THE DOE CSGF ANNUAL

DEIXIS

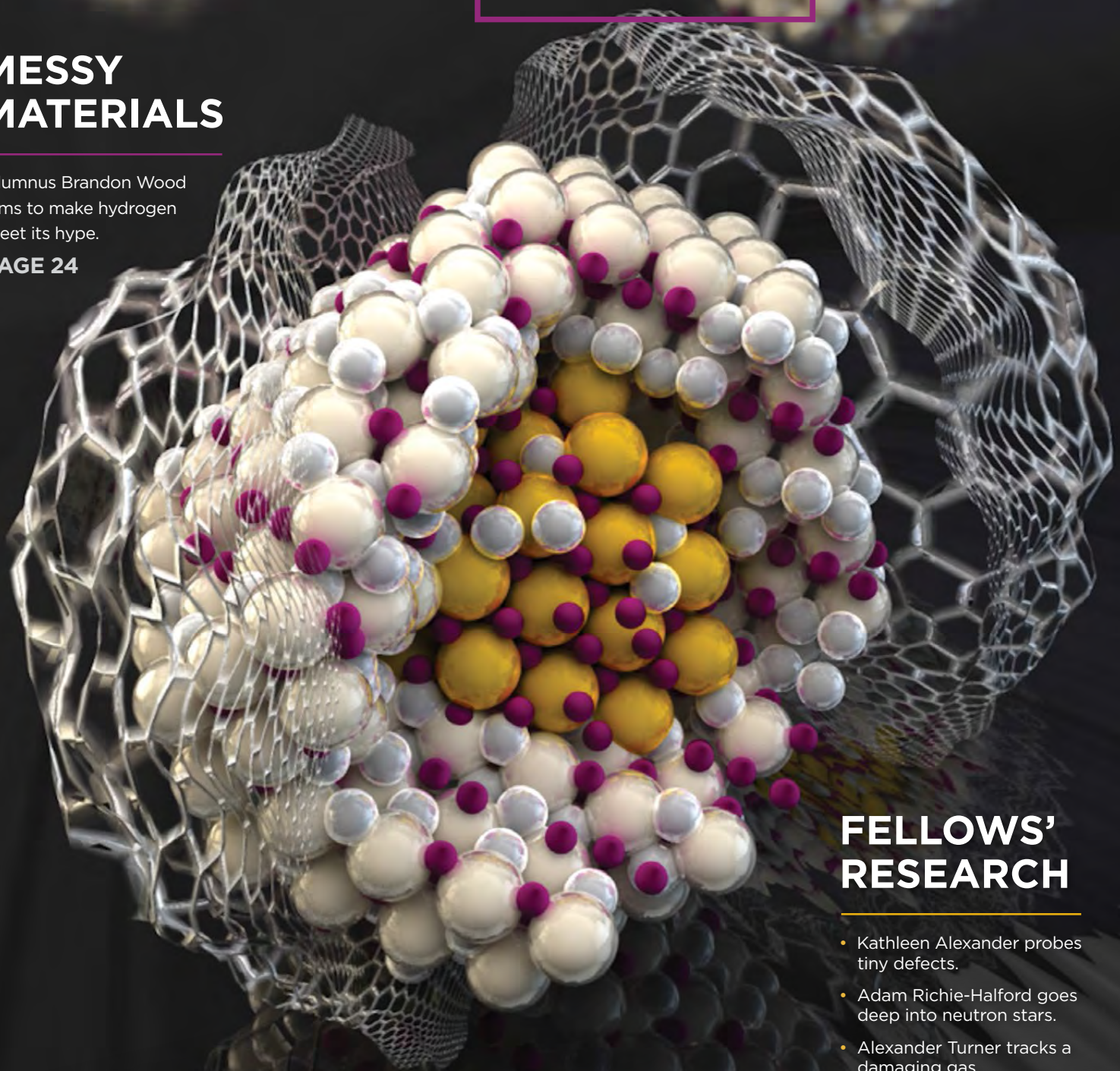
2017

DEPARTMENT OF ENERGY COMPUTATIONAL SCIENCE GRADUATE FELLOWSHIP

MESSY MATERIALS

Alumnus Brandon Wood aims to make hydrogen meet its hype.

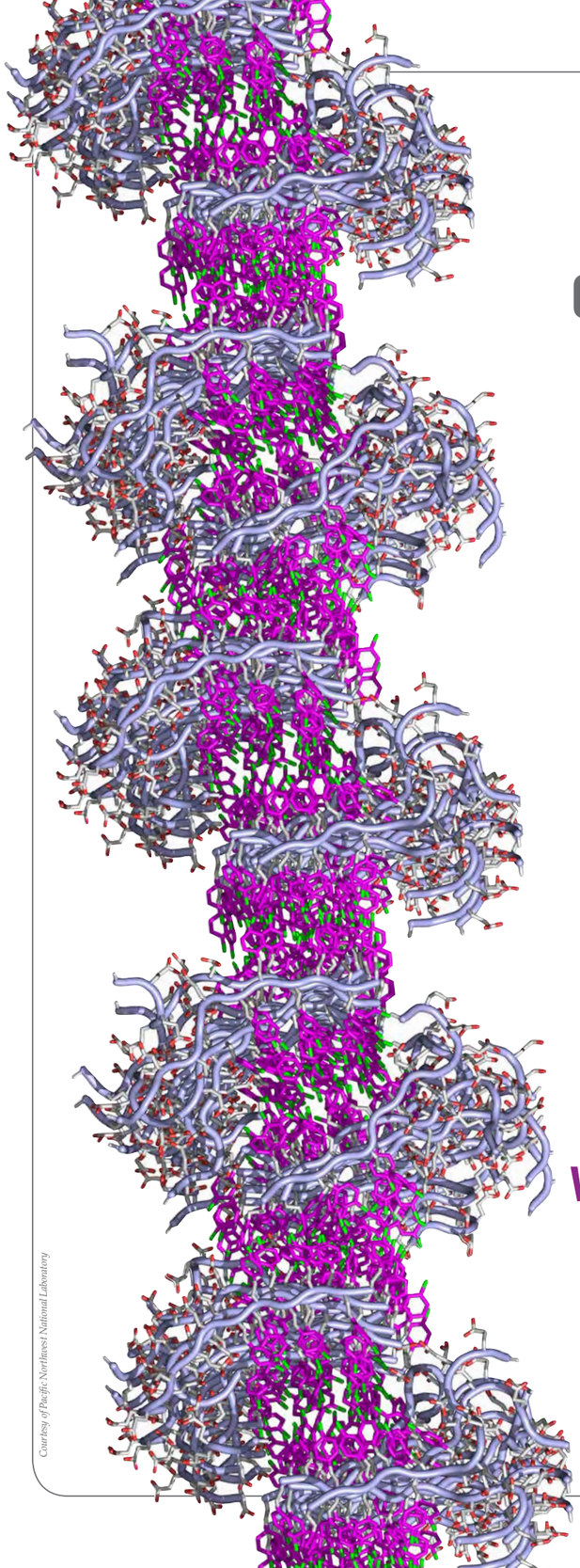
PAGE 24



FELLOWS' RESEARCH

- Kathleen Alexander probes tiny defects.
- Adam Richie-Halford goes deep into neutron stars.
- Alexander Turner tracks a damaging gas.
- David Ozog hunts for HPC speedups.

ALSO: Alumna Amanda Randles' virtual blood vessels, Howes Award winner Eric Isaacs, machine-learning horizons with MIT's Jeremy Kepner, and seeking the Goldilocks exoplanet.



Courtesy of Pacific Northwest National Laboratory

WELCOME TO THE 2017 INCOMING DOE CSGF CLASS

In fall 2017, 20 doctoral students enter the Department of Energy Computational Science Graduate Fellowship (DOE CSGF). These fellows were chosen in a rigorous screening process from more than 450 applicants. Each will receive yearly stipends, full tuition and fees and other benefits for up to four years, freeing them to pursue research and training in applied mathematics, computer science and an application area of their choice.

Peter Ahrens
Massachusetts Institute of Technology
Computer Science

Robert Baraldi
University of Washington
Applied Mathematics

Matthew Carbone
Columbia University
Chemical Physics

Gabriela Correa
Cornell University
Materials Science

Jennifer Coulter
Harvard University
Computational Condensed Matter Theory

Priya Donti
Carnegie Mellon University
Computer Science and Energy Policy

Annie Katsevich
New York University
Applied Mathematics

Jonas Kaufman
University of California, Santa Barbara
Computational Materials Science

Morgan Kelley
University of Texas at Austin
Chemical Engineering

Claire Kopenhafer
Michigan State University
Astrophysics

Alicia Magann
Princeton University
Computational Chemistry

Quentarius Moore
Texas A&M University
Chemistry

Kari Norman
University of California, Berkeley
Ecology

Miriam Rathbun
Massachusetts Institute of Technology
Computational Reactor Physics

Kevin Silmore
Massachusetts Institute of Technology
Chemical Engineering

Benjamin Toms
Colorado State University
Atmospheric Science

Steven Torrisi
Harvard University
Quantum Physics

Annie Wei
Massachusetts Institute of Technology
Quantum Algorithms

Zachary Weiner
University of Illinois at Urbana-Champaign
Cosmology, High Energy

Malia Wenny
Harvard University
Physical Chemistry

DEIXIS, The DOE CSGF Annual is published by the Krell Institute. Krell administers the Department of Energy Computational Science Graduate Fellowship (DOE CSGF) program for the DOE under grant DE-FG02-97ER25308.

