integral Representations for Basic Hypergeometric Functions and Orthogonal Polynomials. *The Ramanujan Journal*, Special Volume Dedicated to Dick Askey, to appear.

[9] H. S. Cohl and R. S. Costas-Santos, On the Relation Between Gegenbauer Polynomials and the Ferrers Function of the First Kind. *Analysis Mathematica*, to appear.

[10] H. S. Cohl, M. E. H. Ismail, and L. Ritter. Two-dimensional Contiguous Relations for Linearization Formulae. In preparation.

[11] J. Bodenschatz, H. S. Cohl, and L. Ritter, Symmetric Elliptic Integrals in Terms of Legendre Incomplete Elliptic Integrals. In preparation.

[12] H. S. Cohl, R. S. Costas-Santos, and J. Zhao. Generalizations of Linearization Formulae for Continuous Hypergeometric Orthogonal Polynomials. In preparation.

[13] H. S. Cohl, R. S. Costas-Santos, and T. H. Koornwinder. Linearization, Dual Addition Theorem and Product Formulas for the Continuous *q*-Jacobi Polynomials. In preparation.

[14] H. S. Cohl and R. S. Costas-Santos. Multi-Integral Representations for Associated Legendre and Ferrers Functions. *Symmetry*, **12** (2020), 1598.

[15] H. S. Cohl and M. E. H. Ismail, eds. *Lectures on Orthogonal Polynomials and Special Functions, OPSFA Sixth Summer School, Norbert Wiener Center, University of Maryland, College Park, Maryland, July 11-15, 2016.* London Mathematical Society Lecture Note Series **464**, Cambridge University Press, 2020.

[16] M. E. H. Ismail, H. S. Cohl and H.-H. Wu, eds. *The Legacy of Dick Askey (1933-2019). Notices of the American Mathematical Society,* to appear.

[17] H. S. Cohl and M. E. H. Ismail, eds. Liber Amicorum, a Friendship Book for Dick Askey. *Celebratio Mathematica,* to appear.

Outreach and Diversity

ACMD staff engage in a variety of efforts that serve to educate the general public about the work of the division and to encourage students to consider careers in science and engineering. We are also involved in internal efforts to improve diversity and inclusivity, which are important for both recruitment and retention of a high-performing workforce. Some of these efforts are described here.

Student Internships in ACMD

Ronald Boisvert

ACMD is committed to helping to prepare the next generation of scientific researchers by providing internships of various types to students at each of the graduate, undergraduate, and high school levels. The NIST programs used to enable such internships include the following:

- *Foreign Guest Researcher Program.* Provides stipends to support visits of guest researchers from foreign institutions for periods of a few weeks to several years.
- *Pathways Program.* Provides temporary Federal appointments to students, typically 1 to 2 years. Allows easy conversion to full-time permanent status. (Restricted to US Citizens.)
- *Professional Research Experience Program (PREP)[26].* A cooperative agreement with nine universities[27] that provides a mechanism for NIST to support internships for students from those institutions on the Gaithersburg campus throughout the year. A similar agreement with four universities[28] exists for the NIST Boulder Labs.
- *Student Volunteer Program.* A mechanism that provides unpaid internships for students.
- *Summer High School Internship (SHIP) Program[29].* SHIP uses the Student Volunteer Program to organize a competitive summer volunteer program for high school students.
- *Summer Undergraduate Research Fellowship (SURF) Program[30].* A competitive program providing undergraduates a 10-week research experience at NIST.

Funding for all of these programs comes from the Division hosting the student. The Pathways Program, the PREP Program, and the Foreign Guest Researcher Program can also be used to support postdoctoral researchers.

Due to the COVID-19 pandemic, the 2021 SHIP and SURF programs were virtual affairs. Nevertheless, Division staff provided research opportunities for eight SURF students and three SHIP students. Nevertheless, we still managed to stay engaged with a large number of students. In total, during the last 15 months, ACMD supported the work of 27 student interns, including 14 graduate students, nine undergraduates, and three high school students. See Table 7 for a complete listing.

ACMD staff members are also active in the education of graduate students, serving both as Ph.D. advisers and as members of thesis committees. See page 160.

Gender, Equity, and Inclusion Survey Study at NIST

Justyna P. Zwolak
Mary F. Theofanos (NIST ADLP)
Jasmine Evans (Morgan State University)
Sandra Spickard Prettyman (University of Akron)

In the fall of 2019, NIST funded three studies to better understand equity and inclusivity at NIST. The purpose of this research effort is to explore the experiences of NIST federal employees to identify the ways in which differential opportunities and outcomes related to gender might exist.

This project focuses specifically on gender and inclusivity and was carried out in three consecutive phases. In the first phase, about 10 years of human resource data was analyzed to compare demographics such as positions, promotions, and awards between men and women at NIST. The results of this analysis informed the development of an interview protocol for the second phase of the project. The in-depth interviews with 40 employees from a range of positions and career paths, in turn, informed the development of an organization-wide quantitative on-line survey of federal employees at NIST. The survey included seven sections

---

[26] https://www.nist.gov/iaao/academic-affairs-office/nist-professional-research-experience-program-prep

[27] Brown University, Georgetown University, Montgomery College, Towson University, the University of the District of Columbia, the University of Maryland College Park, and a consortium of Johns Hopkins University, Morgan State University, and the State University of New York at Binghamton.

[28] Brown University, the Colorado School of Mines, the University of Colorado Boulder, and the University of Colorado Denver

[29] https://www.nist.gov/careers/student-opportunities/summer-high-school-intern-program

[30] https://www.nist.gov/surf

with questions pertaining to the employees' perception of NIST leadership commitment to diversity and inclusivity (4 questions) and NIST culture (8 questions), their experiences at NIST related to gender (11 questions), their perceptions of their work-related interactions at NIST (16 questions), their beliefs about career advancement, work/life balance, and roles at NIST (13 questions), and about their overall sense of satisfactions from working at NIST and perceived respect, and recognition at NIST (2 questions). The last section included a series of demographic questions. The main research questions addressed in the survey are as follows:

1. What are the differences, if any, in the ways in which men and women experience work at NIST?

2. What, if any, gender inequities exist at NIST? To what extent, if at all, do employees at NIST perceive that gender inequities exist?

3. In what ways, if at all, does NIST culture contribute to a lack of inclusivity?

The survey was disseminated via email in mid-August 2020 and remained opened for a period of three weeks.

More than 1500 employees responded to the survey for an overall response rate of 33.2 %, with 1108 completing it. Survey participants demographics were representative of the overall distribution of staff across directorates, operating units, age, years of service, career paths, pay bands, and educational levels.

The survey confirms, clarifies and/or expands on the gender related issues identified in Phase 2 of the study, the interviews. The comparison of responses between genders revealed many responses where women and men experience NIST in the same way. For example, women and men agree that teamwork is valued and rewarded at NIST, they believe they "fit in" with colleagues they work with, that opportunities to be hired do not depend on gender, and that their need to balance work and other life obligations is supported. Both genders also agree that doing interesting, challenging work gives them a sense of accomplishment that provides the most satisfaction. However, for several questions significant differences were found between genders, providing evidence that men and women have different views and understanding of how gender affects certain aspects of the culture, beliefs, and even interactions at NIST. In particular, statistically significant differences ($p < 0.001$) were found for men and women with respect to the diversity and inclusivity, meritocracy, gendered experiences, promotions, and opportunities.

Overall, the survey results align with the qualitative results and provide quantitative data on the differences in which men and women experience the culture, diversity, and inclusivity of NIST. A report summarizing the analysis of the survey data concludes the third and final phase of the project [1].

[1] M. F. Theofanos, J. Evans, J. P. Zwolak, and S. Spickard Prettyman. Survey on Gender, Equity and Inclusion. NISTIR 8362, March 2021. DOI: 10.6028/NIST.IR.8362

Mapping and Analyzing Employee Networks through the NIST Interactions Survey

Justyna P. Zwolak
Laura Espinal (NIST ADLP)
Camila Young (NIST EL)

In the fall of 2019, NIST funded three data-driven initiatives to better understand equity and inclusivity at NIST. The purpose of this research effort is to explore the inclusivity of the NIST workforce through the social network analysis. The goal is to provide the organization with insights about the current climate at NIST as well as suggest strategies that will help promote gender and minority equity and inclusion.

One way to assess inclusivity in a work environment is through the observance of the individuals' direct interactions. Social network analysis (SNA) provides a set of tools that can be naturally used to analyze interaction data. With the focus of this project being inclusivity at the workplace, we consider two types of relationships: (1) interactions related to achieving work-related goals at NIST, and (2) advice about career-related decisions at NIST. We use the measure of composition of employees' self-reported *ego-centric networks*, that is networks
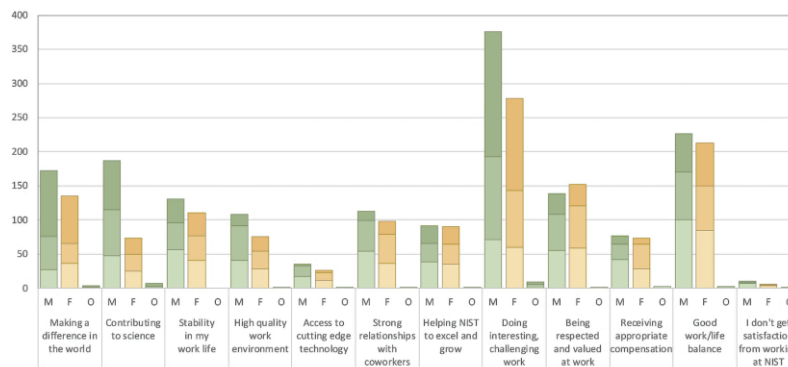

Figure 101. *Results of a sample survey question: Satisfaction from working at NIST. For each question, the responses are divided by gender (M denotes male, F-female, and O-other gender). The respondents were asked to rank their three top choices. The intensity of color represents the rank order, with the darkest color corresponding to rank 1.*
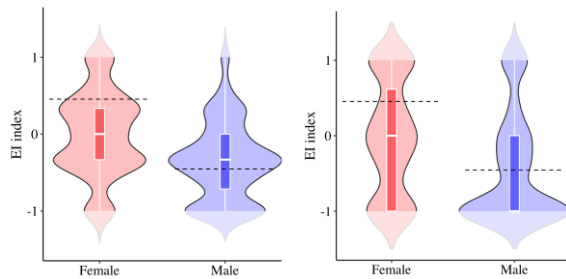
Figure 102. *The distribution of the homophily index between genders. Left: violin plots with overlaid box plots for the homophily index distribution for the advice network. Right: the advice networks. The dashed lines represent theoretical reference point based on the population.*

involving only the responded and their direct connections, as a proxy for inclusivity. Importantly, while the population of interest is NIST federal employees, to ensure that the reported networks are true representations of the NIST environment, we opted to define the network boundaries by the status of being either a NIST federal employee or an associate. In other words, while the survey was only sent out to NIST federal employees, the respondents were allowed to include NIST associates as their connections.

The NIST Interactions Survey was disseminated via email in early-November 2020 and remained opened for a period of three weeks. In addition to the two network questions, the respondents were asked seven demographic and work history questions about themselves (gender, race, age, supervisory status, tenure, location, and OU) as well as five questions about each reported connection (gender, relative age (compared to self), supervisory status, location, and OU). More than 800 employees completed the survey for an overall response rate of 24.6 %. Survey participants demographics were fairly representative of the overall distribution of NIST federal employees.

The ego network analysis was carried out in a series of sequential steps, with results from each phase informing the focus of the next phase. In the first stage, we focused on the networks overall structure and size. No statistically significant differences were found when comparing between respondents' genders and ethnic groups for either of the network type. Next, we analyzed the ego network compositions. For the work-related network most of the differences in composition was correlated with respondent's location and OU. For the advice-network, however, we found statistically significant differences in composition based on gender, age group, and tenure.

The lack of gender- and ethnicity-based differences suggest that the work-related networks can be thought of as a "network of convenience," capturing the snapshot of current research priorities at NIST. The addition of statistically significant differences in the advice-network composition based on gender and age suggest

more a preferential nature of these networks, with the connections potentially being actively sought after. This hypothesis was confirmed in the third phase of analysis focusing on network homophily, that is a measure of similarity between the respondents and their connections. Here, we found that, while men advice networks tens to include mostly men, women's network are more balanced. However, we also found that several women reported networks that include only women, which, given the population of NIST, suggest that these women must have made a conscious effort to establish such networks. As such, the advice-related network can be thought of as a "network of choice." The analysis of the NIST Interactions Survey has been published in NIST Internal Report [1].

[1] L. Espinal, C. Young, and J. P. Zwolak. Mapping Employee Networks through the NIST Interactions Survey. NISTIR 8375, June 2021. DOI: 10.6028/NIST.IR.8375

Restoring Organizational Structure Using Projected Ego-Centric Networks

Robert P. Dalka (University of Maryland)
Justyna P. Zwolak

Organizational network analysis (ONA) is a class of methodologies that aim to explore the social ties within and between formal organizations to better understand phenomena such as knowledge and resource transfer, social influences within teams, or the effect of team building on the dynamics of an organization's social network. Results from ONA can be used to identify how the network structures can be leveraged to better fulfill an organization's mission and, in turn, inform the organization's decisions concerning its own restructuring. However, participants of ONA research are often at a high risk of identification due to managerial dynamics and the type of information that is used to build the network. Additionally, the personally identifiable information (PII) that is typically required to conduct full ONA prohibits this research from being conducted in many organizational contexts (e.g., government institutions).

To overcome these limitations, we propose an alternative method for evaluating interconnectedness within or between organizations. Rather than relying on employees' names to build a network, we use a projection of anonymous ego-centric data onto organizational units as a proxy to capture the organizational network structure. The networks used in this study are built based on the NIST Interactions Survey data [1]. This dataset captures two distinct types of social interactions among NIST employees: their work-related collaborations and
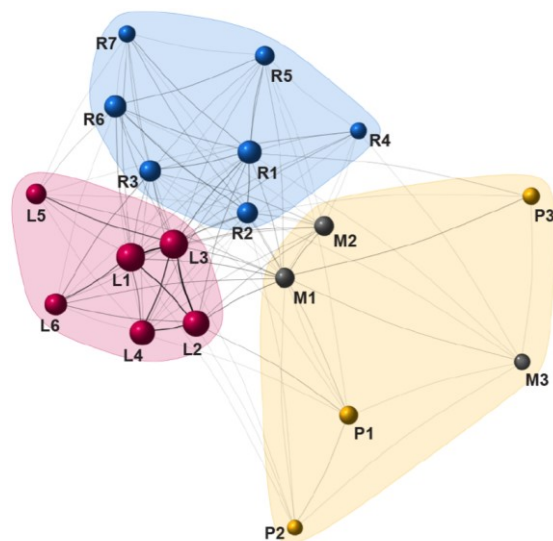
Figure 103. *The second-level communities identified within the collaboration network. The nodes are sized by the number of responses from each unit and the line width represents the number of interactions between two units. Each unit (node) is labeled and colored by the organizational division to which it belongs. The shaded regions indicate the identified communities.*

advice seeking. The two resulting networks are expected to exhibit somewhat different structures: whereas collaborating on a project necessitates interaction with other employees to fulfill the organization's mission, not everyone seeks advice internally or in a systematic way. Thus, we build and compare both networks to demonstrate and validate our approach.

We employ a modular analysis to determine whether the projected unit-based networks accurately reflect the network structure of the organizational units that comprise NIST. For the first type of social interaction investigated, the collaboration networks of NIST employees, we find that the structure restored in the projected network closely resembles the NIST organizational structure. The second level of the hierarchical community structure, depicted in Figure 103, confirms that the projected network exhibits a community assignment that aligns with the organizational divisions of NIST [2]. For the second type of social interaction, the advice network, we find slightly more variability in the communities identified as well as more intermixing between units. We also find that the community assignment for the advice network is less stable and more susceptible to sampling bias, as opposed to the collaboration network.

Our analysis confirms that data capturing truly anonymous (as opposed to anonymized) individuals' ego-centric data suffice to reliably approximate the full organizational network. We also show that, depending on the type of examined interaction, different aspects of the organizational structure can be investigated.

Anonymity is necessary to perform ONA within government agencies, and other institutions in which PII cannot be collected. Our approach is also useful for other spaces, such as national labs and universities which may be most interested in the interconnectedness of research groups or departments rather than individuals. The methodology we propose sets a path forward for performing ONA based on accessible ego-centric data, with minimal identifiable information, to investigate large-scale aspects of organizational structures. An article reporting this work, including an investigation into the stability of the proposed approach, has recently been published online [3].

[1] L. Espinal, C. Young, and J. P. Zwolak. Mapping Employee Networks through the NIST Interactions Survey. NISTIR 8375, June 2021. DOI: 10.6028/NIST.IR.8375

[2] NIST General Information. Online, accessed 17 November 2021. URL: https://www.nist.gov/director/pao/nist-general-information

[3] R. P. Dalka and J. P. Zwolak. Restoring the Structure: A Modular Analysis of Ego-Driven Organizational Networks. Preprint arXiv: 2201.01290, 2022.

Physics Education Survey Validation Through a Network Analytic Approach

Robert P. Dalka (University of Maryland)
Justyna P. Zwolak
Diana Sachmpazidi (University of Maryland)
Charles Henderson (Western Michigan University)

This ongoing project is focused on developing a modular-analysis-based network analytic approach for better understanding of Likert scale survey structure. Multiple surveys aimed at capturing various aspects of student experiences have been developed in physics education research (PER). The validity of such surveys is typically assessed through confirmatory or exploratory analysis. An exploratory approach, such as principal component analysis (PCA) and factor analysis (FA), aims to optimize the grouping of individual variables (here, survey questions) into a set of higher order components. Our ongoing work seeks to provide an alternative exploratory approach through the modular analysis of network representations of the survey responses.

The Aspects of Student Experience Survey (ASES) provides an example of the types of questions and findings that can be asked and uncovered through survey methods [1]. ASES is an instrument designed to assess physics graduate student experiences of departmental support structures. The utility of the ASES was confirmed through analysis of responses of 397 students from 19 physics graduate programs across the USA. PCA revealed four components: mentoring and research
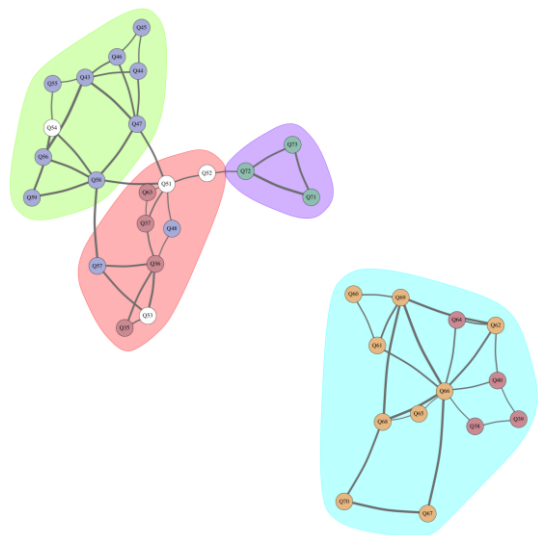
Figure 104. *The modular communities identified in the network of ASES items. The nodes are labeled by the unique question ID and are colored by the earlier identified principal component. Each shaded region corresponds to a community identified in the network.*

experience; professional development; social and academic integration; and financial support.

Our work focuses on developing and evaluating an approach that employs a network analysis of survey data, such as ASES. In this analysis, the items of the survey are treated as nodes in the network and the similarity (or differences) in how respondents answered those items is used to represent weighted edges between the nodes. Survey items that are answered in similar ways (e.g., Q1 receives many of the same type of responses as Q2) are connected through positively weighted edges. Survey items that are answered in dissimilar ways are connected through negatively weighted edges. The module structure of the survey can be found from local connections between survey items through identifying communities of nodes within the network, analogous to the principal components or factors identified in PCA and FA.

The resulting ASES item network is extremely dense—every node is connected to every other node. To preform analysis of the network, we must identify the most important connections between each node, that is the network *backbone*. The backbone is the residual network after eliminating non-significant edges. Several algorithms have been developed to extract the backbone of dense networks. We employ the locally adaptive network sparsification (LANS) algorithm to find the network backbone [2]. The resulting network includes 35 nodes (one for each survey item in ASES) and 55 edges, with 34 (62 %) of the edges created through mutual "agree" selections and 21 (38 %) of the edges created through mutual "disagree" selections. The network is split into two components, with the upper left component, as seen in Figure 104, consisting of the mutual "agree" selections and the other component consisting of the mutual "disagree" selections.

Through modular analysis, we have identified four communities, shown in Figure 104. While these somewhat resemble the components identified through the PCA of the same data set, we find some important differences in the composition of components. Our continuing work focuses on comparing these two different partitions of the ASES items, both qualitatively and quantitatively through their effect size on self-reported student efficacy [3].

[1] D. Sachmpazidi and C. Henderson. Departmental Support Structures for Physics Graduate Students: Development and Psychometric Evaluation of a Self-Report Instrument. *Physical Review Physics Education Research* **17** (2021), 010123.

[2] N. J. Foti, J. M. Hughes, and D. N. Rockmore. Nonparametric Sparsification of Complex Multiscale Networks. *PLoS One* **6** (2011), e16431.

[3] R. P. Dalka, D. Sachmpazidi, C. Henderson, and J. P. Zwolak. Social Networks Perspective on the ASES data. In preparation.

Table 7. *Student interns in ACMD.*

Name	From	Program		Mentor	Topic
Agostino, Isabel	The College of William & Mary	U	SURF	M. Mascagni	Developing a standard family of point sets with known discrepancy
Agrawal, Sristy	U. Colorado Boulder	G	FGR	S. Glancy	Resource based quantum information theory
Alhejji, Mohammad	U. Colorado Boulder	G	PREP	E. Knill	Quantum randomness protocols
Avagyan, Arik	U. Colorado Boulder	G	PREP	E. Knill	Quantum information processing
Bodenschatz, John	U. of Cincinnati	U	SURF	H. Cohl	Legendre elliptic integral representations of symmetric elliptic integrals
Chen, Jeffrey	Poolesville HS	H	SHIP	O. Slattery	Interactive web-nimation for quantum science phenomena
Cong, Alyssa	Brown University	U	DGR	B. Cloteaux	Rewiring systems under infection avoidance
Dalka, Robert	U. Of Maryland	G	PREP	J. Zwolak	Developing a social network analysis toolbox for anonymous ego-networks data
Emiyah, Christian	Morgan State U.	G	PREP	B. Cloteaux	Modeling of contact tracing
Euwart, Elizabeth	Worcester Polytechnic Institute	U	SURF	M. Mascagni	Investigating properties of the walk on spheres algorithm
Eveleth, Jason	Brown U.	U	SURF	A. Kearsley	Optimal Voronoi tessellation
Fan, Julie	Winston Churchill HS	H	SHIP	H. Cohl	Computationally generated mathematical n-ary tree art on smooth 2D manifolds
Fathima, Hashmath	Morgan State U.	G	PREP	A. Kearsley	Artificial intelligence for mass spectrometry
Geller, Shawn	U. of Colorado	G	PREP	E. Knill	Characterization of quantum state preparation and measurement errors
Hu, Mingyu	U. Colorado Boulder	G	PREP	M. Donahue	Modeling of antiferromagnets
Koh, Sophia	Amherst College	U	SURF	J. Zwolak	Machine learning feature detection for ultra-cold atoms
Khrac, Katjana	U. Zagreb, Croatia	G	FGR	K. Sayrafian	Positioning metrology in the human body
Kwiatkowski, Alexander	U. of Colorado	G	PREP	E. Knill	Quantum networking
Lauria, Michael	U. of Colorado	G	PREP	E. Knill	Characterization of quantum computers
Lef, Peter	Clarkson U.	U	SURF	G. Dogan	Computational tools for image analysis
Mooney, Timothy	George Mason U.	U	SURF	L. Brady	Complex deformation of path-integral quantum Monte Carlo
Ornstein, Joel	U. of Colorado	G	PREP	S. Glancy	Quantum information journal club
Schneider, Ryan	U. C. San Diego	G	DGR	B. Schneider	Collocation methods to solve the electronic Schrödinger equation
Seshadri, Akshay	U. Colorado	G	PREP	E. Knill	
Smith, Stephen	U. of South Carolina	G	NSF	M. Coudron	Computational power of low-depth quantum circuits
Su, Ruisi	Carnegie Mellon U.	U	FGR	A. Kearsley	Detecting network anomalies
Suris Rodriguez, Ramon	Towson U.	U	SURF	H. Cohl	Integral and q-integral representations of the Askey-Wilson polynomials and their symmetric subfamilies
Wiesing, Tom	Jacobs U. Bremen	G	FGR	B. Miller	Digital Library of Math Functions
Xiao, Vivian	Poolesville HS	H	SHIP	A. Kearsley	Optimization in chemometrics

Legend				
	G	Graduate Student	PREP	Professional Research Experience Program
	U	Undergraduate	FGR	Foreign Guest Researcher
	H	High School	DGR	Domestic Guest Researcher
			NSF	NSF Mathematical Sciences Graduate Internship Program
			SHIP	Summer High School Internship Program
			SURF	Summer Undergraduate Research Fellowship

Part IV

Activity Data
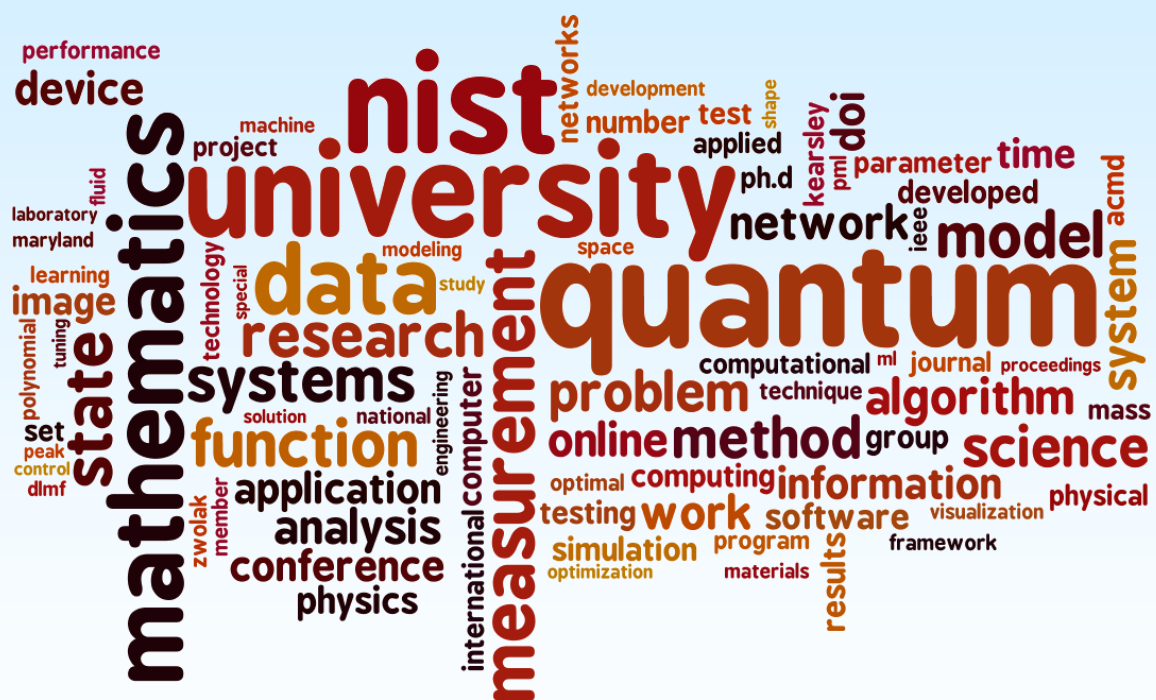
Publications

Note: Names of (co-)authors with a Division affiliation during this reporting period are underlined.

Appeared

Refereed Journals

1. J. Arrazola, V. Bergholm, K. Bradler, T. Bromley, M. Collins, I. Dhand, A. Fumagalli, <u>T. Gerrits</u>, A. Goussev, L. Helt, J. Hundal, T. Isacsson, R. Israel, J. Izaac, S. Jahangiri, R. Janik, N. Killoran, S. Kumar, J. Lavoie, A. Lita, D. Mahler, M. Menotti, B. Morrison, S. Nam, L. Neuhaus, H. Qi, N. Quesada, A. Repingon, K. Sabapathy, M. Schuld, D. Su, J. Swinarton, A. Szava, K. Tan, P. Tan, V. Vaidya, Z. Vernon, Z. Zabaneh, and Y. Zhang. Quantum Circuits with Many Photons on a Programmable Nanophotonic Chip. *Nature* **591**, 54-60 (2021). DOI: [10.1038/s41586-021-03202-1](10.1038/s41586-021-03202-1)

2. I. Bell and <u>B. Alpert</u>. Efficient and Precise Representation of Pure Fluid Phase Equilibria with Chebyshev Expansions. *International Journal of Thermophysics* **42,** 75 (2021). DOI: [10.1007/s10765-021-02824-x](10.1007/s10765-021-02824-x)

3. A. Carasso. Stabilized Leapfrog Scheme Run Backward in Time, and the Explicit $O(\Delta t)^2$ Stepwise Computation of Ill-Posed Time-Reversed 2D Navier-Stokes Equations. *Inverse Problems in Science and Engineering* **29**:13 (2021), 3062-3085. DOI: [10.1080/17415977.2021.1972997](10.1080/17415977.2021.1972997)

4. A. Carasso. Data Assimilation in 2D Viscous Burgers Equation using a Stabilized Explicit Finite Difference Scheme Run Backward in Time. *Inverse Problems in Science and Engineering* **29**:13 (2021), 3475-3489. DOI: [10.1080/17415977.2021.2009476](10.1080/17415977.2021.2009476)

5. <u>H. Cohl</u>, J. Park, and H. Volkmer. Gauss Hypergeometric Representations of the Ferrers Function of the Second Kind. *Symmetry, Integrability and Geometry: Methods and Applications* **17** (2021), 053. DOI: 10.3842/SIGMA.2021.053

6. <u>H. Cohl</u>, M. Schubotz, and O. Teschke. Connecting Islands: Bridging zbMATH and DLMF with Scholix, a Blueprint for Expert Knowledge Systems. *European Mathematical Society Magazine* **120** (2021), 66-67. DOI: [10.4171/MAG-35](10.4171/MAG-35)

7. <u>M. Coudron</u>, J. Stark, and T. Vidick. Trading Locality for Time: Certifiable Randomness from Low-Depth Circuits. *Communications in Mathematical Physics* **382** (2021), 49–86. DOI: [10.1007/s00220-021-03963-w](10.1007/s00220-021-03963-w)

8. <u>R. Evans</u>, A. Balijepalli, and <u>A. Kearsley</u>. Transport Phenomena in Biological Field Effect Transistors. *SIAM Journal on Applied Mathematics* **80**:6 (2020), 2586–2607. DOI: [10.1137/19M1255495](10.1137/19M1255495)

9. <u>J. Fong</u>, N. Heckert, J. Filliben, P. Marcal, and S. Freiman. Estimation of a Minimum Allowable Structural Strength Based on Uncertainty in Material Test Data. *Journal of Research National Institute of Standards and Technology* **126** (2021), 126036. DOI: 10.6028/jres.126.036

10. J. Fowler, G. O'Neil, <u>B. Alpert</u>, D. Bennett, E. Denison, W. Doriese, G. Hilton, L. Hudson, Y. Joe, K. Morgan, D. Schmidt, D. Swetz, C. Szabo, and J. Ullom. Absolute Energies and Emission Line Shapes of the L X-Ray Transitions of Lanthanide Metals. *Metrologia* **58** (2021), 015016. DOI: [10.1088/1681-7575/abd28a](10.1088/1681-7575/abd28a)

11. <u>H. Gharibnejad</u>, N. Douguet, <u>B. Schneider</u>, J. Olsen, and L. Argenti. Multi-Center Quadrature Scheme for the Molecular Continuum. *Computer Physics Communications* **263** (2021), 107889. DOI: [10.1016/j.cpc.2021.107889](10.1016/j.cpc.2021.107889)

12. S. Guo, A. Fritsch, C. Greenberg, I. Spielman, and <u>J. Zwolak</u>. Machine-Learning Enhanced Dark Soliton Detection in Bose-Einstein Condensates. *Machine Learning: Science and Technology* **2**:3 (2021), 035020. DOI: 10.1088/2632-2153/abed1e

13. <u>R. Kacker</u>, D. Kuhn, <u>Y. Lei</u>, and <u>D. Simos</u>. Factorials Experiments, Covering Arrays, and Combinatorial Testing. *Mathematics in Computer Science* **15** (2021), 715-739. DOI: [10.1007/s11786-021-00502-7](10.1007/s11786-021-00502-7)

14. R. Kacker. Towards Correction of the JCGM International Vocabulary of Metrology. *Measurement: Sensors* **18** (2021)**,** 100063. DOI: 10.1016/j.measen.2021.100063

15. K. Keenan, <u>Z. Gimbutas</u>, <u>A. Dienstfrey</u>, K. Stupic, M. Boss, S. Russek, T. Chenevert, P. Prasad, J. Guo, W. Reddick, K. Cecil, A. Shukla-Dave, D. Nunez, A. Konar, M. Liu, S. Jambawalikar, L. Schwartz, J. Zheng, P. Hu, and E. Jackson. Multi-site, Multi-Platform Comparison of MRI T1 Measurement Using a System Phantom. *PLOS One* **16** (6) e0252966. DOI:10.1371/journal.pone.0252966

16. S. Kotler, G. Peterson, <u>E. Shojaee</u>, F. Lecocq, K. Cicak, <u>A. Kwiatkowski</u>, <u>S. Geller</u>, <u>S. Glancy</u>, <u>E. Knill</u>, R. Simmonds, J. Aumentado, and J. Teufel. Direct Observation of Deterministic Macroscopic Entanglement. *Science* **372** (2021)**,** 622-625. DOI:10.1126/science.abf2998

17. A. Li and <u>S. Su</u>. Accelerating Binarized Neural Networks via Bit-Tensor-Cores in Turing GPUs. *IEEE Transactions on Parallel and Distributed Systems* **32**:7 (2021), 1878-1891. DOI: 10.1109/TPDS.2020.3045828

18. L. Li, D. Young, <u>V. Albert</u>, K. Noh, C. Zou, and L. Jiang. Phase-Engineered Bosonic Quantum Codes. *Physical Review A* **103** (2021), 062427. DOI: <u>10.1103/PhysRevA.103.062427</u>

19. D. Lum, M. Mazurek, A. Mikhaylov, K. Parzuchowski, R. Wilson, R. Jimenez, <u>T. Gerrits</u>, M. Stevens, M. Cicerone, and C. Camp. Witnessing the Survival of Time-Energy Entanglement through Biological Tissue and Scattering Media. *Biomedical Optics Express* **12** (2021), 3658-3670. DOI: <u>10.1364/BOE.423743</u>

20. V. Mai, <u>R. La</u>, and A. Battou. Optimal Cybersecurity Investments in Large Networks using SIS Model: Algorithm Design. *IEEE/ACM Transactions on Networking* **29**:6 (2021), 2453-2466. DOI: <u>10.1109/TNET.2021.3091856</u>

21. J. Majikes, <u>P. Patrone</u>, <u>A. Kearsley</u>, M. Zwolak, and J. Liddle. Failure Mechanisms in DNA Self-Assembly: Barriers to Single-Fold Yield. *ACS Nano* **15**:2 (2021), 2384-3294. DOI: <u>10.1021/acsnano.0c10114</u>

22. N. Martys, <u>W. George</u>, <u>S. Satterfield</u>, B. Toman, M. Peltz, S. Jones, and C. Ferraris. Standard Reference Materials for Rheological Measurements of Cement-Based Materials. *ACI Materials Journal* **118**:6 (2021), 325-330. DOI: <u>10.14359/51733132</u>

23. R. Nehra, M. Eaton, C. González-Arciniegas, M. Kim, <u>T. Gerrits</u>, A. Lita, S. Nam, and O. Pfister. Generalized Overlap Quantum State Tomography. *Physical Review Research* **2** (2020), 042002. DOI: <u>10.1103/PhysRevResearch.2.042002</u>

24. K. Parzuchowski, A. Mikhaylov, M. Mazurek, R. Wilson, D. Lum, <u>T. Gerrits</u>, C. Camp Jr., M. Stevens, and R. Jimenez. Setting Bounds on Two-Photon Absorption Cross-Sections in Common Fluorophores with Entangled Photon Pair Excitation. *Physical Review Applied* **15** (2021), 044012. DOI: <u>10.1103/PhysRevApplied.15.044012</u>

25. <u>P. Patrone</u> and <u>A. Kearsley</u>. Classification Under Uncertainty: Data Analysis for Diagnostic Antibody Testing. *Mathematical Medicine and Biology: A Journal of the IMA* **38**:3 (2021), 396–416. DOI: <u>10.1093/imammb/dqab007</u>

26. <u>P. Patrone</u>, A. Li, G. Cooksey, and <u>A. Kearsley</u>. Measuring Microfluidic Flow Rates: Monotonicity, Convexity, and Uncertainty. *Applied Mathematics Letters* **112** (2021), 106694. DOI: <u>10.1016/j.aml.2020.106694</u>

27. C. Qu, <u>B. Schneider</u>, <u>A. Kearsley</u>, W. Keyrouz, and T. Allison, Predicting Kováts Retention Indices Using Graph Neural Networks. *Journal of Chromatography A* **1646** (2021), 462100. DOI: <u>10.1016/j.chroma.2021.462100</u>