For additional information about the DOE CSGF program, the Krell Institute or topics covered in this publication, please go to: www.krellinst.org/csgf

Or contact:
Editor, DEIXIS
Krell Institute
1609 Golden Aspen Drive, Suite 101
Ames, IA 50010
(515) 956-3696

DEIXIS (ΔΕΙΞΙΣ — *pronounced dāksis) transliterated from classical Greek into the Roman alphabet, means a display, mode or process of proof; the process of showing, proving or demonstrating. DEIXIS can also refer to the workings of an individual's keen intellect, or to the means by which such individuals, e.g. DOE CSGF fellows, are identified.*

DEIXIS is an annual publication of the Department of Energy Computational Science Graduate Fellowship program that highlights the work of fellows and alumni.

DOE CSGF funding is provided by the DOE Office of Advanced Scientific Computing Research (ASCR) within the Office of Science and the Advanced Simulation and Computing (ASC) program within the National Nuclear Security Administration.

Manager, Science Media — Bill Cannon
Science Media Editor — Thomas R. O'Donnell
Creative Project Coordinator — Buffy Clatt
Design — Stilt Studio, Inc.

ON THE COVER: A model of an indium phosphide semiconductor in contact with water. The material is a candidate to split water molecules with sunlight to produce hydrogen for energy production. Alumnus Brandon Wood is computing properties of such materials, helping to realize the promise of a hydrogen economy. Read more starting on page 24. *Credit: Brandon Wood.*

JAMES CORONES
1945–2017

Visionary leader oversaw the fellowship for a quarter century, almost from its creation.

With the April death of James Corones, the Department of Energy Computational Science Graduate Fellowship (DOE CSGF) lost perhaps its greatest champion.

Jim managed the program almost from its creation in 1991. He "was a visionary who was able to see early on the impact computational science would have on the scientific community, even when we were still trying to define what computational science meant," says David L. Brown, a program advisor since its start and now director of the Computational Research Division at DOE's Lawrence Berkeley National Laboratory. At the time, few scientists understood how to employ computation in their research or grasped its benefits. DOE laboratories had the world's best machines but usually had to train researchers in their use.

The DOE CSGF helped change that, and in the process contributed to establishing computation as the third leg of science, with theory and experiment. Jim played a vital role, says Roscoe Giles, a Boston University professor of computer and electrical engineering and physics and a fellowship steering committee member: "He had the unique skill to see ways to advance both science and young scientists."

In 1997, Jim founded the Krell Institute to manage the DOE CSGF and other programs. The fellowship expanded further, drawing hundreds of applicants each year. Under Jim's guidance as president until his retirement in late 2016, his staff developed innovative electronic application processing and screening systems, creative recruitment campaigns and far-reaching science communication projects.

A strong advocate for effectively publicizing science information, Jim founded an essay contest for DOE CSGF fellows and alumni, encouraging them to write for a lay audience. He also launched this magazine, an annual journal that recruits applicants as it celebrates the accomplishments of fellows and alumni. Krell later developed a companion website to highlight computational science at the DOE national laboratories.

Shelly Olsan, who succeeded Jim as president, joined the company in 2000 to manage communication and outreach efforts. "We spent many long days brainstorming ways to bring DEIXIS to life. We wanted to convey how cool science is while highlighting the people doing it. Jim was passionate about the people first, whether students or accomplished scientists, and about showcasing their intellect and contributions," she says.

Jim saw Krell as a vehicle to assist with the education of superior scientists for the U.S. workforce, helping the country continue to lead the world in multiple disciplines. The DOE CSGF is proof of his success, with more than 400 fellows and alumni. Each is an ambassador for computational science and a testament to Jim's foresight and energy.

Jim earned a physics degree from Brown University and went on to a doctorate in the same subject from Boston University. After a Fulbright scholarship in Poland and a postdoctoral fellowship at the University of Calgary, he joined Iowa State University in 1973 as an assistant professor of mathematics. He conducted research in linear and nonlinear wave propagation and worked with acoustic and electromagnetic inverse scattering problems, with support from the DOE, the Office of Naval Research, the Air Force Office of Scientific Research and the National Science Foundation.

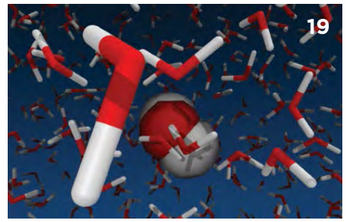
Later, Jim joined the Ames Laboratory, a DOE facility on the ISU campus, and became manager of its Applied Mathematics program. He went on to hold several administrative posts at the lab, including deputy director and acting director.

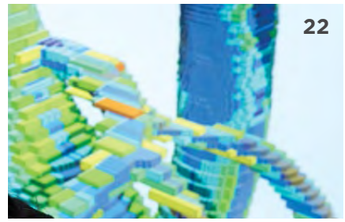
Brown says Jim was "a thoughtful, caring person with a great sense of humor that helped him and the people around him weather the ups and downs in federal politics that often present challenges to science programs," including the DOE CSGF.

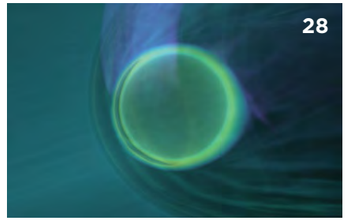
Giles agrees. Jim had a "tenacity of purpose that inspired all of us who worked with him." At Krell, Olsan says, that inspiration "challenged us to always do better, bigger and more inventive things."

Even after retirement, Jim remained engaged with the fellowship and provided advice and support until his death. Computational science, postgraduate education and the nation are stronger for his efforts. His legacy is the programs he developed and the scientists they train.

IN THIS ISSUE







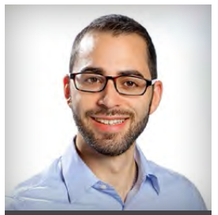
Departments

Features

MAKING A MATERIAL DIFFERENCE

Eric Isaacs' outreach extends across oceans and across town as he uses computers to seek valuable new compounds.

By Thomas R. O'Donnell



Eric Isaacs

The possibility of discovery intrigues Eric Isaacs as he plumbs the properties of materials, both real and computationally imagined.

As a Columbia University doctoral student, he used blended mathematical methods to calculate battery materials' little-understood properties. Now, as a Northwestern University postdoctoral researcher, he mines a materials database, seeking compounds capable of efficiently converting heat directly to electricity.

"It's possible to find a material better than anything that exists today" for a specific purpose, says Isaacs, a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient from 2012-2016. "That's pretty awesome."

Isaacs cooperates in that quest with scientists who create and test the materials computers have forecasted. "To be able to walk down the hall and say, hey, this compound looks really good; do you want to try to make it? – it feels like a very real possibility to actually discover some new, interesting compounds."

For his research excellence and leadership in and out of the laboratory, a panel of DOE CSGF graduates and friends has named Isaacs the 2017 Frederick A. Howes Scholar in Computational Science.

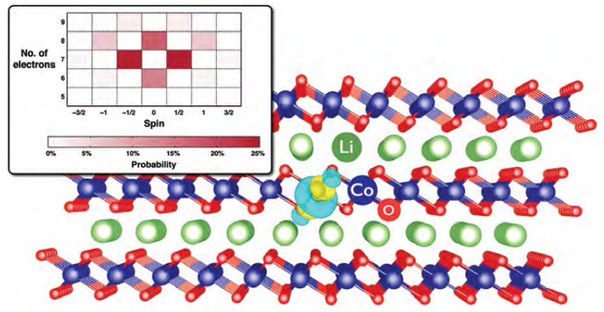
Howes' colleagues established the award after his unexpected death in 2001. Isaacs will receive the honor and deliver a talk on his research at the 2017 DOE CSGF Annual Program Review in Arlington, Virginia, in July.

"It's a weird feeling" to be chosen, Isaacs says. "I know the caliber of my fellow CSGFers and it's a really incredible group of people" with amazing abilities. It's an honor to be counted as exceptional in that company, he adds.

Isaacs was cited for his extensive outreach activities, including work with nanoHUB.org, a national web repository for computational nanoscience materials; helping with post-hurricane rebuilding in New Orleans; and more.

Isaacs went to Ethiopia in 2013 as part of the Joint U.S.-Africa Materials Initiative, which fosters materials science research and collaboration. For two weeks, the American and East African scientists engaged in an extended workshop of classes, talks and discussions, with an emphasis on renewable energy production.

"That's kind of a remarkable thing, to collaborate with someone halfway across the world," Isaacs says, but it's even more unusual because of the trials his counterparts face. For them, "the challenges are not only scientific but also related to infrastructure. You're trying to do an experiment and your power goes out." Isaacs still corresponds with many of the scientists he met.



Crystal structure of lithium cobalt oxide, a widely used cathode for rechargeable batteries. The yellow and cyan bubbles at center show the cobalt d electron orbital involved in reactions during charge and discharge. Density functional theory plus dynamical mean-field theory calculations capture the substantial dynamics of the electron-electron interactions, which result in the probability distribution of different quantum states shown in the inset. *Credit: Eric Isaacs.*

More recently – and more locally – Isaacs worked with a Chicago-area high school freshman as part of Northwestern's Mentorship Opportunities for Research Engagement (MORE) program. He helped the student explore piezoelectric materials, which produce electricity when deformed or strained. The idea was to make sneakers that could produce power with each step.

In a process similar to what Isaacs does in his research, he helped the student do electronic structure calculations of simple piezoelectric materials. The duo did interesting work, but ultimately the science may have been secondary, Isaacs says. The student "got an idea of what it's like to do research, what is an interesting question, how one can answer it, what are the limitations." Few people outside the research world understand that.

The experience reminded Isaacs of his high school and college years and how important adults and fellow students were to his educational and professional development. "Any way I can contribute in that way or send the elevator down to younger students or researchers is important."

As part of Chris Wolverton's group at Northwestern, Isaacs works to speed the search for new, useful materials. He focuses particularly on efficient thermoelectric compounds, which convert heat into electricity.