28. G. Sahonero-Alvarez, A. Singh, <u>K. Sayrafian</u>, L. Bianchi, and A. Roman-Gonzalez. A Functional BCI Model by the P2731 Working Group: Transducer. *Brain-Computer Interfaces Journal* **8**:3 (2021), 92-107. DOI: <u>10.1080/2326263X.2021.1968633</u>

29. L. Scarlett, J. Savage, D. Fursa, I. Bray, M. Zammit, and <u>B. Schneider</u>. Convergent Close-Coupling Calculations of Electrons Scattering on Electronically Excited Molecular Hydrogen. *Phys. Rev. A* **103** (2021), 032802. DOI: <u>10.1103/PhysRevA.103.032802</u>

30. J. Seilert, A. Moorthy, <u>A. Kearsley,</u> and E. Flöter. Revisiting a Model to Predict Pure Triglyceride Thermodynamic Properties: Parameter Optimization and Performance. *Journal of the American Oil Chemists Society* **98**:8 (2021), 837-850. DOI: <u>10.1002/aocs.12515</u>

31. L. Shalm, Y. Zhang, J. Bienfang, C. Schlager, M. Stevens, M. Mazurek, C. Abellán, W. Amaya, <u>M. Mitchell</u>, <u>M. Alhejji</u>, H. Fu, J. Ornstein, R. Mirin, S. Nam, and <u>E. Knill</u>. Device-Independent Randomness Expansion with Entangled Photons. *Nature Physics* **17** (2021), 452-456. DOI: <u>10.1038/s41567-020-01153-4</u>

32. <u>J. Sims</u>, B. Padhy, and M. Ruiz. Exponentially Correlated Hylleraas-Configuration Interaction Nonrelativistic Energy of the ^{1}S Ground State of the Helium Atom. *International Journal of Quantum Chemistry* **121**:4 (2020), e26470. DOI: <u>10.1002/qua.26470</u>

33. J. Sims. Hylleraas-Configuration Interaction (Hy-CI) Non-Relativistic Energies for the 3 ^{1}S, 4 ^{1}S, 5 ^{1}S, 6 ^{1}S, and 7 ^{1}S Excited States of the Beryllium Atom. *Journal of Research of the NIST* **125** (2020), 125006. DOI: <u>10.6028/jres.125.006</u>

34. <u>J. Sims</u> and M. Ruiz. Parallel Generalized Real Symmetric-Definite Eigenvalue Problem. *Journal of Research of NIST* **125** (2020), 125032. DOI: <u>10.6028/jres.125.032</u>

35. J. S. Sims, B. Padhy, and M. B. Ruiz. Exponentially Correlated Hylleraas-Configuration Interaction (E-Hy-CI) Studies of Atomic Systems. II. Non-relativistic Energies of the 1 ^{1}S through 6 ^{1}S States of the Li$^+$ Ion. *International Journal of Quantum Chemistry* **122** (2021), e26823. DOI: <u>10.1002/qua.26823</u>

36. R. Srinivas, S. Burd, H. Knaack, R. Sutherland, A. Kwiatkowski, <u>S. Glancy</u>, <u>E. Knill</u>, D. Wineland, D. Leibfried, A. Wilson, D. Allcock, and D. Slichter. High-fidelity Laser-free Universal Control of Trapped Ion Qubits. *Nature* **597** (2021), 209-213. DOI: <u>10.1038/s41586-021-03809-4</u>

37. K. Stupic, M. Ainslie, M. Boss, C. Charles, <u>A. Dienstfrey</u>, J. Evelhoch, P. Finn, <u>Z. Gimbutas</u>, J. Gunter, D. Hill, C. Jack, E. Jackson, T. Karaulanov,

K. Keenan, G. Liu, M. Martin, P. Prasad, N. Rentz, C. Yuan, and S. Russek. Standard System Phantom for Magnetic Resonance Imaging. *Magnetic Resonance in Medicine* **86** (2021), 1194-1211. DOI: 10.1002/mrm.28779

38. T. Sturges, T. McDermott, A. Buraczewski, W. Clements, J. Renema, S. Nam, T. Gerrits, A. Lita, W. Kolthammer, A. Eckstein, I. Walmsley, and M. Stobińska. Quantum Simulations with Multiphoton Fock States. *Quantum Information* **7** (2021), 91. DOI: 10.1038/s41534-021-00427-w

39. G. Thekkadath, M. Mycroft, B. Bell, C. Wade, A. Eckstein, D. Phillips, R. Patel, A. Buraczewski, A. Lita, T. Gerrits, S. Nam, M. Stobińska, A. Lvovsky, and I. Walmsley. Quantum-enhanced Interferometry with Large Heralded Photon-Number States. *New Journal of Physics Quantum Information* **6** (2020), 89. DOI:10.1038/s41534-020-00320-y

40. L. Tian, E. Elsheikh, P. Patrone, A. Kearsley, A. Gaigalas, S. Inwood, S. Lin-Gibson, D. Esposito, and L. Wang. Towards Quantitative and Standardized Serological and Neutralization Assays for COVID-19. *International Journal of Molecular Sciences* **22**:5 (2021) , 2723. DOI: 10.3390/ijms22052723

41. L. Wang, R. Bhardwaj, H. Mostowski, P. Patrone, A. Kearsley, J. Watson, L. Lim, J. Pichaandi, O. Ornatsky, D. Majonis, S. Bauer, and H. Dagheidy. Establishing CD19 B-cell Reference Control Materials for Comparable and Quantitative Cytometric Expression Analysis. *PLOS One* **16**:3 (2021), e0248118. DOI: 10.1371/journal.pone.0248118

42. B. Weber, S. Kalantre, T. McJunkin, J. Taylor, and J. Zwolak. Theoretical Bounds on Data Requirements for the Ray-based Classification. *SN Computer Science* **3**:1 (2022), 57. DOI: 10.1007/s42979-021-00921-0

43. A. Weiss, S. Gautham, A. Jayakumar, C. Elks, D. Kuhn, R. Kacker, and T. Preusser. Understanding and Fixing Complex Faults in Embedded Systems. *IEEE Computer* **54**:1 (2021), 49-60. DOI: 10.1109/MC.2020.3029975

44. D. Wu and V. Albert. Approximating the Two-mode Two-photon Rabi Model. *Physics Letters A* **422** (2022)**,** 127779. DOI: 10.1016/j.physleta.2021.127779

45. C. You, M. Hong, P. Bierhorst, A. Lita, S. Glancy, S. Kolthammer, E. Knill, S. Nam, R. Mirin, O. Magaña-Loaiza, and T. Gerrits. Multiphoton Quantum Metrology with Neither Pre-nor Post-Selected Measurements. *Applied Physics Reviews* **8** (2021), 041406. DOI: 10.1063/5.0063294

46. Y. Zhai, N. Martys, W. George, J. Curtis, J. Nayem, Y. Zhang, and Y. Liu. Intermediate Scattering Functions of a Rigid Body Monoclonal Antibody Protein in Solution Studied by Dissipative Particle Dynamic Simulation. *Structural Dynamics* **8** (2021), 024102. DOI: 10.1063/4.0000086

47. Y. Zhang, H. Lo, A. Mink, T. Ikuta, T. Honjo, H. Takesue, and W. Munro. A Simple Low-Latency Real-Time Certifiable Quantum Random Number Generator. *Nature Communications* **12** (2021), 1056 . DOI: 10.1038/s41467-021-21069-8

48. J. Zwolak, T. McJunkin, S. Kalantre, S. Neyens, E. MacQuarrie, M. Eriksson, and J. Taylor. Ray-Based Framework for State Identification in Quantum Dot Devices. *PRX Quantum* **2**:2 (2021), 020335. DOI: 10.1103/PRXQuantum.2.020335

Books

1. H. Cohl and M. Ismail, eds. Lectures on Orthogonal Polynomials and Special Functions: Sixth Summer School, Maryland, 2016. *London Mathematical Society Lecture Note Series* **464**, Cambridge University Press, Cambridge, 2021.

Book Chapters

1. A. Moorthy and A. Kearsley. Pattern Similarity Measures Applied to Mass Spectra. In *Progress in Industrial Mathematics: Success Stories* (M. Cruz, C. Parés, and P. Quintela, eds.), SEMA SIMAI Springer Series **5**. Springer, Cham. 2021. DOI: 10.1007/978-3-030-61844-5_4

2. K. Sayrafian and S. Ambroziak. IoT for Healthcare Applications. In *Inclusive Radio Communication Networks for 5G and Beyond* (C. Oestges and F. Quitin, eds.), Elsevier, 2021, 221-252.

In Conference Proceedings

1. S. Bhushan, O. Slattery, X. Tang, and L. Ma. Terahertz Electromagnetically Induced Transparency in Cesium Atoms. In *Proceedings of the 2020 Conference on Frontiers in Optics and Laser Science* (FIO+LS), Washington, DC, September 14-17, 2020, JTu1B.40. DOI: 10.1364/FIO.2020.JTu1B.40

2. J. Chandrasekaran, Y. Lei, R. Kacker, and D. Kuhn. A Combinatorial Approach to Explaining Image Classifiers. In *Proceedings of the 14-th IEEE International Conference on Software Testing, Verification, and Validation Workshop (ICSTW-2021)*, Online, April 12-16, 2021, 35-43. DOI: 10.1109/ICSTW52544.2021.00019

3. J. Chandrasekaran, A. Patel, Y. Lei, R. Kacker, and D. Kuhn. Evaluation of T-Way Testing of DNNs in Autonomous Driving Systems. In *Proceeding of 2021 IEEE International Conference on Artificial*

Intelligence Testing (AITest), Online, August 23-26, 2021, 17-18. DOI: 10.1109/AITEST52744.2021.00013

4. J. Chandrasekaran, Y. Lei, R. Kacker, and D. Kuhn. A Combinatorial Approach to Testing Deep Neural Network-based Autonomous Driving Systems. In *Proceedings of the 14-th IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, Online, April 12-16, 2021, 57-66. DOI: 10.1109/ICSTW52544.2021.00022

5. Z. Chen, J. Wen, A. Kearsley, and A. Pertzborn Smoothing Techniques in Dynamic Building System Simulation. In *Proceedings of the 2021 International Conference on Instrumentation, Control, and Automation*, Online, August 25-7, 2021. DOI: 10.1109/ICA52848.2021.9625664

6. N. J. Coble and M. Coudron. Quasi-polynomial Time Approximation of Output Probabilities of Constant-depth, Geometrically-local Quantum Circuits. In *Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2021)*, Online, February 7-10, 2021. URL: arXiv:2012.05460

7. J. Fong, P. Marcal, R. Rainsberger, N. Heckert, J. Filliben, S. Doctor, and N. Finney Jr. A Multi-Scale Failure-Probability-and-NDE-Based Fatigue Life Model for Estimating Component Co-Reliability of Uncracked and Cracked Pipes. In *Proceedings of the 2021 ASME Pressure Vessels & Piping Division Conference*, Online, July 12-16, 2021, PVP2021-62169, V001T01A029. DOI: 10.1115/PVP2021-62169

8. B. Garn, D. Lang, M. Leithner, D. Kuhn, R. Kacker, and D. Simos. Combinatorially XSSing Web Application Firewalls. In *Proceedings of the 14th IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, Online, April 12-16, 2021, 85-94. DOI: 10.1109/ICSTW52544.2021.00026

9. Z. Grey, S. Mosleh, J. Rezac, Y. Ma, J. Coder, and A. Dienstfrey. Optimizing Unlicensed Band Spectrum Sharing with Subspace-Based Pareto Tracing. In *Proceedings of the IEEE International Conference on Communications*, Montreal, Canada, June 14-23, 2021, 1-7. DOI: 10.1109/ICC42927.2021.9500533

10. A. Greiner-Petter, M. Schubotz, F. Muller, C. Breitinger, H. Cohl, A. Aizawa, and B. Gipp. Discovering Mathematical Objects of Interest-A Study of Mathematical Notations. In *Proceedings of the WWW'20 Web Conference*, Taipei, Taiwan, April 2020, 1445-1446. DOI: 10.1145/3366423.3380218

11. A. Jayakumar, S. Gautham, D. Kuhn, B. Simons, A. Collins, T. Dirsch, R. Kacker, and C. Elks. Systematic Software Testing of Critical Embedded Digital Devices in Nuclear Power Applications. In *Proceedings of the 31st International Symposium on Software Reliability Engineering Workshops (ISSREW 2020)*, Online, October 12-15, 2020, 85-90. DOI: 10.1109/ISSREW51248.2020.00042

12. A. Jayakumar, D. Kuhn, B. Simons, A. Collins, S. Gautham, R. Hite, R. Kacker, A. Rajagopala, and C. Elks. Pseudo Exhaustive Software Testing Framework for Embedded Digital Devices in Nuclear Power. In *Proceedings of the 12th Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies*, June 14-17, 2021, 812-823. URL: https://www.ans.org/store/item-700437/

13. K. Krhac, K. Sayrafian, U. Bengi, and S. Dumanli. A Wearable Wireless Monitoring System for the Detection of Pulmonary Edema. In *Proceedings of the E-Health Symposium*, IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, December 7-11, 2021. DOI: 10.1109/GLOBECOM46510.2021.9685118

14. S. Langer and A. Reid. Calculating Voxel-Polyhedron Intersections for Meshing Images. In *Proceedings of the 29th International Meshing Roundtable*, June 21-25, 2021. DOI: 10.5281/zenodo.5559225

15. E. Lanus, L. Freeman, D. Kuhn, and R. Kacker. Combinatorial Testing Metrics for Machine Learning. In *Proceedings of the 14th IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW),* Online, April 12-16, 2021, 81-84. DOI: 10.1109/ICSTW52544.2021.00025

16. S. Lau, J. Lew, C. Ho and S. Su. Exploratory Investigation on a Naïve Pseudo-Labelling Technique for Liquid Droplet Images Detection using Semi-supervised Learning. In *Proceedings of 2021 IEEE International Conference on Computing (ICOCO 2021),* November 17-19, 2021, 353-359. DOI: 10.1109/ICOCO53166.2021.9673553

17. V. Mai, R. La, and A. Battou. Optimal Cybersecurity Investments for SIS Model. In *Proceedings of the 2020 IEEE Global Communications Conference (Globecom'20),* Taipei, Taiwan, December 8-10, 2020. DOI: 10.1109/GLOBECOM42002.2020.9348109

18. V. Mai and R. La. Story of Two Populations in Epidemics: Is Every Infection Counted? In *Complex Networks & Their Applications X. COMPLEX NETWORKS 2021* (R. M. Benito, C. Cherifi, H. Cherifi, E. Moro, L. M. Rocha, and M. Sales-Pardo, eds). Studies in Computational Intelligence **1016**. Springer, Cham, 229-240. DOI: 10.1007/978-3-030-93413-2_20

19. V. Marbukh. Towards Robust Security Risk Metrics for Networked Systems: Work in Progress. In *Proceedings of IFIP/IEEE International Symposium on Integrated Network Management*, Online,

May 17-21, 2021, 658-661. URL: https://dl.ifip.org/db/conf/im/im2021short/211558.pdf

20. V. Marbukh. Towards Economically Efficient Security Risk Reduction. In *Proceedings of the 31st European Safety and Reliability Conference (ESREL 2021)*, Angers, France, September 19-23, 2021, 1199-1199.

21. V. Marbukh. Towards Economically Efficient Mitigation of Systemic Risk of Undesirable Contagion. In *Proceedings of the 31st European Safety and Reliability Conference (ESREL 2021)*, Angers, France, September 19-23, 2021, 1206-1206.

22. O. Narayan, I. Saniee, and V. Marbukh. Congestion Due to Random Walk Routing. In *Complex Networks & Their Applications X. COMPLEX NETWORKS 2021* (R. M. Benito, C. Cherifi, H. Cherifi, E. Moro, L. M. Rocha, and M. Sales-Pardo, eds). Studies in Computational Intelligence **943**. Springer, Cham. DOI:10.1007/978-3-030-65347-7_46

23. M. Särestöniemi, C. Pomalaza-Raez, K. Sayrafian, and J. Iinatti. In-Body Propagation at ISM and UWB Frequencies for Abdominal Monitoring Applications. In *Proceedings of the 4th IEEE International Workshop on IoT-Health, IEEE International Conference on Communications 2021*, Online, June 14-23, 2021. DOI:10.1109/ICCWorkshops50388.2021.9473828

24. M. Särestöniemi, C. Pomalaza-Raez, K. Sayrafian, T. Myllylä, and J. Iinatti. A Preliminary Study of RF Propagation for High Data Rate Brain Telemetry. In *Proceedings of the 16th EAI International Conference on Body Area Networks (BodyNets 2021)*, Online, October 25-26, 2021.

25. B. Saunders. Complex Variables, Mesh Generation, and 3D Web Graphics: Research and Technology Behind the Visualizations in the NIST Digital Library of Mathematical Functions. In *Proceedings of the Golden Anniversary Celebration of the National Association of Mathematicians*, AMS Contemporary Mathematics Series **759**, 2020, 145-156. DOI: 10.1090/conm/759/15272

26. O. Slattery, L. Ma, X. Tang, T. Gerrits, A. Rahmouni, and S. Bhushan. The Quantum Communications and Networking Project at the Information Technology Laboratory of NIST. In *Proceedings of the 2021 IEEE Conference on Communications and Network Security: Workshop on Quantum Communication and Quantum Cryptography*, Virtual Conference, October 4-6, 2021.

27. S. Su, S. Kase, C. Hung, J. Hare, B. Rinderspacher, and C. Amburn. Mixed Reality Visualization of Friendly vs Hostile Decision Dynamics. In *Proceedings of the Virtual, Augmented and Mixed Reality. HCII 2021*. Lecture Notes in Computer Science **12770**. Springer, Cham. DOI: 10.1007/978-3-030-77599-5_37

28. J. Zwolak, S. Kalantre, T. McJunkin, B. Weber, and J. Taylor. A Ray-based Classification Framework for High-Dimensional Data. In *Proceedings of the Machine Learning and the Physical Sciences Workshop at NIPS 2020 (NeurIPS 2020),* Vancouver, Canada, December 11, 2020.