Isaacs' search is largely focused on the Open Quantum Materials Database (OQMD). Wolverton's group has compiled this collection of compounds and the thermodynamic and structural properties of each as estimated with density functional theory (DFT). DFT calculates atomic interactions while accounting for quantum physics, under which electrons behave as both waves and particles. OQMD has more than 500,000 materials catalogued, many of them hypothetical — they've never been synthesized.

Isaacs uses data mining and machine learning to screen thousands of materials to find ones with the best thermoelectric properties. "It's not trivial to do this because to actually, in a straightforward way, calculate the thermoelectric efficiency is very difficult." It can't be done for every substance, so he seeks simple descriptions for what would comprise a good thermoelectric material, based on its atomic structure, chemistry and electronic properties. "A large database like this really enables that," providing a head start on identifying candidates. With that information, Isaacs can run more detailed calculations on the most promising possibilities.

Isaacs' doctoral work was in a similar vein, but aimed at developing and using mathematical methods to probe battery materials. Working with Chris Marianetti, he combined DFT with another method, dynamical mean-field theory (DMFT), to perform demanding electronic structure calculations for compounds like lithium cobalt oxide and lithium iron phosphate, both popular rechargeable battery materials. "Even though we're using this technology every day, there's a lot of physics of it that's not understood," Isaacs says.

In what may be a first, Isaacs used the DFT-DMFT method to compute the total electronic structure energy in complex oxides. The group derived values for the materials' phase stability (capacity to maintain structure), battery voltage and other properties.

Isaacs expects to stay at Northwestern for a year or two, then seek a position with a DOE laboratory or in academia. Both are appealing, but "regardless of what environment I'm in, it would be great to work collaboratively and help train scientists."

Isaacs' advice to new fellows is based on his experience at his first annual program review. When it came to computational science, "I had no idea what people were talking about." To learn, fellows must be unafraid to ask questions.

"Be assertive and maybe even aggressive, in some sense, in grabbing opportunities" for a practicum and networking. "Those four years go by pretty quickly. It's a unique opportunity to have access to these people who are part of the DOE CSGF."

ABOUT FRED HOWES

The Frederick A. Howes Scholar in Computational Science award, first presented in 2001, has come to stand for research excellence and outstanding leadership. It's a fitting tribute to Howes, who was known for his scholarship, intelligence and humor.

Howes earned his bachelor's and doctoral degrees in mathematics at the University of Southern California. He held teaching posts at the universities of Wisconsin and Minnesota before joining the faculty of the University of California, Davis, in 1979. Ten years later Howes served a two-year rotation with the National Science Foundation's Division of Mathematical Sciences. He joined DOE in 1991 and advocated for the fellowship and for computational science as manager of the Applied Mathematical Sciences Program.

Howes died unexpectedly in 1999 at age 51. Colleagues formed an informal committee to honor him and chose the DOE CSGF as the vehicle. With donations, including a generous contribution from Howes' family, they endowed an award in his name.

LEARNING MACHINES

Jeremy Kepner, a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient from 1993 to 1996, is a Lincoln Laboratory Fellow at the Massachusetts Institute of Technology, where he founded the Lincoln Laboratory Supercomputing Center, and a research affiliate at MIT's Mathematics Department. The topic of his keynote speech at the 2017 DOE CSGF Annual Program Review was machine learning.

DEIXIS: WHAT'S YOUR DEFINITION OF MACHINE LEARNING?

Jeremy Kepner: That's a big question. Machine learning, like many fields, is perhaps best understood in terms of the important questions it tries to answer and how it tries to answer them. Some of these questions have been around since the field's inception. In fact, the first machine-learning meeting in 1955 had four papers. Three were on problems: language, vision and strategy. And one was on an approach: neural networks.

WHY HAS IT COME TO THE FORE LATELY?

Computers have made enormous progress in the past few years, achieving human-level performance in language, like understanding spoken commands, in vision, such as recognizing faces in pictures, and strategy, including playing games like chess or go.

WHY IS THIS IMPORTANT? WHAT CAN MACHINE LEARNING DO FOR US?

I think we feel the impact of machine learning every day. Internet search, social media and self-driving cars depend on it. There's also tremendous interest in applying machine learning to helping solve the nation's problems – developing new medicines, defending against cyber threats and making travel safer.

WHY ARE HPC SYSTEMS NEEDED FOR THIS?

In machine learning's earliest days, experts recognized it would take an enormous amount of computation to train a neural network to solve problems in language, vision or strategy. Much of the recent progress in machine learning has been from vast increases in computing power, such as GPUs (graphics processing units), that originally were developed by the video game industry. These have since become core components of many supercomputing systems.

YOU TITLED A PREVIOUS TALK "THE CONVERGENCE OF BIG DATA, MACHINE LEARNING AND SUPERCOMPUTING." HOW ARE THEY CONVERGING?

Machine learning's vast computing requirements are a key driver in the computing industry and provide key components to supercomputing. Machine learning also requires large data sets to train algorithms, and it's used to categorize ever-larger streams of data. These requirements are also driving big data technologies.

YOU MENTIONED NEURAL NETWORKS. HOW ARE DEEP NEURAL NETWORKS (DNNS) DIFFERENT FROM STANDARD MACHINE LEARNING?

Machine learning consists of many different approaches. Deep neural networks are one that has recently proven useful for many important problems. What's been interesting is how flexible DNNs have proven to be.

HOW DOES A DNN WORK?

At its simplest, a DNN is a black box that allows you to put in a picture of a cat and out pops the word "cat." Mathematically, it's a matrix problem that takes an input vector – the cat picture – and runs it through multiple layers of the DNN, using weighted data matrices and decision-making bias vectors. It produces an output vector that ideally has a value of 1 in the position corresponding to the word "cat" and zeroes for all other values, such as dog, rabbit, or cow. Training DNNs to make these decisions is a big computational challenge that involves iteratively searching for good weight matrices and bias vectors, given a large number of known inputs and outputs. Amazingly, DNNs that achieve human levels of performance are possible, given enough training data and computational power.

IT SOUNDS LIKE PART OF THE TRICK IS MAPPING THE MACHINE-LEARNING PROCESS AND ALGORITHM TO THE MULTIPLE CORES IN A PARALLEL PROCESSOR OR MACHINE.

DNNs are so computationally demanding that exploiting parallel processing is a high priority if we're to run them quickly. But they're often dominated by matrix multiplications, and figuring out how to map these to a parallel processor can require a great deal of expertise. And a good mapping for one problem on a particular processor will not be a good mapping for a different problem on a different processor.

YOUR LAB HAS DEVELOPED SEVERAL REMAPPABLE SOFTWARE PACKAGES FOR FUTURE SYSTEMS. WHY IS REMAPPING USEFUL AND HOW DOES IT RELATE TO DNNS OR MACHINE LEARNING?

When mapping a DNN program to a parallel processor, it's natural to simply write the program so that how the matrix data are divided among processors is part of the code. This lets you precisely hand-tune the code, but a more scalable way to write the program is to separate it from how it's mapped, so mapping is independent. That lets users choose different maps without rewriting the code.

YOU AND YOUR COLLEAGUES ARE APPLYING MACHINE LEARNING TO ENABLE AUTONOMOUS SENSORS AND SYSTEMS. HOW DOES MACHINE LEARNING ENTER INTO THAT?

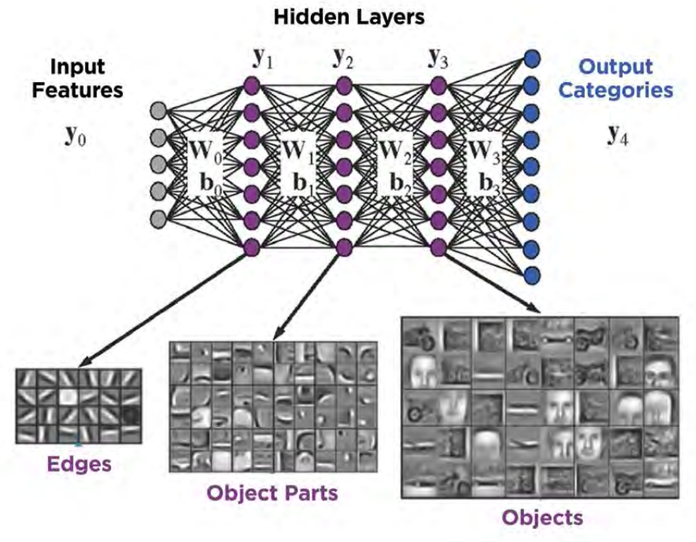
One example is Lincoln Laboratory's work in air traffic control, where machine learning was used to improve weather forecasts used for routing airplanes. The project recently won an R&D 100 award, recognizing it as one of the best new technologies of the year.

WHAT'S NEXT? WHERE IS THIS HEADED?

As the demand for computation continues to increase, the use of more exotic technologies to do specialized computations is going to only grow. Understanding the physics of these technologies and the kinds of the mathematics they support will increasingly determine which areas will be enabled by new computing technologies.

WHAT MAKES YOU EAGER TO GET TO WORK?

I feel like I have the best job at MIT. Every day my team gets to work with all the brilliant scientists and engineers here to help them move forward the frontiers of science and engineering. At the same time, we also get the opportunity help advance the frontiers of computation, algorithms and mathematics.



Four-layer deep neural network architecture for categorizing images. The input features (y_0) of an image are passed through a series of network layers, W_0 through W_3, with bias terms b_0 through b_3, that produce scores for categories y_L=4. *Credit: H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations," in* Proceedings of the 26th annual international conference on machine learning. *ACM, 2009, pp. 609–616.*

GRAIN BY GRAIN

The 'grain boundary' where materials stitch themselves together can introduce tiny defects that exert a big influence on how those materials behave. Kathleen Alexander enlists high-performance computing to expose those flaws.

By Sarah Webb

Kathleen Alexander didn't think much about science or mathematics while growing up in rural northern Idaho. Her small high school didn't even offer a physics course. But for an economics project, she proposed incorporating a composting company into city government. It would collect organic waste, compost it and sell it to farmers, then reinvest the profits to improve recycling technology. "We could then develop a whole new ecosystem for this zero-landfill, closed-loop materials cycle," she says.

The project sparked an interest in materials, the first step on Alexander's unconventional educational path to a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) and a doctorate from the Massachusetts Institute of Technology.

At MIT, Alexander developed computational methods to help researchers understand minuscule defects in metals with the goal of improving their properties, a key to developing higher-efficiency engines and energy storage cells. She hasn't strayed far from her high school project's aims: using science and technology "to really push the limits on what it would mean to be a sustainable society."

Those goals led Alexander to revamp her career path. To find new recycling technologies, "I would need to pursue engineering, and I didn't, clearly, have the background." So Alexander jettisoned plans to attend the University of San Francisco, which had no engineering program. Instead, she shored up her math and science knowledge with two years at American River Community College in Sacramento, California.

Alexander's transfer to a four-year college in the California state university system was virtually guaranteed. But while weighing the options, she attended her brother's graduation from Rose-Hulman Institute of Technology in Indiana.

"One of the speaker's major takeaways was, 'You don't get things you don't ask for, and you should ask for everything,'" Alexander says. So, besides the University of California, Davis, she applied to four-year technical colleges across the country, including MIT, where she was accepted in 2008 to study materials science and engineering.

When Alexander learned it would take three more years of school, not two, to complete her undergraduate degree, she opted for a yearlong exchange program with Great Britain's University of Cambridge. The experience let her dig deeper into metallurgy and other materials science topics.

At MIT, Alexander also delved into both experimental and computational research. With supervision from a postdoctoral researcher in Vladimir Bulović's group, she helped double the efficiency of a device that split water into hydrogen and oxygen gases.

Alexander also did a summer 2010 Sandia National Laboratories internship in New Mexico, using computational modeling to study erbium tritide, a material involved in neutron generation. "Those two experiences combined drove me to be really interested in doing computational work, but as a component of experimental research."

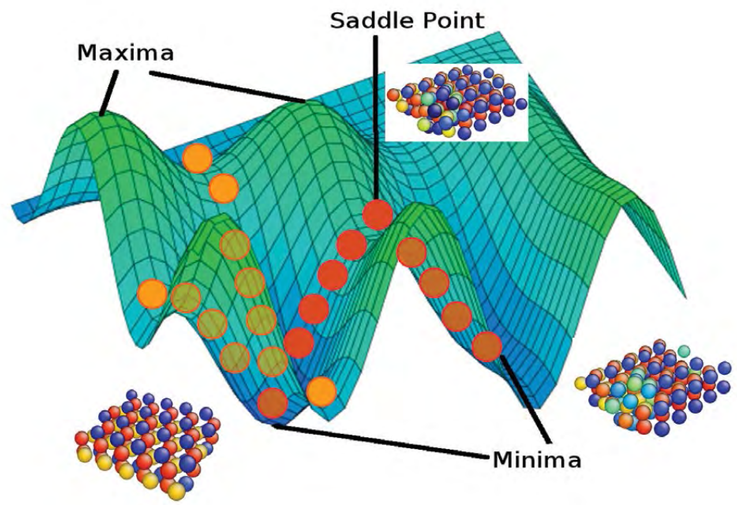
Several Ph.D. programs had accepted Alexander by the time she graduated in May 2011, but she stayed at MIT to work with Christopher Schuh, a metallurgy professor who studies material defects far smaller than the eye can see. Their research philosophies aligned well. At Sandia, Schuh says, Alexander "had become interested in this idea of understanding and controlling defects in materials, and she had fallen in love with the idea of using computers to do it."

But a different love led Alexander on another detour before graduate school. At Sandia, she met William Lane, an engineering student from Canada. They married in 2011 and Alexander moved to New Brunswick while Lane finished his undergraduate degree. It proved to be a valuable professional experience: Alexander took additional mathematics classes and did computational research with Ghislain Deslongchamps, a University of New Brunswick chemistry professor.

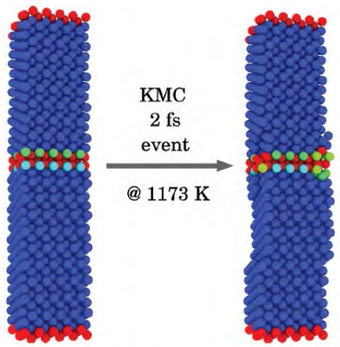
After returning to MIT, Alexander examined the atomic structure and organization of materials and their defects. In metals, such inquiries require an understanding of grain boundaries – defect features that shape fundamental properties of the materials. There are few reliable ways for researchers to study grain boundaries and how they relate to material behavior, so Alexander focused on developing a new computational method to address the challenge.

The angles between grains – the individual crystals that comprise materials – make a big difference in how materials perform. Grain boundaries determine properties related to failure, like cracking and corrosion, as well as qualities that dictate a material's manufacturability, like how easily it welds or how pliable it is.

That means the ability to understand and engineer grain boundaries can vastly improve a substance's performance. In the 1990s, grain-boundary engineering let researchers eliminate corrosion in nickel-based alloys used in caustic



In this potential energy landscape, atomic arrangements of a grain boundary modeled in copper are shown as examples of configurations for the indicated points of interest: saddle point (the energy barrier to an atomic rearrangement) and minima (low-energy, stable configurations). The atoms are colored according to centrosymmetry (blue atoms are in a perfectly symmetric environment and red atoms are in a highly asymmetric environment), and only those on the boundary are shown. The potential energy surface is plotted as energy versus a two-dimensional projection of configuration space. *Credit: Kathleen Alexander; images generated with the AtomEye visualization program.*



The same copper grain boundary (layer of red, green and blue atoms) as the preceding page, modeled before and after a kinetic event at 1,173 degrees Kelvin (900 degrees Celsius) lasting 2 femtoseconds as calculated by the kinetic Monte Carlo (KMC) algorithm. The atoms are colored according to their centrosymmetry parameter (blue atoms are in a perfectly symmetric environment and red atoms are in a highly asymmetric environment). The event leads to a local distortion of the grain boundary plane (visible in the rearranged atoms in the grain boundary), moving it out of the energy ground state configuration. This local distortion leads to a thickening of the grain boundary plane and disrupts its symmetry. *Credit: Kathleen Alexander, using Ovito Open Visualization Tool.*

environments and at the electrodes in lead-acid batteries, drastically improving product lifetimes.

Schuh says Alexander addressed a major issue in computational materials science: the timescale challenge of matching simulations with laboratory results. It often takes weeks for computers to calculate chemical behavior that occurs in a few nanoseconds, but experimentalists typically can't measure phenomena happening over such a short period.

Alexander's method modeled chemical behavior for many thousands of times longer, providing data that are easier to match with experiments, Schuh says. She then validated the method in chemical systems such as chromium and tungsten mixtures. The initial work was done on a 64-core workstation, but the method's kinetic Monte Carlo framework relies on parallel computing and should be well-suited to scale up to thousands of processors on DOE supercomputers.

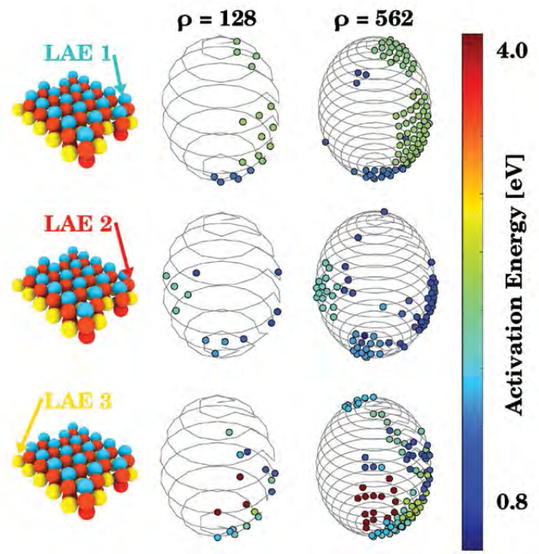
Alexander's DOE CSGF practicum focused on a different grain-boundary problem: how defects influence a promising material for rechargeable lithium batteries. Working with Bobby Sumpter and his Oak Ridge National Laboratory team, Alexander used electronic structure calculations based on fundamental physics to examine composition, organization and transport in lithium lanthanum titanate, a promising solid electrolyte. Such substances are important for batteries because they would be safer and could last longer than liquid electrolytes used today. But researchers must understand what limits ion mobility in the material and whether a simple strategy could overcome those barriers.

In her few weeks at ORNL, Alexander worked with electron microscopists and computational researchers to connect experiments and simulations, finding ways to identify lithium lanthanum titanate defects that improve or inhibit performance. Her accomplishments in such a short time were impressive, Sumpter says. "Kathleen is an extremely smart researcher, an excellent communicator and works very hard, attributes that are important for working in a multidisciplinary environment. She does her homework. Always."

Alexander's range of experiences – at a community college, at MIT, internationally and at Sandia – helped her stand out, Schuh says. "She had seen all these different environments. She had a nuanced understanding of what professional research work looked like."

Alexander finished her doctorate in July 2016 and took a research position at Nucleus Scientific, an MIT spinoff. Its proprietary technology connects to the mission – using science to support sustainability – that launched her career.

But, Schuh says, Alexander has already made important progress toward that goal: "She's done work that paves the way for a lot of exciting new discoveries."



Atomic perturbation directions that led to identifying a potential energy saddle point are plotted on a unit sphere (colored according to the energy needed to activate their corresponding transition state) to create a transition globe. Transition globes of atoms in three different local atomic environments (LAEs) are shown for perturbation meshes with ρ = 128 and ρ = 562 (where ρ is the number of mesh points on each radial shell of the perturbation space). *Credit: Kathleen Alexander.*

JOURNEYMAN

After delaying his studies for the Peace Corps, Adam Richie-Halford plumbs the quantum depths of distant, dense objects.