Technical Reports

1. J. Bernal, J. Lawrence, G. Doğan, and C. Hagwood. On Computing Elastic Shape Distances between Curves in d-dimensional Space. NIST Technical Note 2164, June 2021, 39 pages. DOI: 10.6028/NIST.TN.2164

2. R. Boisvert, ed. Applied and Computational Mathematics Division. Summary of Activities for Fiscal Year 2020. NIST-IR 8373, May 2021, 162 pages. DOI: 10.6028/NIST.IR.8373

3. R. Boisvert (as working group member). Reproducibility Badging and Definitions, NISO RP-31-2021, National Information Standards Organization, January 15, 2021. DOI: 10.3789/niso-rp-31-2021

4. L. Espinal, C. Young, and J. Zwolak. Mapping Employee Networks Through the NIST Interactions Survey. NISTIR 8375, June 2021, 53 pages. DOI: 10.6028/NIST.IR.8375

5. E. Fleisig and G. Doğan. VEMOS: A GUI for Evaluation of Similarity Metrics on Complex Data Sets. NIST Technical Note 2160, June 2021, 34 pages. DOI: 10.6028/NIST.TN.2160

6. A. Kearsley and A. Moorthy. Identifying Fentanyl with Mass Spectral Libraries. 2021. DOI: 10.26434/chemrxiv.14176952.v1

7. M. Theofanos, J. Evans, J. Zwolak, and S. Prettyman. Survey on Gender, Equity and Inclusion. NISTIR 8362*,* March 2021, 54 pages. DOI: 10.6028/NIST.IR.8362

8. J. Wu and R. Kacker. Standard Errors and Significance Testing in Data Analysis for Testing Classifiers. NISTIR 8383, July 2021, 22 pages. DOI: 10.6028/NIST.IR.8383

Blog Posts

1. S. Ressler. Will LiDAR Break the 3D Asset Creation Bottleneck? Khronos Group Blog, September 15, 2021. URL: https://www.khronos.org/blog/will-lidar-break-the-3d-asset-creation-bottleneck

2. M. Theofanos and J. Zwolak. Diversity and Inclusivity at NIST. Women in Standards News, May 30, 2021. URL: https://womeninstandards.org/diversity-and-inclusivity-at-nist/

Accepted

1. H. Ahmadi, A. Nag, Z. Khan, <u>K. Sayrafian,</u> and S. Rahadrja. Networked Twins and Twins of Networks: An Overview on the Relationship Between Digital Twins and 6G. *IEEE Communications Standards Magazine.*

2. <u>D. Anderson</u>, P. Guba, and A. Wells. Mushy Layer Convection. *Physics Today*, February 2022.

3. Z. Chen, J. Wen, <u>A. Kearsley,</u> and A. Pertzborn. Evaluating the Performance of an Inexact Newton Method with a Preconditioner for Dynamic Building System Simulation. *Journal of Building Performance Simulation.*

4. <u>H. Cohl</u> and R. Costas-Santos. Utility of Integral Representations for Basic Hypergeometric Functions and Orthogonal Polynomials. *The Ramanujan Journal, Richard A. Askey Memorial Volume.*

5. <u>H. S. Cohl</u> and R. S. Costas-Santos. On the Relation Between Gegenbauer Polynomials and the Ferrers Function of the First Kind. *Analysis Mathematica.*

6. <u>H. Cohl</u>, M. Ismail, and H. Wu. The Legacy of Dick Askey (1933-2019). *Notices of the American Mathematical Society.*

7. <u>H. Cohl</u> and M. Ismail. Liber Amicorum, a Friendship Book for Dick Askey. *Celebratio Mathematica.*

8. D. Faccini, F. Maggioni, and <u>F. Potra</u>. Robust and Distributionally Robust Optimization. *Computers and Operations Research.*

9. R. Fitzgerald, <u>B. Alpert</u>, D. Becker, D. Bergeron, R. Essex, K. Morgan, S. Nour, G. O'Neil, D. Schmidt, G. Shaw, D. Swetz, R. Verkouteren, and D. Yan. Towards a New Primary Standardization of Radionuclide Massic Activity Using Microcalorimetry and Quantitative Milligram-Scale Samples. *Journal of Research of NIST.*

10. A. Greiner-Petter, <u>H. Cohl</u>, A. Youssef, M. Schubotz, A. Trost, R. Dey, A. Aizawa, and B. Gipp. Comparative Verification of the Digital Library of Mathematical Functions and Computer Algebra Systems. *International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2021).*

11. <u>Z. Grey</u>, S. Mosleh, J. Rezac, Y. Ma, J. Coder, and <u>A. Dienstfrey</u>. Bi-Criteria Radio Spectrum Sharing with Subspace-based Pareto Tracing. *IEEE Transactions on Communications.*

12. Z. Guo, J. Song, G. Barbastathis, M. Glinsky, C. Vaughan, K. Larson, <u>B. Alpert,</u> and Z. Levine. Advantage of Machine Learning over Maximum Likelihood in Limited-Angle Low-Photon X-Ray Tomography. *Machine Learning for Scientific Imaging 2022 Conference.*

13. R. Kacker. On Quantity, Value, Unit, and Other Terms in the JCGM International Vocabulary of Metrology. *Measurement Science and Technology.*

14. <u>R. Kacker</u>, D. Kuhn, <u>Y. Lei</u>, and <u>D. Simos</u>. Measuring Adequacy of a Test Suite with Respect to the Modeled Test Space. *IEEE Software.*

15. V. Mai, <u>R. La</u>, T. Zhang, and A. Battou. End-to-End Quality-of-Service Assurance with Autonomous Systems: 5G/6G Case Study. *IEEE Consumer Communications and Networking Conference (CCNC)*, virtual, January 8-11, 2022.

16. V. Marbukh. Towards Risk Adjusted Wireless Access under Jamming: Reliability through Multiconnectivity. *IEEE Consumer Communications & Networking Conference (IEEE CCNC 2022).*

17. C. Rao, N. Li, Y. Le, J. Guo, Y. Zhang, <u>R. Kacker,</u> and D. Kuhn. Combinatorial Test Generation for Multiple Input Models with Shared Parameters. *IEEE Transactions on Software Engineering.*

18. J. Sadeghi, <u>P. Patrone</u>, <u>A. Kearsley</u>, and G. Cooksey. Optofluidic Flow Meter for Sub-Nanoliter per Minute Flow Measurements. *Journal of Biomedical Optics.*

19. N. Saito, <u>N. Douguet</u>, H. Sannohe, N. Ishii, T. Kanai, Y. Wu, A. Chew, S. Han, <u>B. Schneider</u>, J. Olsen, <u>L. Argenti</u>, Z. Chang, and J. Itatani. Attosecond Electronic Dynamics of Core-Excited States of N_2O in the Soft X-ray Region. *Physical Review Research.*

20. <u>K. Sayrafian</u>, <u>B. Cloteaux</u>, <u>V. Marbukh</u>, and C. Emiyah. Evaluation of the Bluetooth-based Proximity Estimation for Automatic Exposure Determination. *Proceedings of the IEEE Consumer Communications & Networking Conference (IEEE CCNC)*, Online, January 8-11, 2022.

21. H. Zhao, J. Chen, G. Bryant, W. Griffin, and <u>J. Terrill</u>. Evaluating Alternative Glyph Design for Showing Large-Magnitude-Range Quantum Spins. *IEEE Transactions on Visualization and Computer Graphics.*

In Review

1. <u>V. Albert</u>, D. Aasen, W. Xu, W. Ji, J. Alicea, and J. Preskill. Spin Chains, Defects, and Quantum Wires for the Quantum-Double Edge.

2. I. Bezerra, H. Vasconcelos, and <u>S. Glancy</u>. Quadrature Squeezing and Temperature Estimation from The Fock Distribution.

3. <u>D. Brager</u>, E. Camacho, T. Léveillard, and S. Wirkus. Mathematically Modeling Photoreceptor Degeneration in Retinitis Pigmentosa.

4. <u>D. Brager</u>, E. Camacho, T. Léveillard, and S. Wirkus. A Mathematical Investigation of the Progression of Retinitis Pigmentosa.

5. A. M. Childs, <u>M. Coudron</u>, A. S. Gilani, and M. Kovacs-Deak. Quantum Algorithms and the Power of Forgetting.

6. B. Cloteaux. How Much Regularity Forces a Sequence to be Graphic?

7. B. Cloteaux. Fast Graphic Approximation of Very Large Integer Sequences.

8. <u>H. Cohl</u>, R. Costas-Santos, P. Hwang, and T. Wakhare. Generalizations of Generating Functions for Basic Hypergeometric Orthogonal Polynomials.

9. R. Dalka and <u>J. Zwolak</u>. Restoring the Structure: A Modular Analysis of Ego-Driven Organizational Social Networks.

10. L. Ding and <u>L. Ma</u>. Hybrid Quantum-Edge Computing: A New Computing Paradigm.

11. M. DiSalvo, <u>P. Patrone</u>, <u>A. Kearsley,</u> and G. Cooksey. Serial Flow Cytometry in an Inertial Focusing Optofluidic Microchip for Direct Assessment of Measurement Variations.

12. J. Fowler, <u>B. Alpert</u>, G. O'Neil, D. Swetz, and J. Ullom. Energy Calibration of Nonlinear Transition-Edge Sensors with Uncertainty Estimates from Gaussian Process Regression.

13. B. Garn, D. Schreiber, <u>D. Simos</u>, D. Kuhn, J. Voas, and R. Kacker. Combinatorial Methods for Testing IoT Smart Home Systems.

14. B. Garn, <u>D. Simos</u>, S. Zauner, D. Kuhn, and <u>R. Kacker</u>. A Two-Step TLS-Based Browser Fingerprinting Approach using Combinatorial Sequences.

15. <u>S. Geller</u>, D. Cole, <u>S. Glancy,</u> and <u>E. Knill</u>. Improving Quantum State Detection with Adaptive Sequential Observations.

16. <u>T. Gerrits</u>, I. Burenkov, Y. Li-Baboud, A. Rahmouni, D. Anand, Hala, A. Battou, <u>O. Slattery,</u> and S. Polyakov. White-Rabbit-Assisted Quantum Network Node Synchronization with Quantum Channel Coexistence.

17. M. Gockenbach and <u>M. Roberts</u>. Approximating the Generalized Singular Value Expansion: Error Estimates Part 1.

18. M. Gockenbach and <u>M. Roberts</u>. Approximating the Generalized Singular Value Expansion: Error Estimates Part 2.

19. S. Guo, S. Koh, A. Fritsch, I. Spielman, and <u>J. Zwolak</u>. Combining Machine Learning with Physics: A Framework for Tracking and Sorting Multiple Dark Solitons.

20. H. Huang, R. Kueng, G. Torlai, <u>V. Albert,</u> and J. Preskill. Provably Efficient Machine Learning for Quantum Many-Body Problems.

21. L. Kampel, <u>D. Simos</u>, D. Kuhn, and <u>R. Kacker</u>. An Exploration of Combinatorial Testing-based Approaches to Fault Localization for Explainable AI.

22. S. Kase, C. Hung, T. Krayzman, J. Hare, B. Rinderspacher, and <u>S. Su</u>. The Future of Collaborative Human-AI Decision-making for Mission Planning.

23. I. Kim, B. Shi, K. Kato, and <u>V. Albert</u>. Chiral Central Charge from a Single Bulk Wave Function.

24. I. Kim, B. Shi, K. Kato, and <u>V. Albert</u>. Modular Commutator in Gapped Quantum Many-Body Systems.

25. P. Kuo. Noncritical Phasematching Behavior in Thin-film Lithium Niobate Frequency Converters.

26. V. Mai, <u>R. La</u>, T. Zhang, and A. Battou. Distributed Optimization with Global Constraints Using Noisy Measurements.

27. L. Melara, <u>R. Evans</u>, S. Cho, A. Balijepalli, and <u>A. Kearsley</u>. Optimal Bandwidth Selection in Stochastic Regression of Bio-FET Measurements.

28. L. Morais, T. Weinhold, M. de Almeida, A. Lita, <u>T. Gerrits</u>, S. Nam, A. White, and G. Gillett. Precisely Determining Photon-Number in Real-Time.

29. <u>A. Rahmouni</u>, <u>T. Gerrits</u>, A. Migdall, <u>O. Slattery</u>, P. Shaw, and J. Rice. A Self-Validated Detector for Characterization of Quantum Network Components.

30. M. Roudneshin, <u>K. Sayrafian,</u> and A. Aghdam. Maximizing Harvested Energy in Coulomb Force Parametric Generators.

31. <u>B. Saunders</u>, S. Brooks, R. Buckmire, R. Vincent-Finley, F. Backeljauw, S. Becuwe, <u>B. Miller</u>, <u>M. McClain,</u> and A. Cuyt. Validated Computation of Special Mathematical Functions.

32. A. Seif, M. Hafezi, and <u>Y.-K. Liu</u>. Compressed Sensing Measurement of Long-Range Correlated Noise.

33. R. Srinivas, S. Burd, H. Knaack, R. Sutherland, A. Kwiatkowski, <u>S. Glancy</u>, <u>E. Knill</u>, D. Wineland, D. Leibfried, A. Wilson, D. Allcock, and D. Slichter. High-fidelity rf/Microwave-based Universal Control of Trapped Ion Qubits.

34. J. Stone, K. Griffin, J. Amstutz, D. DeMarle, <u>W. Sherman,</u> and J. Günther. ANARI: A 3D Rendering API Standard.

35. Y. Tong, <u>V. Albert</u>, J. McClean, J. Preskill, and Y. Su. Provably Accurate Simulation of Gauge Theories and Bosonic Systems.

36. M. Vazquez, T. Berry, T. Sauer, and <u>G. Doğan</u>. Texture Segmentation from a Manifold Learning Perspective.

37. C. Qu, <u>A. Kearsley</u>, <u>B. Schneider</u>, W. Keyrouz, and T. Allison. Graph Convolutional Neural Network Applied to the Prediction of Normal Boiling Points.

38. <u>J. Ziegler</u>, T. McJunkin, E. Joseph, S. Kalantre, B. Harpt, D. Savage, M. Lagally, M. Eriksson, J. Taylor, and <u>J. Zwolak</u>. Toward Robust Autotuning of Noisy Quantum Dot Devices.

Inventions

Patents Awarded

1. G. Cooksey, <u>A. Kearsley,</u> and <u>P. Patrone</u>. Optical Flow Meter for Determining a Flow Rate of a Liquid. Patent US 11035707, June 15, 2021.

2. J. Ullom, G. O'Neil, L. Avila, K. Silverman, D. Swetz, R. Jimenez, W. Doriese, G. Hilton, C. Reintsema, D. Schmidt, <u>B. Alpert</u>, J. Uhlig, Y. Joe, W. Fullagar, V. Sundstrom, I. Maasilta, and J. Fowler. X-Ray Spectrometer. Patent US 10,914,694 B2, Feb. 9, 2021.

Invention Disclosures and Patents in Review

1. G. Cooksey, <u>A. Kearsley,</u> and <u>P. Patrone</u>. Multiplexed Amplitude Modulation Photometer and Performing Multiplexed Amplitude Modulation Photometry. Provisional Patent Application 20210055201.

2. G. Cooksey, <u>A. Kearsley,</u> and <u>P. Patrone</u>. Serial Flow Cytometer. Provisional Patent Application 20210302300.

3. <u>A. Kearsley</u>, <u>P. Patrone</u>, E. Romsos, and P. Vallone. System and Method for Data Analysis in Quantitative PCR Measurements. Provisional Patent Application 20210395807.

4. K. Sayrafian. Lung Fluid Monitor and Determining Fluid Level in a Lung. Provisional Patent Application Serial Number 63/272,783, NIST Docket # 21-046P1, October 28, 2021.

5. A. Seif, M. Hafezi, and <u>Y.-K. Liu</u>. Compressed Sensing Measurement of Long-Range Correlated Noise. NIST Invention Disclosure, NIST Docket #21-020, February 19, 2021.

6. <u>J. Ziegler</u> and <u>J. Zwolak</u>. Framework for Robust and Efficient Calibration of Quantum Dot Devices to the Single-Electron Regime. NIST Invention Disclosure, NIST Docket #21-057, August 11, 2021.

Presentations

Note: When multiple presenters are listed, names of co-presenters with an ACMD affiliation during this reporting period are <u>underlined.</u>

Invited Talks

1. V. V. Albert. "Spin Chains, Defects, and Quantum Wires for the Quantum-Double Edge." Caltech Condensed Matter Theory Seminar, California Institute of Technology, Online, September 23, 2021.

2. V. V. Albert. "Overview of Quantum Research At UMD." Joint Quantum Institute Friday Seminar, University of Maryland, College Park, MD, September 3, 2021.

3. V. V. Albert. "Robust Encoding of a Qubit in a Molecule." 52nd Annual Meeting of the American Physical Society Division of Atomic, Molecular and Optical Physics (DAMOP), Online, June 3, 2021.

7. V. V. Albert. "Applications of QSTEM at UMD." Society of Physics Students Faculty Talk, University of Maryland, College Park, MD, May 7, 2021.

4. V. V. Albert. "Robust Encoding of a Qubit in a Molecule." Joint Center for Quantum Information and Computer Science (QuICS) Stakeholder Day, University of Maryland, Online, March 3, 2021.

5. V. V. Albert. "Lindbladians with Multiple Steady States." Workshop on Many-Body Physics in Open Quantum Systems, Princeton Center for Theoretical Science, Online, January 21, 2021.

6. D. Anderson, S. Coriell, P. Guba, <u>G. McFadden,</u> and B. Murray. "Convective Instabilities in Ternary Alloy Solidification." Oxford Center for Industrial and Applied Mathematics Seminar, Oxford University, Online, November 25, 2020.

7. R. Boisvert. "Special Functions at NIST: From Mathematical Tables to Digital Libraries." Department of Mathematics, University of Wyoming, Online, November 13, 2020.

8. <u>D. Brager</u>, E. Camacho, T. Léveillard, and S. Wirkus. "Mathematically Investigating Retinitis Pigmentosa." Applied and Computational Math Seminar, George Mason University, Fairfax, VA, May 7, 2021.

9. D. Brager, E.T. Camacho, T. Léveillard, and S. Wirkus. "Mathematically Investigating Retinitis Pigmentosa." Systems Medicine Seminar, University of Florida, Gainesville, FL, December 16, 2021.

10. R. Buckmire, S. Brooks, R. Vincent-Finley, and B. Saunders. "Validated Computation of Special Functions II: Error Analysis." ADJOINT (African Diaspora Joint Mathematics Workshop) Research Showcase, 2021 Joint Mathematics Meetings, Online, January 9, 2021.

11. M. Coudron. "Computations with Greater Quantum Depth Are Strictly More Powerful (Relative to an Oracle)." Mathematics Department, University of Maryland, November 2020.

12. M. Coudron. "Approximating Output Probabilities of Shallow, Geometrically-local Quantum Circuits." QuICS Stakeholder Day, University of Maryland, Online, March 3, 2021.

13. M. Coudron. "Simulating Shallow Quantum Circuits" Combinatorics, Algorithms, and AI for Real Problems Research Experience for Undergraduates (REU) Seminar, June 2021.

14. G. Doğan. "Topology Optimization for Image Analysis." SIAM Conference on Computational Science and Engineering, Online, March 1-5, 2021.

15. J. Fong. "What is ASME Boiler and Pressure Vessel Code Section V (NDE) Article 14 and Why It Needs to be Revised." ASME BPVC Section V Code Committee Task Group on Assisted Analysis Committee Meeting, Online, February 5, 2021.