By Thomas R. O'Donnell

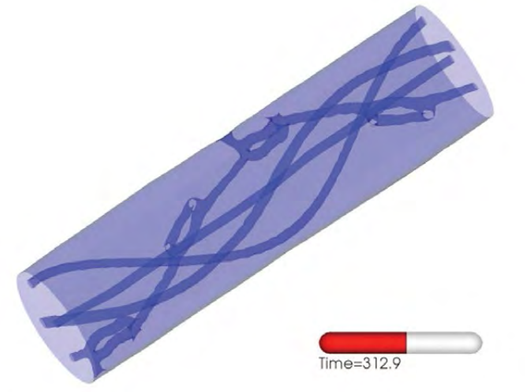
Adam Richie-Halford was gung-ho for the Air Force when he joined the ROTC program at Embry-Riddle Aeronautical University. As the son of an airline aviator, he dreamt of becoming a test pilot.

Once he entered the engineering physics curriculum, however, his priorities shifted to science and engineering, Richie-Halford says. "I realized I'm more of a nerd than I thought."

After graduation, Richie-Halford was assigned to Los Angeles Air Force Base, home of the Space and Missile Systems Center. It was interesting but sometimes frustrating work. For example, he was helping to resolve interference in signals from a communications satellite when his superior asked why the project was taking so long. "I started trying to explain the problem. And he said, 'Adam, we don't want to understand the problem. We want to fix it.'"

Richie-Halford's hunger to grasp a puzzle's roots led him to a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) and research at the University of Washington in Seattle that explores some of matter's strangest behaviors.

While in the Air Force, Richie-Halford had earned a physics master's degree and became fascinated with quantum mechanics – the rules that govern subatomic interactions. In the quantum world, energy and matter interact as both particles and waves, and two particles can influence each other over great distances.



Quantized vortex connection and separation in the Unitary Fermi Gas. As two vortices approach each other, they combine, exchange vortex material and separate. This is the underlying mechanism of quantum turbulence. *Credit: A. Bulgac, Y.-L. Luo, P. Magierski, K.J. Roche, and Y. Yu, Real-Time Dynamics of Quantized Vortices in a Unitary Fermi Superfluid,* Science, *332, 1288 (2011).*

"Usually when you think about quantum physics, you're talking about microscopic things you can't see," Richie-Halford says. "But there are some phenomena that would not exist without quantum effects, like superconductivity and superfluidity, that you can see macroscopically."

With UW physics professor Aurel Bulgac, Richie-Halford uses high-performance computing (HPC) to model quantum neutron interactions under extreme conditions, such as inside neutron stars. These dead star remnants are extraordinarily dense – a half-gallon carton of neutron star matter would weigh as much as Seattle-area landmark Mount Rainier – and can have powerful magnetic fields. Neutron stars often reveal themselves as pulsars, sending out jets of particles as they rotate and appearing to radio astronomers as regularly spaced bursts of radiation, like the sweep of a lighthouse.

Neutron stars are the biggest manifestation of the dense nuclear systems Bulgac and Richie-Halford simulate. What they learn also could help explain how protons and neutrons interact in atomic nuclei.

Scientists theorize that neutrons inside these stars become superfluid – flowing with no loss of kinetic energy. They also may develop quantum vortices – superfluid quantum tornados with empty cores large enough to engulf atomic nuclei. The vortices may become pinned on density irregularities in the star's crystal crust, and their pinning and depinning is believed to cause pulsar glitches in which the stars occasionally spin faster than usual.

Physicists disagree on the mechanisms behind pulsar glitches and vortex pinning. Computer simulations provide clues, but it's a massive many-body problem, with each neutron strongly interacting with others in a complex dance.

Bulgac and his colleagues model superfluid vortices with a technique based on density functional theory (DFT), which dramatically reduces a problem's dimensions. The simulations, however, start with neutrons at a low-energy initial state – a "snapshot at time zero," Richie-Halford says. "It was only so long before I said, well, how do we get the initial state?"

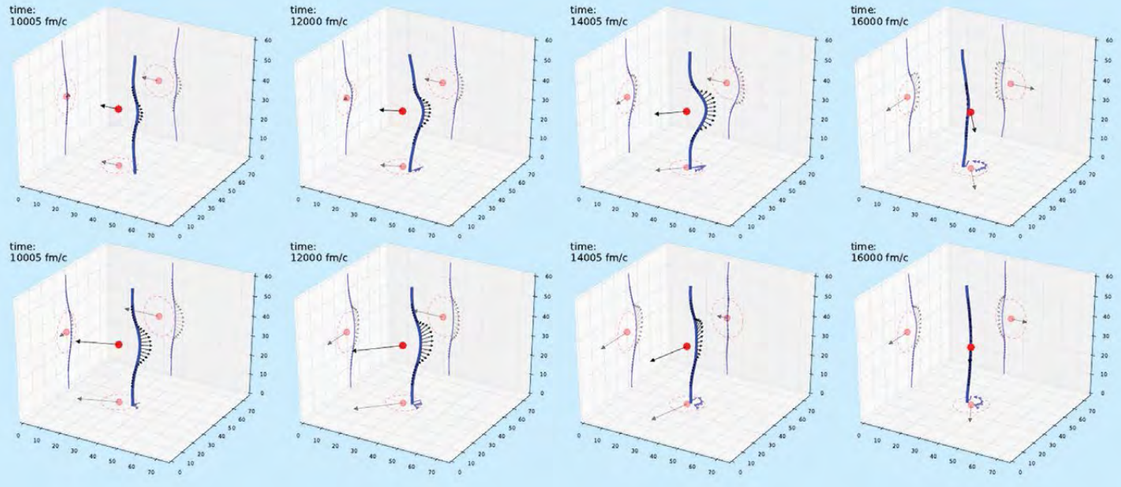
But finding that, Bulgac says, "is not a trivial thing in quantum mechanics, especially if you have a system of many particles strongly interacting with each other," as in a neutron star. As with the quantum vortex case, mathematicians and physicists seek ways to manage the problem.

That brings in a second mathematical technique: quantum Monte Carlo simulation. Monte Carlo methods randomly sample the myriad possible parameter values to arrive at the most likely low-energy state. Their name comes from Monaco's Monte Carlo casino because they rely on chance.

"I love doing Monte Carlo simulation," Richie-Halford says. "I like the idea that you can use the dice-rolling, the random computational noise to simulate actual physical thermal noise. Once I made that connection I was hooked."

To understand these methods, it helps to think of neutron matter's energy as a landscape with hills and valleys. Richie-Halford wants to find the true ground state, in the lowest valley. He starts with a random initial state and then perturbs it slightly, calculating the resulting energy differences. In essence, the algorithm takes a random walk across the hills until it reaches the valley – the low-energy ground state.

While working in Morocco for the Peace Corps, Adam Richie-Halford focused on teaching hygiene and health skills, like brushing teeth. *Credit: Zoë True.*



Simulated vortex-nucleus pinning in low-density neutron matter. From left to right, frames are about 6 x 10^{-21} seconds apart. The blue lines indicate the position of the vortex core while the red dot indicates the center of mass of protons inside the nucleus. Nuclei repel vortices in the neutron star crust, leading to interstitial vortex pinning. *Credit: From Gabriel Wlazłowski, Kazuyuki Sekizawa, Piotr Magierski, Aurel Bulgac, Michael McNeil Forbes, "Vortex pinning and dynamics in the neutron star crust,"* Phys. Rev. Lett. 117, *232701 (2016).*

It may not yield the precise quantity for the system's minimum energy, but the method provides "an upper bound in our approach." Comparisons with other techniques show it's usually near the true minimum.

Richie-Halford's research goes a step further to explore quasiparticle properties, portraying the strongly interacting neutrons as a system of weakly interacting particles, like a gas. They're called quasiparticles "because they're not real, but we can pretend they're real" to simplify the calculations.

Richie-Halford calculates what properties a quasiparticle formulation must have – what its mass should be, for example – to accurately capture real system interactions. "Then, hopefully, people can use that quasiparticle picture" to learn about strongly interacting systems without huge simulations. He's found that the effective mass of the strongly interacting neutron system differs little from the noninteracting version. That's surprising because they behave so differently.

Richie-Halford's research has run on Titan, a Cray XK7 at Oak Ridge National Laboratory; Stampede, a Dell at the University of Texas; Hyak, UW's HPC cluster; and other systems.

Bulgac says Richie-Halford's interest in learning multiple modeling techniques sets him apart from graduate students who typically learn one major tool and devote their careers to it. "Because he has so many tools at his disposal, he'll be able to apply to (work at) more places than other people."

After leaving the Air Force, Richie-Halford was determined to pursue a science career. He approached Bulgac about joining his group but delayed entry so he and his wife, Zoë True, could serve two years in Morocco with the Peace Corps.

"It sounds trite to say I wanted to make a difference, but that's why we went," Richie-Halford says. He was disillusioned with the military and wanted to try development work while pursuing a direction more in line with his wife's career than his own. The couple was in Iknioun, a village whose population swells to thousands on market days. They provided hygiene and AIDS education and helped build school bathrooms. To disperse their message further, the couple trained students to spread health information through skits and songs.

The experience led True to earn a master's degree in public administration with an emphasis on global health and nonprofit management. She now works with a Seattle nonprofit consulting firm. She and Richie-Halford bought a house last fall and recently welcomed their first child, Ida Grace True.

The DOE CSGF, he says, "afforded me an opportunity to explore some things I wouldn't have otherwise," including computer science and applied mathematics courses and computational science research outside his field.

GRASPING A GAS

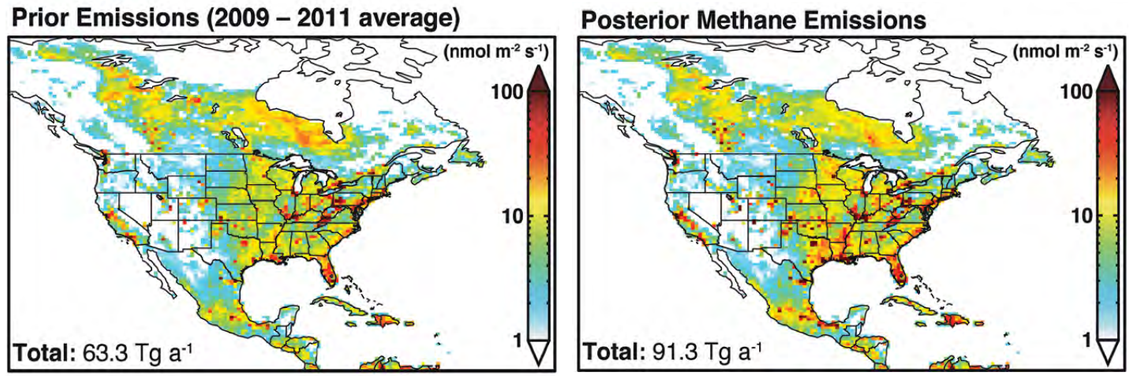
Methane is a leading greenhouse gas, but it's difficult to track its origins. Fellow Alexander Turner takes an inverse approach, using current levels to generate past source scenarios.

By Andy Boyles

Atmospheric scientist Alexander Turner first skied at age 3, making a wobbly beginning of a life outdoors. As he grew and developed his skills, he traded resort slopes for backcountry trails. In warmer seasons, he backpacked and rock-climbed.

Turner stayed active outside as an engineering student at his hometown University of Colorado at Boulder. There, climate change cast its shadow on the Rocky Mountain foothills. Pine beetles, thriving in the warming temperatures, were killing trees.

Since his sophomore year, Turner has worked with atmospheric and climate scientists to bring computational power to their research. Today, as a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient at Harvard University, he has followed his own trail in inverse modeling, building codes that turn around existing representations. Instead of a simulation that predicts events, inverse models use current observations to generate past scenarios. Using statistical methods, the atmospheric models identify emission sources that most likely led to the observations. As Turner puts it, "it's like measuring the depth of a river and trying to figure out where it rained."



A bottom-up estimate of average annual North American methane emissions for 2009-2011 (prior, left) and average annual emissions for the same period as estimated from satellite data (posterior, right). Total average annual emissions are 63.3 teragrams in the bottom-up estimate. The researchers' estimate is more than 20 teragrams per year higher, at 91.3 teragrams annually. Colors indicate where methane emissions are highest, as measured in nanomoles (billionths of a mole) per square meter per second. Red indicates highest emissions at 100 or more nanomoles per square meter per second, while blue and white are the lowest. *Credit: From Turner et al., Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data,* Atmos. Chem. Phys., *15, 7049-7069, 2015.*

With atmospheric chemist Daniel Jacob as his advisor, Turner has focused mainly on methane, the most abundant greenhouse gas after water vapor and carbon dioxide. Jacob wants to understand the main sources of atmospheric methane, including livestock, fossil fuels, landfills, wastewater treatment and rice cultivation, as well as the main natural source, wetlands. In Jacob's lab, Turner has mapped emission sources in greater detail than ever. Besides methane, he has contributed to research on carbon dioxide, through practicums at Lawrence Berkeley National Laboratory. (See sidebar, "Sensitive Model.")

Like other Bayesian atmospheric models, Turner's code superimposes patchy data onto a complete grid of prior data derived bottom-up from emissions sources and best-available assumptions. His method exploits the fact that similar regions host similar underlying processes. A smattering of observational data points in a large wetland, for example, can carry great weight and represent a large portion of the wetland. If the data are dense enough, a large wetland may be represented as more than one feature. "It's a balance between using a super-fine grid and a coarser grid that represents just regional-scale processes," Turner says.

Turner and Jacob worked with 22 collaborators to develop a high-resolution estimate of global methane emissions. For a 31-month period from mid-2009 through 2011, they found that sources worldwide had released methane at a rate of 593 million tons per year.

The team's results add to growing evidence that enormous amounts of methane escape notice in the Environmental Protection Agency (EPA) estimate of continental U.S. emissions. That bottom-up inventory is a tally of estimated discharges from known methane sources, and it pegs annual human-caused releases at 27 million tons per year. Turner's model indicates emissions were at least 60 percent higher, between 44 million and 47 million tons.

With a resolution of 50 km by 50 km, Turner's model confirmed that U.S. methane emissions mainly come from south-central states that are highly active in raising livestock and extracting and processing fossil fuels. Sources popped out in unprecedented detail – hot spots in that region as well as in California's Central Valley and the Four Corners area. The group concluded that livestock accounts for most of the methane emissions attributable to human-directed activity, followed by oil and gas, landfills and wastewater, and finally coal.

Turner and Jacob worked with six other researchers to measure methane emissions from 2002 through 2014. They combined data from surface, aircraft and satellite instruments and subtracted background methane levels (over oceans and upwind) from totals measured over land to find overall emission levels. They found a 30 percent increase for the 12-year period, potentially making the U.S. responsible for 30 to 60 percent of the past decade's global methane increase. The result contradicts EPA's bottom-up inventory, which holds that U.S. emissions were essentially flat.

Sensitive Model

As U.S. cities come to life each day, they trigger a carbon dioxide (CO_2) outflow that overshadows even power plant emissions.

To get a handle on urban CO_2 emissions, Ronald Cohen, a chemistry and earth and planetary sciences professor at the University of California, Berkeley, leads a research team that's building a CO_2 sensor network in the San Francisco Bay Area. The network will measure carbon dioxide emissions at a resolution high enough to attribute them to specific neighborhoods and to point and line sources like factories and highways.

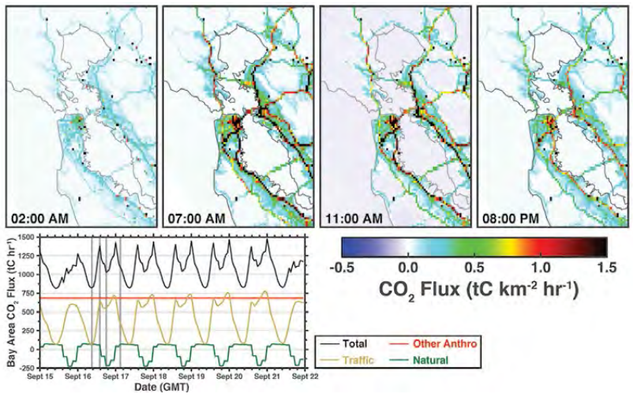
But how many sensors will such a network need to be reliable, yet affordable? And how precise, and therefore costly, do the sensors have to be?

Harvard University doctoral student Alexander Turner worked with Cohen during two DOE CSGF practicums at Lawrence Berkeley National Laboratory (where Cohen also holds a research post) to develop an inverse model that balances the trade-offs between the number of sensors and their precision.

Turner's code coupled a weather forecasting model with one designed to simulate the dispersion of gases and airborne particles. He wasn't the first to wed the models into a new one, but his version yields data at a one-square-kilometer scale. "We are using it at a much higher resolution than a lot of groups typically do because we're considering urban sources," Turner says. "There's a lot of variability on small scales."

He wrote the code from scratch in three months. Then, using Cori, a Cray XC40 supercomputer at the National Energy Research Scientific Computing Center, and university computing clusters, he simulated the performance of various hypothetical sensor networks. A web comprising many moderate-precision sensors worked best. To follow emissions from small sources within a city, "you don't need to have very precise measurements," Turner says.

Cohen praises Turner's contribution. "He built an inverse model that allowed us to build a better emissions inventory. And he did that at a speed that was extraordinary."



September 2013 carbon dioxide fluxes from bottom–up inventories. The top row shows fluxes in the Bay Area at four representative hours in local time. The panel at right shows the atemporal EDGAR (Emissions Database for Global Atmospheric Research) carbon dioxide flux. Bottom: the total Bay Area carbon dioxide flux (black), auto traffic (orange), other anthropogenic sources (red), and natural sources (green). Vertical gray shading indicates the time slices plotted in the top panels. *Credit: From Turner et al., Network design for quantifying urban CO_2 emissions: assessing trade-offs between precision and network density.* Atmos. Chem. Phys., *16, 13465–13475, 2016.*

The work on global and U.S. emissions led Turner to a long-standing mystery: why the global methane burden increased steadily through the 1980s and 1990s, then was inexplicably flat from 1999 to 2007, when levels began rising again for no obvious reason.

Researchers have offered a range of explanations without consensus. Turner and Jacob decided to consider a novel possibility: fluctuations of a compound that degrades methane – the hydroxyl radical (OH). This oxidant forms when ultraviolet light interacts with ozone, but once developed, it reacts so quickly with nearby compounds that researchers can't measure it directly. They can track OH indirectly by measuring one of its targets, methyl chloroform, but have assumed it's been constant for about 35 years.

Turner, Jacob and two collaborators ran an inverse model covering the decades leading up to and through the renewed growth. This time, they included methyl chloroform to infer OH levels. They didn't have enough data to rule out all proposed explanations, but OH emerged as the probable cause. "It looks like OH increased in the early 2000s, which caused the plateau, and then decreased by about 7 percent since 2003 or so," Turner says.

A paper on the project has been accepted by *Proceedings of the National Academy of Sciences* – just the latest trail marker on Turner's journey to open more backcountry routes in atmospheric science.

HITTING THE ACCELERATOR

Sluggish computational chemistry models prompted David Ozog to switch fields in search of improved performance.