16. J. Fong. "Reliability and Reliability Targets for Four Cases of AISI 4340 Steel Pipes with Cracks." ASME BPVC Section XI (Inspection) Subgroup on Reliability and Integrity Management Working Group on Monitoring and NDE Subcommittee Meeting, Online, February 7, 2021.

17. J. Fong. "Reliability and Reliability Targets for Four Cases of AISI 4340 Steel Pipes with Cracks." ASME BPVC Section XI (Inspection) Subgroup (SG) on Reliability and Integrity Management (RIM) Committee Meeting, Online, February 8, 2021.

18. J. Fong. "A Progress Report on Research to Upgrade Table T-1472.1 of ASME BPVC Section V (NDE) Article 14 (System Qualification)." ASME BPVC Section XI (Inspection) Subgroup on Reliability and Integrity Management Working Group on Monitoring and NDE Subcommittee Meeting, Online, October 31, 2021.

19. T. Gerrits and M. Stephens. "Optical Power Traceability and Calibrations for Quantum Network Components." Quantum Network Grand Challenge Seminar Series, NIST, Online, July 30, 2021.

20. T. Gerrits. "Optical Quantum Network Metrology." SPIE Photonics West – Optical and Quantum Sensing and Precision Metrology, Online, March 2021.

21. Z. Grey, S. Mosleh, J. Rezac, Y. Ma, J. Coder, and A. Dienstfrey. "Optimizing Unlicensed Band Spectrum Sharing with Subspace-Based Pareto Tracing." IEEE International Conference on Communications, Online, June 14-23, 2021.

22. Z. Grey, "Global Adjoint-Based Shape Sensitivity Analysis." SIAM Conference on Optimization, Online, July 2021.

23. Z. Grey, "Subspace-based Dimension Reduction with Manifold Extensions." NASA JPL Verification and Validation Uncertainty Quantification Seminar, August 2021.

24. Z. Grey, "Multi-Criteria Radio Spectrum Sharing with Subspace-based Pareto Tracing." NIST CTL MLComSIG Seminar, September 2021.

25. Z. Grey, "Active Manifold-Geodesics with Aerospace Applications." Air Force Research Laboratory Miracle Forum, October 2021.

26. A. Kearsley. "A Survey of Applied Mathematics Applications in Metrology at NIST." MAA MD-DC-VA Section Meeting, Online, November 7, 2020.

27. A. Kearsley, D. Anderson, and J. Benson. "Control of Inward Solidification in Cryobiology." American University, Department of Mathematics Seminar, November 10, 2020.

28. A. Kearsley and P. Patrone. "Optimization and Data Analysis for Improved COVID19 Detection and Measurement." SIAM Annual Meeting, Workshop on Celebrating Diversity, July 21, 2021.

29. A. Kearsley, D. Anderson, and J. Benson. "Control of Inward Solidification in Cryobiology." University of Colorado, Boulder, Mathematical Biology Seminar, October 10, 2021.

30. A. Kearsley. "NIST Applications of Optimization and Data Analysis to COVID19 Detection and Measurement." Shippensburg University, Student Mathematics Seminar. October 25, 2021.

31. P. Kuo. "Domain-engineered PPLN for Tailored Spontaneous Parametric Down Conversion." Novel Optical Materials and Applications Conference, 2021 OSA Advanced Photonics Congress, Online, July 26-30, 2021.

32. P. Kuo. "Photons for Quantum Networking." Physics Department Seminar, University of Maryland, College Park, MD, February 24, 2021.

33. R. La, "Minimization of Aggregate Cybersecurity Costs in Complex Systems: SIS Model-based Approach." SIAM Conference on Applied Linear Algebra, Online, May 20, 2021.

34. L. Ma. "Quantum Communications and Networks for Edge Computing." Keynote Presentation, Sixth ACM/IEEE Symposium on Edge Computing, San Jose, CA, December 17, 2021.

35. J. Nolan, "Hitting Objects with Random Walks." Math/Stat Department Colloquium, American University, Washington, DC, October 13, 2020.

36. P. Patrone and A. Kearsley. "Optimal Decision Theory for Antibody Testing." Seronet Monthly Meeting, May 13, 2021.

37. P. Patrone. "Take Data Analysis Seriously: You Might Save Lives." MAA MD-DC-VA Section Meeting, April 24, 2021.

38. M. Roberts. "Approximating the Generalized Singular Value Expansion." Applied and Computational Mathematics Seminar, George Mason University Fairfax, VA, October 22, 2021.

39. B. Saunders. "From Abramowitz and Stegun to the NIST Digital Library of Mathematical Functions and Beyond." History of Mathematics Class, Prof. Hortensia Soto (MAA Associate Secretary), Colorado State University, Online, November 10, 2020.

40. B. Saunders, S. Brooks, R. Buckmire, and R. Vincent-Finley. "Validated Computation of Special Functions I: Overview." African Diaspora Joint Mathematics Workshop (ADJOINT) Research Showcase, 2021 Joint Mathematics Meetings, Online, January 9, 2021.

41. B. Saunders. "Complex Functions, Mesh Generation, and Hidden Figures in the NIST Digital Library of Mathematical Functions." AWM-MAA Etta Zuber Falconer Lecture, MAA MathFest 2021, Online, August 6, 2021.

42. B. Saunders. "Discussion on Diversity and Inclusion for Careers in CSE." SIAM Conference on Computational Science and Engineering, Online, March 1, 2021.

43. B. Saunders. "Careers in Mathematics." Mathematical Association of America Texas Section Meeting, Online, April 9, 2021.

44. B. Saunders. "Career Paths in Business, Industry, and Government." MAA MathFest 2021, Online, August 6, 2021.

45. O. Slattery. "Quantum Communications and Networking Project - An Introduction to Quantum Communications and Cryptography." at Advanced Technology Academic Research Center, Online, December 22, 2021.

46. K. Sayrafian. "Wireless Wearables and Implants Technology: Shaping the Future of Health-IT." IEEE Eta Kappa Nu 2020, Online, November 3, 2020.

47. B. Schneider, S. Pamadighantam, A. Scrinzi, K. Hamilton, K. Bartschat, O Zatsarinny, I. Bray, L. Carr, F. Martin, J. Gonzalez, R. Lucchese, A. Brown, C. Fischer, N. Douguet, and S. dos Santos. "A Science Gateway for Atomic and Molecular Physics: Democratizing Atomic and Molecular Physics Research and Education." 74th Gaseous Electronics Conference, October 7, 2021.

48. W. Sherman. "Immersive Visualization: the role of VR in HPC data analysis." Keynote presentation, Collaborative Conference on Computational and Data Intensive Science, Online, July 7, 2021.

49. W. Sherman. "Immersive Visualization: The Role of VR in Data Analysis." NASA XR Community of Interest, July 29, 2021.

50. E. Shojaee. "Tomography of Entangled Macroscopic Mechanical Objects." Quantum Information Technologies in Louisiana Day, Online, December 20, 2020.

51. O. Slattery "The Quantum Communications and Networking Project at the Information Technology Laboratory at NIST." IEEE International Workshop on Quantum Communication and Quantum Cryptography, October 6, 2021.

52. R. Vincent-Finley, S. Brooks, R. Buckmire, and B. Saunders. "Validated Computation of Special Functions III: DLMF Tables." African Diaspora Joint Mathematics Workshop (ADJOINT) Research Showcase, 2021 Joint Mathematics Meetings, Online, January 9, 2021.

53. R. Vincent-Finley, S. Brooks, R. Buckmire and B. Saunders. "Validated Computation of Special Mathematical Functions." Mathematical Sciences Research Institute (MSRI) African Diaspora Joint Mathematics Workshop (ADJOINT) Colloquium, Online, April 27, 2021.

54. J. Zwolak and G. Doğan. "Mathematical and Computational Science Research at NIST." Center of Mathematics and Artificial Intelligence Industry Day, George Mason University, Online, June 25, 2021.

55. J. Zwolak. "Ray-based Framework for Tuning Quantum Dot Devices: Two Dots and Beyond."

Seminar of the Chair of Mathematical Physics, Nicolaus Copernicus University, Toruń, Poland, November 10, 2020.

56. J. Zwolak. "Ray-based Framework for Tuning Quantum Dot Devices: Two Dots and Beyond." Quantum Physics Department Seminar, Nicolaus Copernicus University, Toruń, Poland, January 12, 2021.

57. J. Zwolak. "Learning the States of Quantum Dot Systems: A Ray-Based Framework." 2021 SIAM Conference on Computational Science and Engineering, Online, March 1-5, 2021.

58. J. Zwolak. "Ray-based Framework for Tuning Quantum Dot Devices: Two Dots and Beyond." Machine Learning for Quantum Conference, Online, March 1-5, 2021.

59. J. Zwolak. "Ray-based Framework for Tuning Quantum Dot Devices: Two Dots and Beyond." American Association of Physics Teachers Summer Meeting 2021, Online, August 8, 2021.

Conference Presentations

1. V. Albert. "Robust Encoding of a Qubit in a Molecule." Young Researchers Symposium, International Congress of Mathematical Physics, Online, July 30, 2021.

2. V. Albert. "Filling Cavities to Prevent Decay: Bosonic Quantum Error Correction." Quantum Network Grand Challenge Seminar Series, NIST, Online, June 25, 2021.

3. V. Albert. "Quantum Doubles in One Dimension." 2021 APS March Meeting, Online, March 17, 2021.

4. V. Albert. "Robust Encoding of a Qubit in a Molecule." APS March Meeting, Online, March 17, 2021.

5. S. Bhushan, L. Ma, X. Tang and O. Slattery. "Electromagnetically Induced Transparency Based Quantum Memory in Vapor Cell." IEEE International Workshop on Quantum Communication and Quantum Cryptography, Online, October 6, 2021.

6. D. Brager, E. Camacho, T. Léveillard, and S. Wirkus. "Mathematically Investigating Retinitis Pigmentosa." Society for Mathematical Biology Annual Meeting, Online, June 13-17, 2021.

7. D. Brager, E. Camacho, T. Léveillard and S. Wirkus. "Mathematically Investigating Retinitis Pigmentosa." SIAM Annual Meeting, Online, July 19-23, 2021.

8. N. J. Coble and M. Coudron. "Quasi-polynomial Time Approximation of Output Probabilities of Constant-depth, Geometrically-local Quantum Circuits." 24th Annual Conference on Quantum Information Processing (QIP/21), Online, February 2, 2021.

9. R. DeJaco. "Resolving the Shock Layer in Fixed-Bed Adsorption with Boundary Layer Theory." Center for Mathematics and Artificial Intelligence Colloquium, George Mason University, Online, May 7, 2021.

10. R. DeJaco. "Resolving the Shock Layer in Fixed-Bed Adsorption with Boundary Layer Theory," SIAM Conference on Mathematical Aspects of Materials Science, Online, May 20th, 2021.

11. R. DeJaco. "Boundary Layer Theory in Fixed-Bed Adsorption," American Institute of Chemical Engineering Annual Meeting, Online, November 10th, 2021.

12. G. Doğan. "Scikit-Shape: Python Toolbox for Shape Analysis and Segmentation." SIAM Conference on Geometric and Physical Modeling, Online, September 27-29, 2021.

13. M. Donahue and D. Porter. "High-Order Methods for Computing the Demagnetization Tensor for Periodic Boundaries." Magnetism and Magnetic Materials Conference, Online, November 6, 2020.

14. J. Fong. "A Multi-Scale Failure-Probability-and-NDE-Based Fatigue Life Model for Estimating Component Co-Reliability of Uncracked and Cracked Pipes." ASME Pressure Vessel and Piping Division Conference, Online, July 13, 2021.

15. T. Gerrits, A. Migdall, J. Bienfang, J. Lehman, S. Nam, J. Splett, I. Vayshenker, and J. Wang. "Calibration of Free-space and Fiber-coupled Single-photon Detectors." International Conference on New Developments and Applications in Optical Radiometry (Newrad), Online, June 21-24, 2021.

16. H. Gharibnejad, N. Douguet, B. Schneider, J. Olsen, and L. Argenti. "A Multi-Center Quadrature Scheme for the Molecular Continuum." International Conference on the Physics of Electronic and Atomic Collisions, Online, July 20-23, 2021.

17. Z. Gimbutas. "Far Field Translation Operators for the Helmholtz Equation in Three Dimensions." SIAM Conference on Computational Science and Engineering, Online, March 1, 2021.

18. Z. Gimbutas. "Edge and Corner Pre-conditioners in Three Dimensions." SIAM Annual Meeting, Online, July 21, 2021.

19. M. Hu, E. Hoefer, and <u>M. Donahue</u>. "Antiferromagnetic Spin Wave Propagation on Nonuniform Magnetic Backgrounds." Magnetism and Magnetic Materials Conference, Online, November 5, 2020.

20. A. Kwiatkowski. "Constraints on Gaussian Error Channels for Entanglement Swapping with Gaussian Measurements." APS March Meeting, Online, March 18, 2021.

21. R. La. "Story of Two Populations in Epidemics: Is Every Infection Counted?" International Conference on Complex Networks and Their Applications, Hybrid (Madrid, Spain), November 30 - December 2, 2021.

22. S. Langer. "Calculating Voxel-Polyhedron Intersections for Meshing Images." 29th International Meshing Roundtable, Online, June 21-25, 2021.

23. J. Lundstrom, M. Poorman, <u>A. Dienstfrey,</u> and K. Keenan. "Nonlinear Optimization for Enhanced Parameter Retrieval for Magnetic Resonance Fingerprinting." 2021 Annual Meeting of the APS Four Corners Section, Boulder, CO, October 8-9, 2021.

24. N. Martys, <u>W. George</u>, S. Jones, and A. Goyal. "Physics Based Approach to DEM." American Concrete Institute Concrete Convention, Online, March 28 - April 1, 2021.

25. M. Poorman, <u>Z. Gimbutas</u>, D. Ma, <u>A. Dienstfrey,</u> and K. Keenan. "Uncertainty Analysis Framework for Quantifying Error Propagation in Magnetic Resonance Fingerprinting." The International Society for Magnetic Resonance in Medicine Annual Meeting, Online, May 17, 2021.

26. D. Porter. "IMMUTABLE Values & Data Structures." SQLite & Tcl Conference 2021, Online, November 17, 2021.

27. <u>A. Rahmouni</u>, <u>T. Gerrits</u>, A. Migdall and <u>O. Slattery</u>. "A Self-validated Detector for Characterization of Quantum Network Component." IEEE International workshop on Quantum Communication and Quantum Cryptography, Online, October 6, 2021.

28. J. Randazzo, <u>J. Olsen</u>, <u>H. Gharibnejad</u>, <u>B. Schneider</u> and <u>L. Argenti</u>. "ASTRA, a New Close-Coupling Approach for Molecular Ionization. International Conference on the Physics of Electronic and Atomic Collisions, Online, July 20-23, 2021.

29. M. Ruiz, <u>J. Sims,</u> and B. Padhy. "Exponentially Correlated Hylleraas-Configuration-Interaction Nonrelativistic Energy of the singlet S Ground State of the Helium Atom." Warsaw Electronic Structure Conference, Online, September 1-4, 2020.

30. W. Sherman. "ANARI: A Future for High-End Visualization Rendering." 6th Annual Campus Alliance for Advanced Visualization Conference, West Lafayette Indiana, November 3, 2021.

31. W. Sherman. "Building an At-home CAVE-style VR Display to Broaden User Access." 6th Annual Campus Alliance for Advanced Visualization Conference, West Lafayette Indiana, November 3, 2021.

32. <u>J. Ziegler</u>, T. McJunkin, E. Joseph, S. Kalantre, B. Harpt, D. Savage, M. Lagally, M. Eriksson, J. Taylor, and <u>J. Zwolak</u>. "Toward Robust Autotuning of Noisy Quantum Dot Devices." Silicon Quantum Electronics Workshop, Online, October 25-29, 2021.

33. <u>J. Ziegler</u>, S. Kalantre, T. McJunkin, M. Eriksson, J. Taylor, and <u>J. Zwolak</u>. "Towards Autonomous Tuning of Noisy Quantum Dots." APS March Meeting, Online, March 15-19, 2021.

34. <u>J. Zwolak</u>, T. McJunkin, S. Kalantre, S. Neyens, E. MacQuarrie, L. Edge, M. Eriksson, and J. Taylor. "Learning the States of Double Quantum Dot Systems: The Ray-based Approach." APS March Meeting 2021, Online, March 15-19, 2021.

Poster Presentations

1. V. Albert. "Quantum Doubles in One Dimension." 23rd Annual Conference on Quantum Information Processing, Online, February 4, 2021.

2. <u>L. Argenti</u>, <u>N. Douguet</u>, <u>H. Gharibnejad</u>, <u>B. Schneider</u>, J. Randazzo <u>J. Olsen</u>, and S. Bhushan. "Terahertz Electromagnetically Induced Transparency in Cesium Atoms." Conference on Frontier in Optics and Laser Science, Online, September 14-17, 2020.

3. <u>H. Gharibnejad</u>, <u>N. Douguet</u>, <u>B. Schneider</u>, <u>J. Olsen</u> and, <u>L. Argenti</u>. "A Multi-Center Quadrature Scheme for the Molecular Continuum." APS Division of Atomic, Molecular and Optical Physics Annual Meeting, Online, May 31 - June 4, 2021.

4. <u>D. Brager</u>, E. Erisman, A. Moorthy, and <u>A. Kearsley</u>. "Numerical Scaling for Pre-processing Mass Spectrometry Data." American Society for Mass Spectrometry Annual Conference, Philadelphia, PA, October 31 – November 4, 2021.

Figure 105. *Posters in COVID times. Danielle Brager (top) and Jake Roberts (bottom) presented posters at the American Society for Mass Spectroscopy Annual Meeting in Philadelphia in November 2021. Pictured below right is mentor Tony Kearsley.*

5. S. Cho, R. Evans, A. Kearsley, and A. Balijepalli. "High-Performance Electronic Measurements of Protein Kinetics and Affinity." APS March Meeting, Online, March 16, 2021.

6. G. Doğan. "Shape and Geometry Analysis in Images using Scikit-Shape." SIAM Conference on Optimization, Online, July 20-23, 2021.

7. Z. Grey, "Optimizing Unlicensed Band Spectrum Sharing with Subspace-Based Pareto Tracing." 28th Annual NIST Sigma Xi Postdoctoral Poster Presentations, Online, March 31, 2021.

8. S. Guo, A. Fritsch, C. Greenberg, I. Spielman, and J. Zwolak. "Towards Automated Phase Transition Detection with ML Tools." Harnessing Quantum Matter Data Revolution Symposium and Summer School, Online, June 7-11, 2021.

9. S. Guo, A. Fritsch, C. Greenberg, I. Spielman, and J. Zwolak. "Towards Automated Phase Transition Detection with ML Tools." 28th Annual NIST

Sigma Xi Postdoctoral Poster Presentation, Online, March 31, 2021.

10. S. Guo, A. Fritsch, C. Greenberg, I. Spielman, and J. Zwolak. "Towards Automated Phase Transition Detection with ML Tools." QuICS Stakeholder Day, Online, March 3, 2021.

11. P. Kuo. "Towards Noncritical Phasematching in Thin-film Lithium Niobate Frequency Converters." OSA Frontiers in Optics + Laser Science Conference, Online, November 1-4, 2021.

12. V. Marbukh. "Towards Robust Security Risk Metrics for Networked Systems: Work in Progress." IFIP/IEEE International Symposium on Integrated Network Management, Online, May 17-21, 2021.

13. P. Patrone and A. Kearsley. "Improving Serology Tests Through Optimal Decision Theory." Cold-Springs Harbor Labs COVID/SARS-CoV-2 Rapid Research Reports #5, January 26-27, 2021.

14. M. Roberts, A. Moorthy, E. Sisco, and A. Kearsley. "Continuous Spectral Approximation Orthogonal Distance Minimization in High Resolution DART-MS." American Society for Mass Spectrometry Annual Meeting, Philadelphia, Pennsylvania, November 1, 2021.

15. J. Ziegler, S. Kalantre, T. McJunkin, M. Eriksson, J. Taylor, and J. Zwolak. "Towards Autonomous Tuning of Noisy Quantum Dots." 28th Annual Sigma Xi Postdoctoral Poster Presentation, Online, March 31, 2021.

16. J. Ziegler, S. Kalantre, T. McJunkin, M. Eriksson, J. Taylor, and J. Zwolak. "Towards Autonomous Tuning of Noisy Quantum Dots." SIAM Conference on Computational Science and Engineering 2021, Online, March 3, 2021.

17. J. Ziegler, S. Kalantre, T. McJunkin, M. Eriksson, J. Taylor, and J. Zwolak. "Towards Autonomous Tuning of Noisy Quantum Dots." Machine Learning for Quantum, Online, March 2, 2021.

18. J. Zwolak, S. Kalantre, T. McJunkin, B. Weber and J. Taylor. "Ray-based Classification Framework for High-Dimensional Data." Third Workshop on Machine Learning and the Physical Sciences, NeurIPS 2020, Online, December 11, 2020.

NIST News Releases

The following news items released by the NIST Public Affairs Office describe work in which ACMD staff members have participated.

1. Thermal MagIC: New NIST Project to Build Nano-Thermometers Could Revolutionize Temperature Imaging. October 9, 2020.

2. [NIST Innovation Could Improve Detection of COVID-19 Infections](). October 5, 2020.

3. [Error-Prone Quantum Bits Could Correct Themselves, NIST Physicists Show](). December 8, 2020.

4. [NIST Team Directs and Measures Quantum Drum Duet](). May 6, 2021.

5. [NIST-led Study Finds Variations in Quantitative MRI Scanner Measurements](). June 30, 2021.

Web Services

Note: ACMD provides a variety of information and services on its website. Below is a list of major services provided that are currently under active maintenance.

1. [*Digital Library of Mathematical Functions*]*:* a repository of information on the special functions of applied mathematics.

2. [*Digital Repository of Mathematical Formulae*]: a repository of information on special function and orthogonal polynomial formulae.

3. [*DLMF Standard Reference Tables on Demand*]: an online software testing service providing tables of values for special functions, with guaranteed accuracy to high precision.

4. [*Error-correction Zoo*]: a repository of information about classical and quantum error-correcting codes.

5. [*muMAG*]: a collection of micromagnetic reference problems and submitted solutions.

Software Released

Note: ACMD distributes a large number of software packages that have been developed in the course of its work. Listed below are particular packages which have seen new releases during the reporting period.

1. ACTS*:* Advanced Combinatorial Testing System. Version 3.2. R. Kacker. URL: [https://csrc.nist.gov/projects/automated-combinatorial-testing-for-software/downloadable-tools]

2. FreeVR: Open-source VR Integration Library. Versions 0.6g (04/28/21) and 0.7c (11/01/21). W. Sherman. URL: [http://freevr.org/]

3. Itcl: C++ Inspired Object Oriented Commands for Tcl. Version 4.2.1 (1/2/2021), D. G. Porter. URL: [https://www.tcl.tk/]

4. [LaTeXML]: A LaTeX to XML, HTML, MathML+ Converter. Release 0.8.6 (9/30/2021). B. R. Miller. URL: [https://dlmf.nist.gov/LaTeXML/]

5. MPI_GEVP: Software for Solving Large-scale Generalized Eigenvalue Problems on Distributed Computers. Version 1 (9/9/2020). J. S. Sims DOI: [10.18434/mds2-2293]

6. OOMMF: The Object Oriented MicroMagnetic Framework. Version 2.0a3 (9/30/21). M. J. Donahue and D. G. Porter. URL: [https://math.nist.gov/oommf/software-20.html]

7. Scikit-shape: Python Package for Shape and Image Analysis, Version 0.2 (9/13/21). G. Doğan. URL: [http://scikit-shape.org]

8. Tcl/Tk: Extensible scripting language and GUI toolkit. Versions 8.6.11 (1/2/2021), 8.7a5 (6/18/2021), and 9.0a3 (6/23/2021). D. G. Porter. URL: [https://sourceforge.net/projects/tcl/files/Tcl/]

9. TDBC: Database Connection Commands for Tcl. Version 1.1.2 (1/2/2021). D. G. Porter. URL: [https://sourceforge.net/projects/tcl/files/Tcl/]

10. Thread: Thread Management Commands for Tcl. Version 2.8.6 (1/2/2021). D. G. Porter. URL: [https://sourceforge.net/projects/tcl/files/Tcl/]

Data Released

1. S. Ressler. Photogrammetry test: Stuffed Animal. URL: [https://sketchfab.com/3d-models/curious-george-vintage-1970s-stuffed-animal-47fae6ca9ff64f1b823682799da54b82]

2. Justyna Zwolak. Dark Solitons in BECs Dataset (2022). DOI: [10.18434/mds2-2363]

Conferences, Minisymposia, Lecture Series, Courses

ACMD Seminar Series

Stephen Langer served as Chair of the ACMD Seminar Series. There were 19 talks presented during this period; talks are listed chronologically.

1. Barry Schneider (ACMD). "Calculating Molecular Integrals in Polyatomic Molecules: Developing a 3D Adaptive Grid Based on Atomic Weighting Functions." February 23, 2021.

2. Stephen Schmidt (Humboldt University, Berlin). "Higher Order and Non-Smooth Optimization of Shells with Applications in 3D Imaging and Micro Fluids." March 2, 2021.

3. Tyrus Berry (George Mason University). "Model-free Forecasting with Applications to Multi-sensor Arrays." March 16, 2021.

4. Robert DeJaco (ACMD). "Resolving the Shock layer in Fixed-bed Adsorption with Boundary Layer Theory." March 30, 2021.

5. Anca Radulescu (SUNY New Paltz). "Architecture-Dependent Bifurcations and Clustering in Brain Networks." April 13, 2021.

6. Adarsh Kumhari (University of Sydney). "Regulation of the PD-L1 Inside the Tumor Microenvironment." April 27, 2021.

7. Danielle Brager (ACMD). Mathematically Investigating Retinitis Pigmentosa." May 11, 2021.

8. Adaku Uchendu (Pennsylvania State University). "Reverse Turing Tests for Distinguishing AI-Generated Texts from Human-Written Texts." May 25, 2021.

9. Matthew Daniels (NIST PML). "Efficient Neural Network Training: Streaming Hyperspherical Rotation Circuits for Neural Network Gradient Compression." June 15, 2021.

10. Howard Cohl (ACMD). "The Utility of Series and Integral Representations for the Askey-Wilson Polynomials and their Symmetric Sub-families." June 29, 2021.

11. André Grenier-Petter (University of Wuppertal). "Semantic Preserving Translations between NIST's Digital Library of Mathematical Functions and Computer Algebra Systems." July 20, 2021.

12. Robert Maier (University of Arizona). "Legendre Functions: New Formulas, Identities, and Transformations." August 3, 2021.

13. Josef Sifuentes (University of Texas Rio Grande Valley). "GMRES Convergence and Spectral Properties of Approximate Preconditioners for KKT Matrices." August 17, 2021.

14. Bob Sorenson (Hyperion Research). "Highlights from the Advanced Computing World: Exascale Rolls Out and Quantum Computing Developments." August 31, 2021.

15. Rommie Amaro (UCSD) and Gregory Voth (University of Chicago). "Computational Molecular Biology and SARS-CoV-2." September 14, 2021.

16. Pulkit Grover (Carnegie Mellon University). "Information-theoretic Techniques for Examining and Affecting Biological and Artificial Computing Systems for Human-centered Goals." October 5, 2021.

17. Harbir Antil (GMU). "Optimization Based Deep Learning Neural Networks with Applications." November 9, 2021.

18. Paul Patrone (ACMD). "Optimal Decision Theory for SARS-CoV-2 Antibody testing." November 16, 2021.

19. Andrew Glaws (National Renewable Energy Laboratory). "Invertible Neural networks for Wind Turbine Airfoil and Blade Design." November 30, 2021.

Shortcourses

1. <u>V. Albert</u> and M. Gullans. "Advanced Topics in Theory of Computing: Quantum Error Correction and Fault-Tolerance." University of Maryland, College Park, MD, Fall 2021.

2. V. Albert. "Many-Body Physics in a Nutshell." Advanced Topics in Electrophysics: Quantum Technology, Graduate course, University of Maryland, College Park, MD, April 21, May 3, and May 5, 2021.

Conference Organization

Leadership

1. T. Gerrits. Lead Co-organizer, Quantum Economic Development Consortium (QED-C) Workshop on Single-Photon Measurement Infrastructure for Quantum Applications, Online, September 14-21, 2021.

2. P. S. Kuo. Subcommittee Chair, FS 3: Quantum Photonics, 2022 Conference on Lasers and Electro-optics, San Jose, CA, May 15-20, 2022.

3. L. Ma. Lead Organizer, IEEE International Workshop on Quantum Communication and Quantum Cryptography, IEEE Conference on Communications and Network Security (CNS) 2021, Online, October 6, 2021.

4. K. Sayrafian. Lead Organizer and Co-Chair, 4th International IoT-Health Workshop, IEEE International Conference on Communications. Online, June 14-23, 2021.

5. K. Sayrafian. Co-Chair. Vertical Applications and Internet of Things Track, IEEE European Conference on Networks and Communications (EuCNC 2021). Online, June 8-11, 2021.

Committee Membership

1. M. Coudron. Member, Program Committee, 16th Conference on the Theory of Quantum Computing, Communication, and Cryptography (TQC 2021), Online, July 5-8, 2021.

2. J. Fong. Member, Organizing Committee, Symposium on Particle and Finite Element Models for

Interaction, Simulation and Statistical Design, 15th World Congress on Computational Mechanics (WCCM-XV), Yokohama, Japan, July 31 – August 5, 2022.

3. T. Gerrits. Member, Program Committee, Quantum Information & Communications, CLEO 2021, San Jose, CA, May 9-14, 2021.

4. T. Gerrits. Member, Program Committee, IEEE International Workshop on Quantum Communication and Quantum Cryptography, IEEE Conference on Communications and Network Security (CNS) 2021, Online, October 6, 2021.

5. S. Glancy. Member, Program Committee, Southwest Quantum Information Technology Workshop, Online, October 14-15, 2021.

6. R. Kacker. Member Steering Committee and Member Program Committee, 10th International Workshop on Combinatorial Testing (IWCT 2021), 14th IEEE International Conference on Software Testing, Verification, and Validation (ICST 2021), Online, April 12-16, 2021.

7. P. Kuo. Subcommittee Member, FS 3: Quantum Photonics, 2021 Conference on Lasers and Electro-optics, Online, May 9-14, 2021.

8. P. Kuo. Member, Program Committee, Quantum Computing, Communication, and Simulation, 2021 SPIE Photonics West, Online, March 6-11, 2021.

9. R. La. Member, Technical Program Committee, IEEE International Conference on Computer Communications (INFOCOM 21), Online, May 10-13, 2021.

10. R. La. Member, Technical Program Committee, IEEE International Conference on Computer Communications (INFOCOM 22). Online, May 2-5, 2022.

11. R. La. Member, Technical Program Committee, International Conference on Complex Networks and Their Applications (Complex Networks 21). Hybrid (Madrid, Spain), November 30 – December 2, 2021.

12. Y. Liu. Member, Program Committee, Third NIST PQC Standardization Conference. Online, June 7-9, 2021.

13. K. Sayrafian. Member, Program Committee, 15th EAI International Conference on Body Area Networks (BODYNETS 2020). Online, October 21-22, 2020.

14. K. Sayrafian. Member, International Advisory Board, International Symposium on Medical Information and Communication Technology.

15. K. Sayrafian. Member, Organizing Committee, Vertical Track 1: Health and Well-Being, COST CA20120: Intelligence-Enabling Radio Communications for Seamless Inclusive Interactions, 2021.

16. W. Sherman. Member, Program Committee, 2021 IEEE Conference on Virtual Reality, Online, March 27 – April 2, 2021.

17. W. Sherman. Member, Executive Board, 2021 Campus Alliance for Advance Visualization (CAAV) Annual Conference, Purdue University, November 1-4, 2021.

18. O. Slattery. Member, Program Committee, Quantum Communications and Imaging XI, SPIE Optics and Photonics, San Diego, CA, Online, August 2021.

19. S. Su. Member, Program Committee, Conference Paper Track. IEEE Virtual Reality 2022, Online, March 12-16, 2022.

20. J. Zwolak. NIST Booth Moderator, Grace Hopper Celebration of Women in Computing, Online, September 27 – October 1, 2021.

Session Organization

1. H. Cohl. Co-Organizer, Special Session on the Legacy of Dick Askey, 2021 Joint Mathematics Meetings, Online, January 6-9, 2021.

2. G. Doğan. Co-Organizer, Minisymposium 21: Advances in Shape Optimization Algorithms. SIAM Conference on Optimization, Online, July 20-23, 2021.

3. A. Kearsley. Organizer, Minisymposium: Modelling and Methods in Mathematical Biology. SIAM Annual Meeting, Online, July 19-23, 2021.

4. A. Kearsley. Organizer, Minisymposium: Modelling and Methods in Mathematical Biology, Society for Mathematical Biology Annual Meeting, Online, June 13-17, 2021.

5. K. Sayrafian. Co-Organizer and Co-Chair, Measurement & Modelling of Radio Waves Propagation in Body Area Networks, URSI General Assembly and Scientific Symposium, Online, Sept. 4, 2021.

6. K. Sayrafian. Co-Chair, VAP4: IoT for Industrial and Business Applications, IEEE European Conference on Networks and Communications (EuCNC 2021), Online, June 10, 2021.

Other Professional Activities

Internal

1. R. Boisvert. Coordinator, ITL Quantum Information Program.

2. R. Boisvert. Member, NIST Computing Research Advisory Committee.

3. B. Cloteaux. Member, Washington Editorial Review Board.

4. T. Gerrits. Organizer, Quantum Optics Meeting Series.

5. S. Glancy. Member, Boulder Summer Undergraduate Research Fellowship Committee.

6. S. Glancy. Member, ITL Diversity Committee.

7. P. Kuo. Member, ITL Space Task Force.

8. P. Kuo. Member, NIST Quantum Networking Grand Challenge Small Team (advisory board) .

9. L. Ma. Organizer, Quantum Repeater Journal Club.

10. L. Orr. Member, ITL Diversity Committee.

11. D. Porter. Member, Editorial Board, Journal of Research of NIST.

12. D. Porter. Member, ITL Awards Committee

13. B. Schneider. Chair, HPC Task Force, NIST Research Computing Advisory Committee.

14. O. Slattery. Laser Safety Representative, ITL Safety Committee.

15. O. Slattery. Member, ITL Space Task Force.

16. O. Slattery. ITL Representative. NIST Laser Safety Committee.

17. J. Zwolak. At-Large Member for ITL, NIST AI Community of Interest Steering Committee.

External

Editorial

1. R. Boisvert. Associate Editor, *ACM Transactions on Mathematical Software*.

2. H. Cohl. Member, Editorial Board, *The Ramanujan Journal*.

3. H. Cohl. Co-Editor, OP-SF NET Newsletter, SIAM Activity Group on Orthogonal Polynomials and Special Functions.

4. H. Cohl. Co-Editor, Liber Amicorum, a Friendship Book for Dick Askey.

5. H. Cohl. Co-Editor, Richard A. Askey Memorial Volume, *The Ramanujan Journal*.

6. T. Gerrits. Associate Editor, *Optics Express*.

7. Z. Gimbutas. Member, Editorial Board, *Advances in Computational Mathematics*.

8. S. Glancy. Associate Editor, *Quantum Information Processing*.

9. R. La. Associate Editor, *IEEE/ACM Transactions on Networking*.

10. F. Potra. Regional Editor for the Americas, *Optimization Methods and Software*.

11. F. Potra. Associate Editor, *Journal of Optimization Theory and Applications*.

12. F. Potra. Associate Editor, *Numerical Functional Analysis and Optimization*.

13. F. Potra. Associate Editor, *Optimization and Engineering*.

14. B. Saunders. Associate Editor, *MAA Mathematics Magazine*.

15. B. Saunders. Webmaster, SIAM Activity Group on Orthogonal Polynomials and Special Functions.

16. B. Saunders. Moderator, OP-SF Talk, SIAM Activity Group on Orthogonal Polynomials and Special Functions.

17. K. Sayrafian. Associate Editor, *International Journal of Wireless Information Networks*.

18. B. Schneider. Associate Editor in Chief, *IEEE Computing in Science and Engineering*.

19. B. Schneider. Specialist Editor, *Computer Physics Communications*.

Boards and Committees

1. R. Boisvert. Member, International Federation of Information Processing Working Group 2.5 (Numerical Software).

2. R. Boisvert. Member, ACM Presidential Task Force on Digital Library Metadata, Metrics, and Reporting Data.

3. R. Boisvert. Member, Reproducibility Badging and Definitions Working Group, National Information Standards Organization (NISO).

4. B. Cloteaux, Member, Advisory Board, Department of Computer Science, New Mexico State University.

5. A. Dienstfrey. Vice-Chair, International Federation of Information Processing Working Group 2.5 (Numerical Software).

6. J. Fong. Member, American Society of Mechanical Engineers (ASME) Boiler and Pressure Vessel Code Committee.

7. T. Gerrits. Member, Scientific Advisory Board, European (NMI-driven) SIQUST (Single-photon Sources as New Quantum Standards) Project.