By Thomas R. O'Donnell

As a chemistry graduate student, David Ozog grasped what he didn't know. Then he took steps to learn. Ozog, double-majoring in applied math and physics at Whitman College in southeastern Washington, first dipped into chemistry during an undergraduate research program at Marina Guenza's University of Oregon laboratory. She introduced Ozog to computer simulations of proteins and the theory of polymer dynamics – how lengthy chains of similar atomic units behave. It led Ozog, now a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient, to pursue a chemistry doctorate at Oregon.

"There's great power and opportunity in being able to use computer models to simulate chemical systems," he says. "Our goal was to design new materials, pharmaceuticals and chemical processes." But there was a problem. "These simulations were incredibly computationally expensive," sometimes taking months.

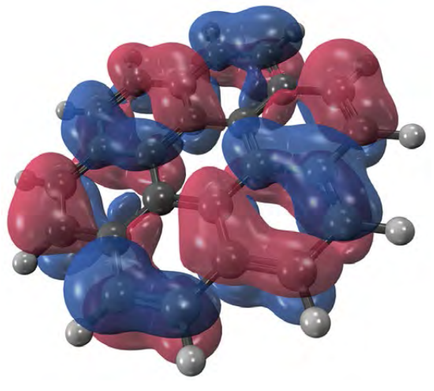
Ozog knew of ways to accelerate such computations, including parallel processing, which breaks problems into pieces and sends each to a different processor for simultaneous solution. High-performance computing (HPC) system designers also have adopted graphics processing units (GPUs), relatives to video game chips, to accelerate calculations.

"I decided there was too much computer science I didn't know" to address the issue, Ozog says. After earning a chemistry master's degree, he left the university to work in programming,

A 1,000-molecule water system simulation with a quantum mechanics/molecular mechanics (QM/MM) approach. The enlarged water molecule in the middle is modeled with QM; all other molecules are modeled with MM. *Credit: David Ozog.*

creating iPhone applications and websites and mastering the Python language. A year later, he returned to Oregon's graduate school, this time in computer science. His goal: Make computational chemistry codes run more efficiently, enabling bigger and more precise simulations. "It's like hunting for speedups. I'm all about taking the original code – thinking about its algorithms and how it maps onto the hardware – and looking for optimizations, ways to speed up the execution time."

Ozog's doctoral advisor, Computer and Information Science Professor Allen Malony, says his student's background is perfect for the job. "I saw in David somebody who could bring background knowledge from (chemistry) to look at computational chemistry as a high-performance computing problem" while focusing on new technology like GPUs. "Once he was here, we continued to emphasize that having connections to chemistry was important."



A graphene molecule from one of David Ozog's Hartree-Fock quantum mechanical calculations. The blue and red blobs show the lowest unoccupied molecular orbital. *Credit: David Ozog.*

One of Ozog's Oregon mentors, Sameer Shende, also recognized Ozog's unusual skill combination. Shende suggested an internship at Argonne National Laboratory, where former DOE CSGF recipient Jeff Hammond was working with NWChem, DOE's leading computational chemistry code.

NWChem runs on massively parallel computers and calculates the effects of quantum physics, which governs particles at the tiniest scales, when simulating atomic and molecular interactions. But parts of the code can overwork some processors while leaving others idle, a problem known as load balancing.

In his 2012 Argonne internship, Ozog focused on NWChem's tensor contraction engine (TCE), part of its coupled cluster code describing electron interactions. Tensors characterize relationships between data in multiple dimensions. (A matrix is a two-dimensional tensor.) Tensor contraction sums the products of tensor components over one or more indices to reduce the answer's dimensions.

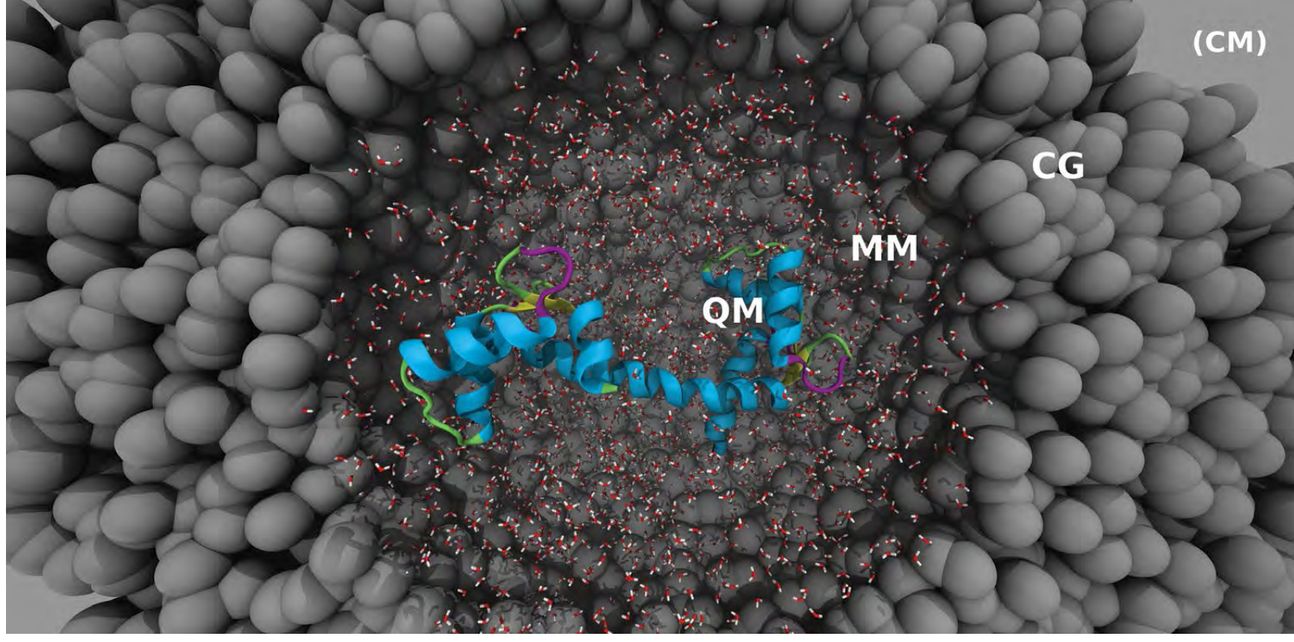
The TCE distributes calculations across thousands or even millions of processor cores, Ozog says. "If one processor wants to do a particular task, it might have to get data from two different processors. It might even have to get data from four, eight or more," slowing calculations.

NWChem also used a centralized approach: Every processor ready to do work interacted with a single core in memory to determine which tasks needed completion.

After studying how data moved and processors contended for access to this processor, Ozog and Hammond developed an "inspector and executor" approach. The inspector algorithm uses simple models to estimate how long tensor contractions take for particular data. The executor algorithm assigns tasks to processors accordingly. "It's not going to be perfect, but it improves the situation of just relying on one processor to assign tasks and to keep (processors) busy," Ozog adds. The new approach often made the code run significantly faster.

The summer internship went so well it was extended through the end of the year. Ozog later was chosen for the DOE CSGF based partly on his work and Hammond's recommendation.

In more recent research with Malony, Ozog automated operations in a coarse-graining code, which makes atomic-detail simulations easier to solve by treating several atoms as a unit rather than individually. The model omits some details but is valid for calculating things like a material's thermodynamic properties.



An example of a quantum mechanics/molecular mechanics/coarse grain (QM/MM/CG) system in which the calmodulin protein is modeled with QM, a water layer with MM, and the water solvent with CG. The outer region also is possibly modeled with continuum mechanics. *Credit: David Ozog.*

A coarse-grained representation could capture those exactly with greater computational efficiency than a detailed model, Ozog says.

Suppose, however, that someone wanted to return to a higher-resolution model. "Can you reconstruct that original representation without messing things up?" Ozog asks. In an April 2015 paper published in the *Journal of Computational Science*, he and his collaborators presented a technique to manage this arduous transition.

Say a researcher is modeling a polymer using a standard molecular dynamics code like LAMMPS from Sandia National Laboratories. Such models omit quantum mechanical properties yet demand computational power. "We have codes that efficiently coarse-grain" the system being modeled, but "these codes don't exist in LAMMPS," Ozog says. A researcher must manually implement the transitions from atomistic to coarse-grained and back again – steps that could be required through multiple iterations.

With Ozog's workflow system, programmers can specify how they want the transition to happen. "Instead of waiting for the LAMMPS simulation to finish and then running a separate code, it'll just happen automatically," Ozog says. The algorithm also ensures that the transition is done correctly. The method helped accelerate a polymer model on Edison, a Cray XC30 at the National Energy Research Scientific Computing Center.