8. S. Glancy. Member, IEEE Working Group on Metrics and Benchmarks for Quantum Computing Devices and Systems.

9. S. Glancy. Member, Standards Technical Advisory Committee, Quantum Economic Development Consortium.

10. P. Kuo. Member, Executive Committee. American Physical Society Mid-Atlantic Section.

11. D. Porter. Member, Tcl Core Team.

12. S. Ressler. Member, World Wide Web Consortium (W3C) Immersive Web Working Group.

13. S. Ressler. NIST Representative, W3C Advisory Committee.

14. S. Ressler. NIST Representative, Khronos Group.

15. S. Ressler. Member, Khronos Group 3D Formats Working Group.

16. B. Saunders. Member, Board of Trustees, Society for Industrial and Applied Mathematics (SIAM).

17. B. Saunders. Member, SIAM Systems Oversight Committee.

18. B. Saunders. MD-DC-VA Section Representative, Mathematical Association of America (MAA) Congress.

19. B. Saunders. Member, AWM-MAA Falconer Lecture Nominating Committee.

20. K. Sayrafian. Co-Chair. IoT-Health Sub-Group, COST CA15104: Inclusive Radio Communication Networks for 5G and Beyond.

21. W. Sherman. Member, Khronos Group OpenXR Working Group.

22. W. Sherman. Member, Khronos Group, ANARI Working Group.

23. B. Schneider. Co-Chair, Fast Track Action Committee on Strategic Computing, National Science and Technology Council.

24. S. Su. Member, Khronos Group 3D Formats Working Group.

25. S. Su. Member, Khronos Group OpenXR Working Group.

26. S. Su. Member, Khronos Group, ANARI Working Group.

Adjunct Academic Appointments

1. V. Albert. Adjunct Assistant Professor, Department of Physics, University of Maryland, College Park, MD.

2. M. Coudron. Adjunct Assistant Professor, Department of Computer Science, University of Maryland, College Park, MD.

3. S. Glancy. Lecturer, Department of Physics, Colorado University, Boulder, CO.

4. E. Knill. Lecturer, Department of Physics, Colorado University, Boulder, CO.

5. P. Kuo. Adjunct Associate Professor, Department of Physics, University of Maryland, College Park, MD.

6. Y. Liu. Co-Director, Joint Center for Quantum Information and Computer Science (QuICS), and Adjunct Associate Professor, Department of Computer Science, University of Maryland, College Park, MD.

7. K. Sayrafian. Affiliate Associate Professor, Electrical and Computer Engineering Department, Concordia University, Montreal, Canada.

8. S. Su. Adjunct Professor, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County.

9. X. Tang. Adjunct Professor, Department of Physics, University of Limerick, Ireland.

Thesis Direction

1. V. Albert. Advisor, University of Maryland: S. Jain.

2. V. Albert. Co-Advisor, University of Maryland: J. Kunjummen.

3. V. Albert. Co-Advisor, University of Maryland: S. Nezhadi.

4. V. Albert. Co-Advisor, University of Maryland: J. Iosue.

5. M. Donahue. Member, Ph.D. Thesis Committee, University of Colorado, Boulder: Mingyu Hu.

6. Y.-K. Liu. Member, Ph.D. Thesis Committee, University of Maryland: E. Schoute.

7. Y.-K. Liu. Co-Advisor, University of Maryland: K. Huang.

8. B. Miller, Member, Ph.D. Thesis Committee, Deyan Ginev, University of Erlangen, Germany: D. Ginev.

9. F. Potra. Chair, Ph.D. Thesis Committee, UMBC: J. Praniewicz.

10. K. Sayrafian. Co-Advisor, University of Zagreb, Zagreb, Croatia: K. Krhac.

Community Outreach

1. D. Brager and A. Kearsley. Participants, ACM Richard Tapia Conference.

2. M. Coudron. Project Leader, Combinatorics, Algorithms, and AI for Real Problems Research Experience for Undergraduates, University of Maryland, College Park, Summer 2021.

3. G. Doğan and J. Zwolak. "Mathematical and Computational Science Research at NIST," Center for Mathematics and Artificial Intelligence Industry Day, George Mason University, Online, June 6, 2021.

4. B. Saunders. Advisor, Virginia Standards of Learning Tutoring Program, Northern Virginia Chapter (NoVAC), Delta Sigma Theta Sorority, Inc. and the Dunbar Alexandria-Olympic Branch of the Boys and Girls Clubs of Greater Washington.

5. B. Saunders. SIAM Visiting Lecturer.

6. J. Zwolak. NIST Booth Moderator, Grace Hopper Celebration of Women in Computing, Online, September 27 – October 1, 2021.

Awards and Recognition

External

1. L. Brady. USRA Q2B Applied NISQ Computing Paper Award, 2021.

2. S. Glancy. Outstanding Referee. American Physical Society, December 2020.

3. S. Guo, A. Fritsch, C. Greenberg, I. Spielman and J. Zwolak. Best Poster Award, Harnessing Quantum Matter Data Revolution Virtual Symposium and Summer School, June 2021.

4. A. Kearsley. Excellence in Research in Mathematics and Computer Science Award, Washington Academy of Sciences, September 2021.

5. R. La. University of Maryland Institute for Systems Research Outstanding Faculty Award, September 2021.

6. S. Ressler. 100 Original Voices in XR, Avi Bar-Zeev Blog[31], September 2021.

7. B. Saunders Society for Industrial and Applied Mathematics, 2021 Black History Month Honoree.[32]

8. B. Saunders. 2021 AWM-MAA Etta Z. Falconer Lecturer[33], June 2021.

9. K. Sayrafian. Certificate of Appreciation. IEEE Eta Kappa Nu, December 2020.

Internal

1. R. Evans and A. Kearsley. DOC Silver Medal (Group Award) , January 2022.

2. T. Gerrits. Dept of Commerce Silver Medal (Group Award), January 2022.

3. W. George and S. Satterfield. NIST Judson C. French Award (Group Award), January 2021.

4. S. Glancy and E. Knill. Department of Commerce Gold medal (Group Award), January 2021.

5. D. Lozier. NIST Portrait Gallery of Distinguished Scientists, Engineers and Administrators, October 2020.

6. B. Saunders, NIST Diversity, Inclusion and EEO Award, January 2021.

7. B. Schneider. Department of Commerce Bronze Medal (Group Award), January 2021.

Funding Received

During FY 2021 ACMD's yearly allocation of base funding from the NIST Information Technology Laboratory was supplemented with funding from a variety of internal and external competitions. Such funding represented 16 % of the Division's FY 2021 budget.

Note: For multi-year awards and joint awards, only projects with new funding received by ACMD during FY 2021 are listed. For joint projects, names of ACMD participants are underlined.

External

1. M. Donahue. NIST Support for DARPA Magnetic Miniaturized and Monolithically Integrated Component (M3IC) Program, DARPA. (Joint with PML)

2. B. Schneider. XSEDE Computing Allocation for AMP Gateway: A Gateway for Atomic and Molecular Physics.

---

[31] https://avibarzeev.medium.com/100-original-voices-in-xr-55b5a609f4ae
[32] https://sinews.siam.org/Details-Page/honoring-bonita-v-saunders-1

[33] https://awm-math.org/wp-content/uploads/2021/06/PR-Falconer-2021-Saunders.pdf

Internal

1. T. Blattner, W. Keyrouz, J. H Scott, <u>M. Mascagni</u>, D. Juba, P. Bajcsy, and J. Villarrubia. Ray Tracing for Accelerated Metrology. ITL Building the Future Program.

2. C. Dennis, T. Moffat, A. Biacchi, A. Hight Walker, S. Woods, W. Tew and <u>M. Donahue</u>. Thermal MagIC: An SI-Traceable Method for 3D Thermal Magnetic Imaging and Control. NIST Innovations in Measurement Science.

3. A. Dienstfrey. Advanced Hardware for AI Advantage. ITL AI Initiative. (Joint with PML)

4. <u>R. Evans</u>, A. Balijepalli, <u>C. Schanzle</u>, and <u>A. Kearsley</u>. Modeling, Uncertainty Quantification, and Calibration for BioFETs. ITL Building the Future Program.

5. R. Fitzgerald, D. Bergeron, S. Nour, D. Schmidt, D. Swetz, G. Shaw, <u>B. Alpert</u>, and M. Verkouteren. True Becquerel: A New Paradigm for 21st Century Radioactivity Measurements. NIST Innovations in Measurement Science.

6. <u>T. Gerrits</u>, <u>O. Slattery</u>, <u>L. Ma</u>, <u>X. Tang</u>, <u>A. Rahmouni</u>, and <u>S. Bhushan</u>. A Robust, Shippable Reference Single-Photon Source. ITL Building the Future program.

7. C. Greenberg, H.-K. Liu, <u>J. Zwolak</u>, P. Fontana, Y Lee, and J. Zhang. Generative Models and Simulation for AI Measurement and Evaluation. ITL Building the Future Program.

8. D. Juba, W. Keyrouz, and <u>M. Mascagni</u>. Accelerated Floating-Point Uncertainty Quantification. ITL Building the Future Program.

9. <u>A. Kearsley</u>, T. Allison, W. Keyrouz, and <u>B. Schneider</u>. Machine Learning (ML) for Predicting Normal Boiling Point. ITL Building the Future Program.

10. A. Kearsley. Computational Classification of Forensic Evidence. NIST Special Programs Office.

11. K. Keenan, S. Russek, K. Stupic, <u>A. Dienstfrey</u> and A. Peskin. AI for Low-Field Magnetic Resonance Imaging. ITL Use-inspired AI Program.

12. <u>E. Knill</u> and <u>S. Glancy</u>. Practical Statistical Intervals with Proven Coverage for Quantum Measurements, ITL Building the Future Program.

13. <u>P. Kuo</u> and <u>J. Zwolak</u>. Machine Learning for Designing Entangled Photon-Pair Sources. ITL Building the Future Program.

14. D. Leibfried, A. Wilson, <u>E. Knill</u>, and <u>S. Glancy</u>. A Practical Quantum Repeater Unit. NIST Quantum Networking Grand Challenge.

15. K. Lehnert, K. Silverman, D. Moody, J. Teufel, R. Mirin, S.-W. Nam, <u>E. Knill</u>, P. Hale, and T. Dennis. Establishing the Science of Networks for Superconducting Quantum Computers. NIST Innovations in Measurement Science Program.

16. <u>V. Marbukh</u>, <u>B. Cloteaux</u> and <u>D. Mitra</u>. Towards Actionable Cybersecurity Risk Metrics in a Hyperconnected World. ITL Building the Future Program.

17. <u>P. Patrone</u>, <u>A. Kearsley</u>, <u>G. McFadden</u>, G. Cooksey, S. Sarkar, and L. Wang. NIST-in-a-drop: Revolutionizing Measurements of Single-cell Kinetics. NIST Innovations in Measurement Science.

18. <u>P. Patrone</u>, <u>A. Kearsley</u>, <u>R. DeJaco</u>, <u>C. Schanzle</u>, E. Romsos, and P. Vallone. Mathematics of Metrology for Quantitative Polymerase Chain-Reaction Measurements. ITL Building the Future Program.

19. K. Sayrafian. A Wireless Wearable Technology to Detect Accumulation of Fluid in the Lungs. ITL Building the Future Program.

20. <u>O. Slattery</u>, A. Battou and N. Zimmerman. Quantum Network Testbeds on NIST Campus and on DC-QNet. NIST Quantum Networking Grand Challenge.

21. <u>J. Zwolak</u>, C. Greenberg, and I. Spielman. Machine Learning System for Information Extraction from Experimental BEC Images. ITL Building the Future Program.

Grants Funded

ACMD awards a small amount of funding through the NIST Grants Program for projects that make direct contributions to its research programs. During FY 2021 the following cooperative agreements were active.

1. Clarkson University: *Restoration of Orientation Data from Polycrystalline Materials*. PI: Prashant Athavale.

2. Concordia University: *Application of Machine Learning in Kinetic-Based Micro Energy Harvesting for Wearable IT-Health Devices*. PI: Amir Aghdam.

3. Prometheus Computing LLC: *Security, Resiliency and Dynamics of Interdependent Self-Organizing Networks*. PI: Assane Gueye.

4. Theiss Research: *Exploiting Alternate Computing Technologies*. PI: Alan Mink.

5. University of Edinburgh: *Rigorous and Presentable Asymptotics for Special Functions and Orthogonal Polynomials*. PI: Adri Olde-Daalhuis.

6. University of Maryland: *In-situ Visualization for Immersive Environments*. PI: Amitabh Varshney.

7. University of Maryland: *Joint Center for Quantum Information and Computer Science (QuICS)*. PI: Andrew Childs.

8. University of Texas at Arlington: *SENTINEL: Security Interaction Testing for IoT Systems and Blockchains*. PI: Yu Lei.

External Contacts

ACMD staff members interact with a wide variety of organizations in the course of their work. Examples of these follow.

Industrial Labs

Amazon AWS
Boeing Company
Chakra Consulting
Computational Physics, Inc.
Deltares (Netherlands)
Duke Energy
Fraunhoffer IGD (Germany)
GE Global Research Center
Google
HC-Photonics Corp.
Hyperfine
Honeywell
Intel Corporation
Intelligent Automation Inc.
Interdigital
Kitware
Microsoft Research
MPACT Corp.
NTT Corporation (Japan)
PhaseCraft (UK)
Photon Spot Inc.
Prometheus Computing LLC
Prorenata Labs
PsiQuantum
Rolls-Royce Corporation
Xanadu Quantum Technologies, Inc.
XYZ Scientific

Government/Non-profit Organizations

Advanced Technology Academic Research Center
American Society of Mechanical Engineers
American Telemedicine Association
Argonne National Laboratory
Army Research Laboratory
Association for Computing Machinery
Center for Disease Control
Consejo Nacional de Investigaciones
    Científicas y Técnicas (COCINET, Argentina)

Defense Advanced Research Projects Agency
El Centro Nacional de Metrología (CENAM, Mexico)
Federal Communications Commission
Food and Drug Administration
IEEE Computer Society
ETH Zurich
International Federation for Information Processing
Johns Hopkins University Applied Physics Laboratory
Kavli Institute for Theoretical Physics
Laboratory for Telecommunication Sciences
Lawrence Berkeley National Laboratory
Lawrence Livermore National Laboratory
Mathematical Sciences Research Institute
Nanohub.org
National Cancer Institute
National Aeronautics and Space Administration
National Institute of Biomedical Imaging and Bioengineering
National Institutes of Health
National Physical Laboratory (UK)
National Research Council (Canada)
National Security Agency
National Science Foundation
National Renewable Energy Laboratory (NREL)
Naval Research Laboratory
NTT Corporation
Oak Ridge National Laboratory
OpenMath Society
Pacific Northwest National Laboratory
Physikalisch-Technische Bundesanstalt (Germany)
Quantum Economic Development Consortium
Sandia National Laboratories
SBA-Research, Austria
Society for Industrial and Applied Mathematics
Stanford Research International (SRI)
Theiss Research
US Department of Energy
US Geological Survey
US Holocaust Memorial Museum
US Office of Naval Research (USNO)
World Wide Web Consortium

Universities

Aarhus University (Denmark)
Alfréd Rényi Institute of Mathematics (Hungary)
American University
Beijing Institute of Technology (China)
Boğaziçi University (Turkey)
Brown University
Caltech
Carnegie Mellon University
China University of Petroleum (China)
Clarkson University
Clemson University
Colorado School of Mines
Colorado State University

Columbia University
Comenius University (Slovakia)
Concordia University (Canada)
Coppin State University
Courant Institute of Mathematical Sciences
Curtin University (Australia)
Dalian University of Science and Tech (China)
Delft University of Technology (The Netherlands)
Drake University
Drexel University
Duke University
East China U. of Science and Tech (China)
European University Cyprus (Cyprus)
Federal University of Ceará (Brazil)
Flatiron Institute
Florida State University
Freie Universität Berlin
Gdansk University of Technology (Poland)
George Mason University
George Washington University
Harvard University
Haverford College
Humboldt University (Germany)
Imperial College London (UK)
Indiana University
Indian Institute of Technology (India)
Institute of Mathematics and its Applications
Jacobs University Bremen (Germany)
Johannes Kepler University (Austria)
Johns Hopkins University
Khallikote College (India)
Kennesaw State University
Louisiana State University
Ludwig-Maximilians University (Germany)
Massachusetts Institute of Technology (MIT)
Morgan State University
Nagoya University
National University of Singapore (Singapore)
New York University
Norfolk State University
Occidental College
Ohio State University
Old Dominion University
Open University (UK)
Osaka University
Paderborn University (Germany)
Polytechnic University of Valencia (Spain)
Rollins College
Rose-Hulman Institute of Technology
Shandong University (China)
ShanghaiTech University (China)
Shippensburg University
Southern University and A&M College
Stanford University
SUNY Binghamton
Technical University of Denmark (Denmark)
Technical University of Munich (Germany)

Technische Universität Berlin (Germany)
Texas A&M University
Texas Tech University
Tokyo University of Science (Japan)
Tsinghua University (China)
Tufts University
University of Akron
University of Alcala
University of Amsterdam
University of Antwerp (Belgium)
Universidad Autonoma de Madrid (Spain)
University of Bergamo (Italy)
University of Bologna (Italy)
University of British Colombia (Canada)
University of California, Berkeley
University of California, Los Angeles
University of California, Santa Barbara
University of California, San Diego
University of Central Florida
University of Colorado, Boulder
University of Delaware
University of Edinburgh (UK)
University of Erlangen (Germany)
University of Freiburg (Germany)
University of Illinois, Urbana-Champaign
University of Limerick (Ireland)
University of Lisbon (Portugal)
University of Manchester (UK)
University of Maryland, Baltimore County
University of Maryland, College Park
University of Massachusetts
University of Michigan
University of New Mexico
University of Oregon
University of Oulu (Finland)
University of Oxford (UK)
University of Rochester
University of Saskatchewan (Canada)
University of Scranton
University of Sheffield (UK)
University of Southern California
University of South Carolina
University of Strathclyde (UK)
University of Texas at Arlington
University of Texas at Rio Grande Valley
University of Toronto (Canada)
University of Washington
University of Wisconsin, Madison
University of Wisconsin, Milwaukee
University of Wuppertal (Germany)
University of Wyoming
University of York (UK)
University of Zagreb (Croatia)
Virginia Polytechnic Institute
Western Michigan University
Worcester Polytechnic University
Yale University

Part V

Appendix

Staff

ACMD consists of full-time permanent Federal staff located at NIST laboratories in Gaithersburg, MD and Boulder, CO. This full-time staff is supplemented with a variety of special appointments. The following list reflects all non-student appointments held during any portion of the reporting period (October 2020 – December 2021). Students and interns are listed in Table 7 page 139.

* Denotes staff at NIST Boulder.
† Denotes part-time Federal staff.

Division Staff

Ronald Boisvert, *Chief*, Ph.D. (Computer Science), Purdue University, 1979
Catherine Graham, *Administrative Assistant*
Elsie (Meliza) Lane, *Administrative Assistant*
Lochi Orr, *Administrative Assistant*, A.A. (Criminal Justice), Grantham University, 2009
†Alfred Carasso, Ph.D. (Mathematics), University of Wisconsin, 1968
Roldan Pozo, Ph.D. (Computer Science), University of Colorado at Boulder, 1991
Kamran Sayrafian-Pour, Ph.D. (Electrical and Computer Engineering), University of Maryland, 1999
Christopher Schanzle, B.S. (Computer Science), University of Maryland Baltimore County, 1989

Mathematical Analysis and Modeling Group

Timothy Burns, *Leader*, Ph.D. (Mathematics), University of New Mexico, 1977
*Daniel Flynn, *Administrative Assistant*, B.S. (Political Science), Iowa State University, 2016
*Bradley Alpert, Ph.D. (Computer Science), Yale University, 1990
*Andrew Dienstfrey, Ph.D. (Mathematics), New York University, 1998
Ryan Evans, Ph.D. (Applied Mathematics), University of Delaware, 2016
†Jeffrey Fong, Ph. D. (Applied Mechanics and Mathematics), Stanford University, 1966
*Zydrunas Gimbutas, Ph.D. (Applied Mathematics), Yale University, 1999
*Zachary Grey, Ph.D. (Computational and Applied Mathematics), Colorado School of Mines, 2019
Raghu Kacker, Ph.D. (Statistics), Iowa State University, 1979
Anthony Kearsley, Ph.D. (Computational and Applied Mathematics), Rice University, 1996
Geoffrey McFadden, *NIST Fellow,* Ph.D. (Mathematics), New York University, 1979
Danielle Middlebrooks, Ph.D. (Applied Mathematics, Statistics and Scientific Computing), University of Maryland, 2020
Paul Patrone, Ph.D. (Physics), University of Maryland, 2013

Faculty Appointee (Name, Degree / Home Institution)
    Daniel Anderson, Ph.D. / George Mason University
    Michael Mascagni, Ph.D. / Florida State University
    John Nolan, Ph.D. / American University
    Florian Potra, Ph.D. / University of Maryland Baltimore County

NRC Postdoctoral Associates
    Danielle Brager, Ph.D. (Mathematical Biology), Arizona State University, 2020
    Robert DeJaco, Ph.D. (Chemical Engineering), University of Minnesota, 2020
    Matthew Roberts, Ph.D. (Mathematical Sciences), Michigan Technological University, 2019

Guest Researchers (Name, Degree / Home Institution)
    Sebastian Barillaro, Ph.D. / Industrial Technology National Institute, Argentina
    Prajakta Bedekar, Ph.D. / Johns Hopkins University
    Natesh Ganesh, Ph.D. / University of Colorado
    Fern Hunt, Ph.D. / *NIST Scientist Emeritus*

Yu (Jeff) Lei, Ph.D. / University of Texas at Arlington
Rayanne Luke, Ph.D. / Johns Hopkins University
Dimitrios Simos, Ph.D. / SBA Research, Austria

Mathematical Software Group

Michael Donahue, *Leader* Ph.D. (Mathematics), Ohio State University, 1991
Javier Bernal, Ph.D. (Mathematics), Catholic University, 1980
Howard Cohl, Ph.D. (Mathematics), University of Auckland, 2010
Günay Doğan, Ph.D. (Applied Mathematics and Scientific Computing), University of Maryland, 2006
Stephen Langer, Ph.D. (Physics), Cornell University, 1989
Marjorie McClain, M.S. (Mathematics), University of Maryland College Park, 1984
Bruce Miller, Ph.D. (Physics), University of Texas at Austin, 1983
Donald Porter, D.Sc. (Electrical Engineering), Washington University, 1996
Bonita Saunders, Ph.D. (Mathematics), Old Dominion University, 1985
Barry Schneider, Ph.D. (Physics), University of Chicago, 1969

Faculty Appointees (Name, Degree / Home Institution)
Abdou Youssef, Ph.D. / George Washington University

NRC Postdoctoral Associates
Lisa Ritter, Ph.D. (Mathematics), University of Hawai'i at Mānoa, 2021

Guest Researchers (Name, Degree / Home Institution)
Nicolas Douguet, Ph.D. / Kennesaw State University
Heman Gharibnejad, Ph.D. / Computational Physics Inc.
Deyan Ginev, M.Sc. / Chakra Consulting
Mark Alexander Henn, Ph.D. / University of Maryland
Justin Kauffman, Ph.D. / Virginia Polytechnic Institute and State University
Daniel Lozier, Ph.D. / NIST, Retired
Todd Martinez, Ph.D. / Stanford University
Adri Olde Daalhuis, Ph.D. / University of Edinburgh
Jeppe Olsen, Ph.D. / Aarhus University, Denmark
Chen Qu, Ph.D. / University of Maryland
Moritz Schubotz, Ph.D. / University of Karlsruhe, Germany

Computing and Communications Theory Group

Ronald Boisvert, *Acting Leader*, Ph.D. (Computer Science), Purdue University, 1979
Victor Albert, Ph.D. (Physics), Yale University, 2017
Lucas Brady, Ph.D. (Physics), University of California at Santa Barbara, 2018
Brian Cloteaux, Ph.D. (Computer Science), New Mexico State University, 2007
Matthew Coudron, Ph.D. (Computer Science), Massachusetts Institute of Technology, 2017
Thomas Gerrits, Ph.D. (Physics), Radboud University Nijmegen, 2004
*Scott Glancy, Ph.D. (Physics), University of Notre Dame, 2003
*Emanuel Knill, *NIST Fellow,* Ph.D. (Mathematics), University of Colorado at Boulder, 1991
Paulina Kuo, Ph.D. (Physics), Stanford University, 2008
Yi-Kai Liu, Ph.D. (Computer Science), University of California, San Diego, 2007
Lijun Ma, Ph.D. (Precision Instruments and Machinery), Tsinghua University, 2001
Vladimir Marbukh, Ph.D. (Mathematics) Leningrad Polytechnic University, 1986
Oliver Slattery, *Project Leader*, Ph.D. (Physics), University of Limerick, 2015
†Xiao Tang, Ph.D. (Physics), Chinese Academy of Sciences, 1985

Faculty Appointees (Name, Degree / Home Institution)
James Lawrence, Ph.D. / George Mason University

Richard La, Ph.D. / University of Maryland
Debasis Mitra, Ph.D. / Columbia University

Guest Researchers (Name, Degree / Home Institution)
Isabel Beichl, Ph.D. / NIST, Retired
Sumit Bhushan, Ph.D. / Indian Institute of Technology
Sesha Challa / Pusan National University (India)
Assane Gueye, Ph.D. / Prometheus Computing
*Karl Mayer, Ph.D. / Honeywell
Alan Mink, Ph.D. / Theiss Research
Anouar Rahmouni, Ph.D. / Moroccan Foundation for Advanced Sci., Innovation and Research
*Ezad Shojaee, Ph.D. / University of Colorado
Francis Sullivan, Ph.D. / IDA Center for Computing Sciences
*James Van Meter, Ph.D. / HRL Laboratories

High Performance Computing and Visualization Group

Judith Terrill, *Leader*, Ph.D. (Information Technology), George Mason University, 1998
William George, Ph.D. (Computer/Computational Science), Clemson University, 1995
Terence Griffin, B.S. (Mathematics), St. Mary's College of Maryland, 1987
Sandy Ressler, M.F.A. (Visual Arts), Rutgers University, 1980
Steven Satterfield, M.S. (Computer Science), North Carolina State University, 1975
William Sherman, M.S. (Computer Science), University of Illinois, 1989
Simon Su, Ph.D. (Computer Science), University of Houston, 2001
Justyna Zwolak, Ph.D. (Physics), Nicolaus Copernicus University, Poland, 2011

NRC Postdoctoral Associates
Joshua Ziegler, Ph.D. (Physics), University of Oregon, 2020

Guest Researchers (Name, Degree / Home Institution)
John Hagedorn, M.S. / Chakra Consulting
James Sims, Ph.D. / NIST (retired)

Glossary of Acronyms

1D	one-dimensional
2D	two-dimensional
2DEG	2D electron gas
3D	three-dimensional
ACI	American Concrete Institute
ACM	Association for Computing Machinery
ACMD	NIST/ITL Applied and Computational Mathematics Division
ACTS	Advanced Combinatorial Testing System (software)
ADC	apparent diffusion coefficient
ADJOINT	African Diaspora Joint Mathematics Workshop
ADLP	NIST Associate Director for Laboratory Programs
ADS	autonomous driving system
AI	artificial intelligence
AM	analytical methods
AMOS	atomic molecular and optical science
AMOSG	Atomic Molecular and Optical Science Gateway
AMP	atomic and molecular physics
AMS	American Mathematical Society
AN	access network
ANARI	Analytic Rendering Interface for Data Visualization
API	application programming interface
APS	American Physical Society
AR	augmented reality
ARDS	acute respiratory distress syndrome
ASES	Aspects of Student Experience Survey
ASME	American Society of Mechanical Engineers
AWM	Association for Women and Computing
arXiv	preprint archive housed at Cornell University (http://arxiv.org/)
BCI	brain-computer interface
Be	beryllium
BEC	Bose-Einstein Condensate
Bio-FET	biological field effect transistor
BMP	Bateman Manuscript Project
BP	normal boiling point
Bq	becquerel: absolute activity of radionuclide mixtures
Caltech	California Institute of Technology
CAD	computer aided design
CAS	computer algebra system
CAVE	CAVE Automatic Virtual Environment
CCM	Combinatorial Coverage Measurement (software)
CENAM	Center for Metrology of Mexico
CI	configuration interaction
CICM	Conference on Intelligent Computer Mathematics
CFD	combination frequency differencing
CFSF	Continued Fractions for Special Functions
CLEO	Conference on Lasers and Electro-Optics
CMA	Computational Mathematics Research Group at the University of Antwerp
CN	core network
CNN	convolutional neural network
CONICET	National Scientific and Technical Research Council of Argentina
COVID	coronavirus disease
CPU	central processing unit
CRADA	cooperative R&D agreement

CSF	configuration state function
CST	Class-shape transformation
CT	combinatorial testing
CT	computed tomography
CTL	NIST Communications Technology Laboratory
CY	calendar year
CyEVaR	entropic cybersecurity value at risk
CyVaR	cybersecurity value at risk
DARPA	Defense Advanced Research Projects Agency
DART-MS	Direct Analysis in Real Time Mass Spectrometry
DC-QNet	DC area Quantum Network testbed
DGR	NIST Domestic Guest Researcher
DL	deep learning
DLMF	Digital Library of Mathematical Functions
DNA	deoxyribonucleic acid
DNN	deep neural network
DOC	Department of Commerce
DOI	digital object identifier
DQC	data quality control
DRMF	Digital Repository of Mathematical Formulae
E2E	end-to-end
EBSD	electron backscatter diffraction
eCF	Wolfram Computational Knowledge of Continued Fractions Project
eDNA	environmental DNA
EEO	equal employment opportunity
E-Health	electronic health care
E-Hy-CI	exponentially correlated Hy-CI
EI	electron ionization
EIT	electromagnetically induced transparency
EL	NIST Engineering Laboratory
EMD	earth mover distance
EOS	equation of state
EVar	entropic value at risk
FA	factor analysis
FEDVR	finite element discrete variable
FEM	finite element method
FFT	fast Fourier transform
FGR	NIST Foreign Guest Researcher Program
FY	fiscal year
GDML	Global Digital Mathematical Library
GHG	greenhouse gas
G-L	Gordon-Loeb optimization model
GMRES	generalized minimal residual iterative method
GNN	graph neural network
GPU	graphical processing unit
GSVE	generalized singular value expansion
GUI	graphical user interface
HEC	High End Computing
HEV	high end visualization
HMD	head-mounted display
HOM	Hong-Ou-Mandel interference effect
HPC	high performance computing
HPCVG	ACMD High Performance Computing and Visualization Group
HTML	hypertext markup language
HVAC	heating, ventilation, and air conditioning
HVACSIM+	software package and computing environment for simulating HVAC system
Hy	Hylleraas

Hy-CI	Hylleraas-Configuration Interaction technique
IARPA	Intelligence Advanced Research Projects Agency
ICST	International Conference of Software Testing
IEEE	Institute of Electronics and Electrical Engineers
IFIP	International Federation for Information Processing
IMA	Institute of Mathematics and Its Applications
IMKT	International Mathematical Knowledge Trust
IMS	NIST Innovations in Measurement Science program
IoT	Internet of things
IS	photoreceptor inner segment
is-CID	centroided in source collision induced (a type of mass spectra)
ISO	International Standards Organization
IT	information technology
ITL	NIST Information Technology Laboratory
IU	Indiana University
IVE	immersive visualization environment
IWCT	International Workshop on Combinatorial Testing
JHU	Johns Hopkins University
JPL	NASA Jet Propulsion Laboratory
JSON	an open standard file format and data interchange format
KKT	Karush-Kuhn-Tucker
KLS	Koekoek, Lesky and Swarttouw
KPI	key performance indicator
LAA	license-assisted access
LANS	locally adaptive network sparsification
LaTeX	a math-oriented text processing system
LaTeXML	a LaTeX to Math ML converter
LBNL	Lawrence Berkeley National Laboratory
LE	lattice element
Li	lithium
LN	lithium niobate
M3IC	DARPA Magnetic, Miniaturized, and Monolithically Integrated Components program
MAA	Mathematical Association of America
mAb	monoclonal antibody
MAC	medium access control
MathML	Mathematical Markup Language (W3C standard)
MDS	multi-dimensional scaling
MEGNet	materials graph network
MGI	Materials Genome Initiative
MIT	Massachusetts Institute of Technology
ML	machine learning
MLP	mathematical language processing
MML	NIST Material Measurement Laboratory
MOS	Magnus, Oberhettinger, and Soni
MOT	magneto-optical trap
MPS	matrix product states
MPI	Message Passing Interface
MR	mixed reality
MRAM	magneto-resistive random-access memory
MR	magnetic resonance
MRI	magnetic resonance imaging
MS	mass spectrometry
MSGI	NSF Mathematical Sciences Graduate Internship program
MSRI	Mathematical Sciences Research Institute (Berkeley)
muMAG	Micromagnetic Activity Group
nanoHUB	Web portal for nanotechnology research at https://nanohub.org/
NBS	National Bureau of Standards (former name of NIST)

NCCOE	NIST National Cybersecurity Center of Excellence
NCNR	National Center for Neutron Research
NDE	non-destructive evaluation
NISO	National Information Standards Organization
NISQ	noisy intermediate-scale quantum
NIST	National Institute of Standards and Technology
NISTIR	NIST Internal Report
NISTmAb	NIST monoclonal antibody reference material
NITRD	Networking and Information Technology Research and Development
nm	nanometer
NMI	national metrology institute
NRC	National Research Council
NREL	National Renewable Energy Laboratory
NSA	National Security Agency
NSF	National Science Foundation
OD	object detector
ODE	ordinary differential equation
ONA	organizational network analysis
OOF	Object-Oriented Finite Elements (software)
OOF3D	3D version of OOF
OOMMF	Object-Oriented Micromagnetic Modeling Framework (software)
OP	orthogonal polynomials
OPSF	orthogonal polynomials and special functions
OS	photoreceptor outer segments
OSA	Optical Society of America
PAML	physics-assisted machine learning
PCA	principal component analysis
PCR	polymerase chain reaction
PDE	partial differential equation
PDF	probability density function
PER	physics education research
PET	positron emission tomography
PH	Powell's hybrid method
PHY	physical layers (in a network)
PI	principal investigator
PIE	physics-informed excitation
PLOS	Public Library of Science
PML	NIST Physical Measurement Laboratory
PNNL	Pacific Northwest National Laboratory
POD	probability of detection
PQC	post-quantum cryptography
PREP	NIST Professional Research Education Program
PT	Painlevé transcendents
PUF	physically unclonable function
Q2B	quantum to business
QA	quantum annealing
QAO	quantum adiabatic optimization
QAOA	quantum approximate optimization algorithm
QD	quantum dot
QDPD	Quaternion-based Dissipative Particle Dynamics simulation code
QIP	quantum information processing
QoS	quality of service
qPCR	quantitative polymerase chain reaction
QuICS	UMD-NIST Joint Center for Quantum Information and Computer Science
R&D	research and development
RAVEN	IARPA Rapid Analysis of Various Emerging Nanoelectronics program
RBC	ray-based classification

REU	Research Experience for Undergraduates
REFPROP	Reference Fluid Thermodynamic and Transport Properties Database
RGM	reduced gradient method
RIM	reliability and integrity management
RPE	retinal pigment epithelium
RSA	Rivest-Shamir-Adelman public key cryptographic algorithm
RWC	region of wireless coexistence
SAM	self-assembled monolayer
SARS	severe acute respiratory syndrome
SARS-CoV-2	the virus that causes the respiratory disease COVID-19
SAVG	an internal HPCVG visualization file format
SDE	stochastic differential equation
SEM	scanning electron microscope
S-GRWA	symmetric generalized rotating wave approximation
SH	second harmonic
SHG	second-harmonic generation
SHIP	Summer High School Internship Program
SHM	structural health monitoring
SIGACT	ACM Special Interest Group on Computer Architecture
SIL	short iterative Lanczos
SIAM	Society for Industrial and Applied Mathematics
SiC	silicon carbide
SIS	susceptible-infectious-susceptible
SIGGRAPH	ACM Special Interest Group on Graphics
SNA	social network analysis
SNSPD	superconducting nanowire single-photon detectors
SPD	single-photon detectors
SPDC	spontaneous parametric down conversion
SPIE	International Society for Optical Engineering
SRM	standard reference material
STEM	science, technology, engineering, and mathematics
SUNY	State University of New York
SURF	NIST Student Undergraduate Research Fellowship program
SVD	singular value decomposition
SVM	support vector machine
SVOP	several variable orthogonal polynomials
SVP	NIST Student Volunteer Program
TDE	ThermoData Engine
TDSE	time domain Schrodinger equation
TES	transition edge sensor
TFLN	thin-film lithium niobate
TLS	transport layer security
TRC	NIST Thermodynamics Research Center
UC	University of California
UCSD	University of California at San Diego
UMD	University of Maryland
UMIACS	University of Maryland Institute for Advanced Computer Studies
UQ	uncertainty quantification
URL	universal resource locator
USGS	US Geological Survey
USRA	Universities Space Research Association
VaR	value at risk
VEMOS	Visual Explorer for Metric of Similarity (software)
VR	virtual reality
W3C	World Wide Web Consortium
WAF	web application firewall
WAS	Washington Academy of Sciences

WRF	weather research and forecasting
X3DOM	a graphics library
XAI	explainable AI
XR	extended reality
XSEDE	NSF eXtreme Science and Engineering Discovery Environment
XSS	cross-site scripting
ZE	Zeta and related functions