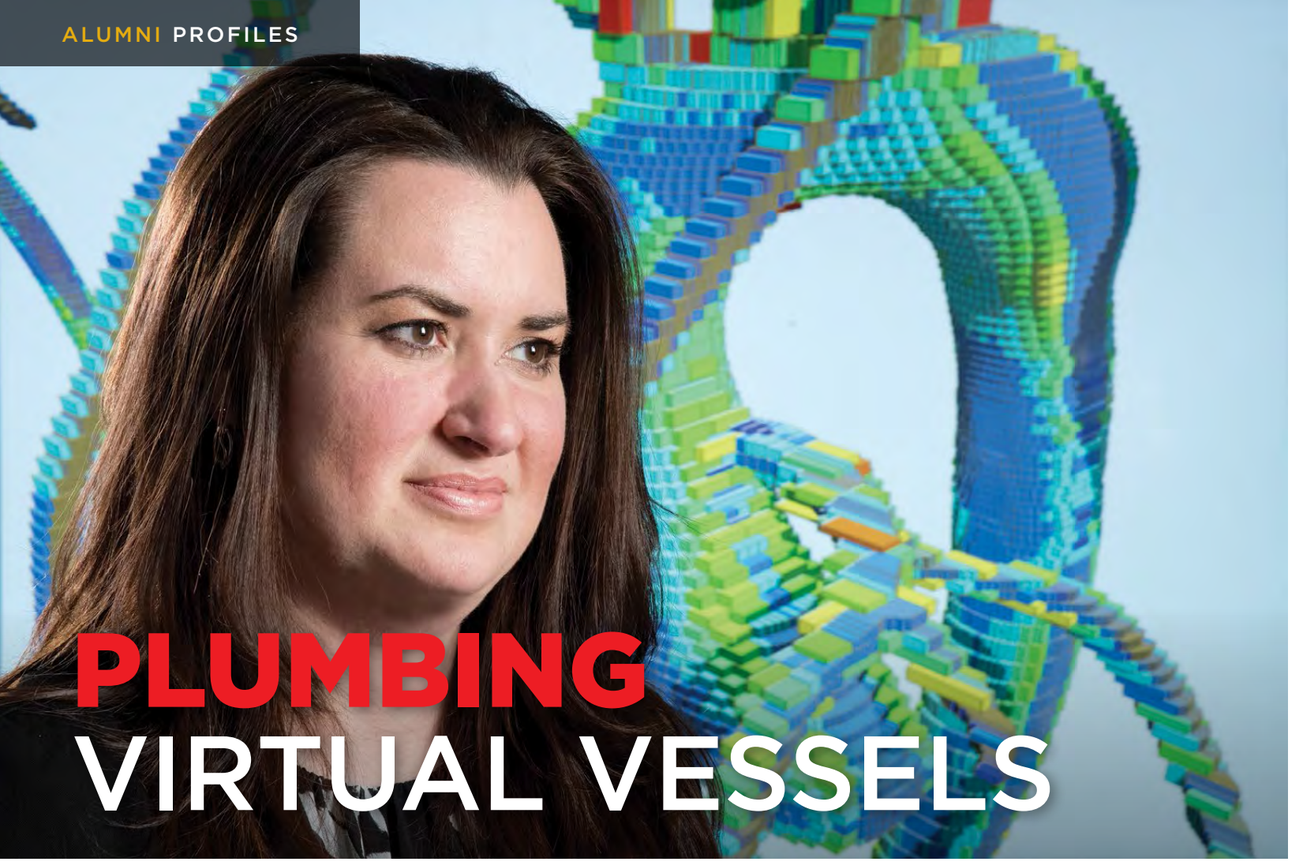
Ozog's dissertation focuses on these scientific workflow problems, such as coupling quantum codes like NWChem with molecular dynamics codes so they can share data or managing interactions between codes written in different languages. Linking the components is complicated, especially since each could address the problem at different resolutions in time and space, Malony says. "Trying to build these coupled systems that support multiresolution and workflows is really leading edge." The task will become even more complex when exascale computers, hundreds of times more powerful than today's systems, come on line, he says.

Significantly, Ozog is collaborating on much of the work with Guenza's group, bringing him full circle in theoretical chemistry.

Ozog's career may have been inevitable, given his background. Growing up in Colorado and Montana, he learned about chemistry from his paternal grandfather, a chemistry professor at Denver's Regis University, and computers from his maternal grandfather, an electrical engineer.

After finishing his dissertation in December, Ozog moved to Massachusetts to start work with Intel Federal LLC, an Intel Corp. subsidiary that supports HPC at DOE and other government agencies. He'll optimize applications and design future processor hardware.

Ozog's background, with one foot in computational science and another in any one of many subjects, will serve him well, Malony says. "What he's done transcends just computational chemistry."



PLUMBING VIRTUAL VESSELS

The circulatory simulations Amanda Randles pioneers at Duke University could help physicians choose the best treatments before operating.

by Karyn Hede

When a patient shows signs that a major blood vessel wall is weakening, a physician must weigh the risk of surgery against that of the vessel wall's rupturing catastrophically. That clinical decision is based on experience, images of the vessels and the patient's other risk factors. But what if a physician could experiment with different surgical options before operating?

Computational modeling of the human circulatory system has arrived, and physicians are eager to test it at Duke University, where Amanda Randles is an assistant professor of biomedical engineering. Randles, a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) alumna, has gained global acclaim for her circulation codes, earning two nominations for the prestigious Gordon Bell Prize for high-performance computing (HPC). Her latest, in 2015, was for HARVEY, a program named for the 16th Century English physician who first described blood circulation. It modeled a fully functioning human arterial system.

Randles started HARVEY as a Harvard University graduate student. With skills gained while programming IBM's Blue Gene supercomputer for four years, she modeled patient-specific coronary artery geometries on 294,912 processors.

She says HARVEY might not exist without support from the fellowship, which "allowed me the freedom to pursue a line of research tangential to the main focus of my advisor's lab."

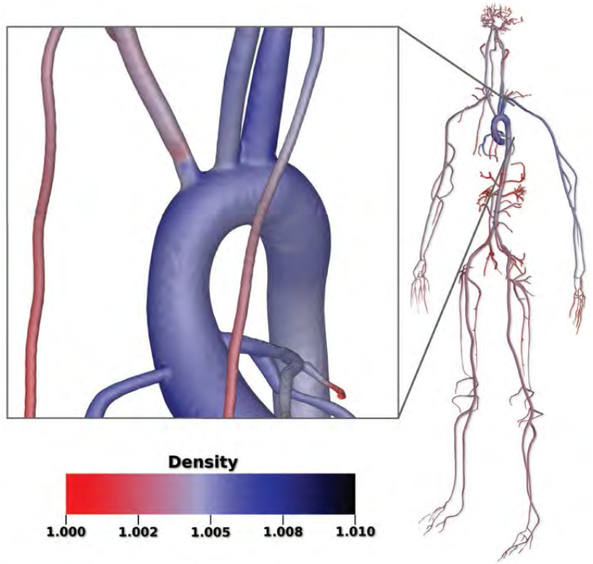
Now entering her third year at Duke, Randles has set even more ambitious goals, including modeling the capillary beds that deliver oxygen to tissue and incorporating veins, which return oxygen-depleted blood to the heart. That will likely require more computational power than available on existing HPC systems, so Randles is driving the computer science side of her research.

Several group members are working to port HARVEY to next-generation computing architectures.

These computational challenges serve a larger goal in Randles' research. As a Duke undergraduate, she majored in both physics and computer science but was drawn to biological science.

"I've always wanted to work on medical problems that ultimately could help people," she says. "Every project we're working on has a clinician on it, just to make sure we're asking the right questions."

Her research uses computational science to evaluate approaches in ways that would be difficult or unethical to try in people. Take the question of whether to operate on a patient developing an aortic aneurysm, a weakening in the central vessel delivering blood from the heart. A physician might make different decisions about whether to line the vessel wall with an artificial support if he or she could visualize areas of weakness and model the potential risk.



Vessels larger than 1-mm wide for the entire human arterial system. The results are from a simulation at 200-micron resolution on the Vulcan supercomputer at Lawrence Livermore National Laboratory (LLNL). Inset: density along the upper aortic arch. *Credit: Liam Krauss, LLNL.*

Randles wondered if by changing a diseased vessel's geometry "to represent different treatment options before the doctor ever goes into the operating room, can we see how it would affect the patient under different physiological conditions?"

That scenario is playing out in a project to engage cardiologists in simulated case studies. In Randles' lab, an interactive model lets physicians see the diseased artery and where HARVEY predicts the vessel wall might undergo low or oscillatory shear stress, leading to dangerous plaque accumulation. The research team wants to know whether that information changes the physician's decision of where to place a surgical mesh to support the vessel.

To complement research on forces affecting arterial blood flow, Randles is modeling how individual cells twist, fold and deform as they flow through vessels. With support from a National Institutes of Health Early Independence award, she's using computing to study the transport of individual cancer cells in the bloodstream.

"We're moving from the large scale to the small scale," Randles says.

The DOE CSGF, Randles says, made her aware of the Lawrence Fellowship, a program that finances postdoctoral researchers at Lawrence Livermore National Laboratory (LLNL). She won the award in 2013 and began collaborating with LLNL scientists Erik Draeger, Maxim Shusteff, Ted Laurence, Monica Moya and Sonny Ly to model rigid vessel wall geometries compared with three-dimensional physical models.

'I've always wanted to work on medical problems that ultimately could help people.'

The team continues collaborating. With support from two laboratory-directed research and development awards, they're studying fluid flow inside networks of bioprinted tissue – "3-D printed vasculatures made of clear polyurethane to mimic the geometry," she says. LLNL scientists are assembling sophisticated capillary networks of endothelial cells to provide real-world test beds for validating results Randles obtained running HARVEY on LLNL's Vulcan HPC system.

Randles is at home working with chips and processors but isn't afraid – well not too afraid – to dive into the medical side.

"I've been in the cardiac catheterization lab," she says. "I've been invited to the operating room, but I'm afraid I might get queasy."

MESSY MATERIALS

At Lawrence Livermore National Laboratory, Brandon Wood probes material properties that could help make the hydrogen economy a reality.

By Monte Basgall

Materials scientist Brandon Wood obsesses over hydrogen's paradoxical properties. With only one proton and electron, it's the universe's most abundant element, notes the 2007 Department of Energy Computational Science Graduate Fellowship (DOE CSGF) alumnus. "But to use it, say, in a fuel cell, it needs to be in a pure form. And that's challenging, because it's usually chemically bound to something else. So it needs to be extracted."

At Lawrence Livermore National Laboratory's Quantum Simulations Group, he approaches such conundrums via high-performance computing (HPC). His group pushes past the usual, unrealistically simple materials models, seeking "beyond ideal" traits instead. He introduces messiness – "defects and interfaces, where atomic arrangements are disturbed," just as in real-world materials. Those "oftentimes can have huge and profound consequences for the applications I'm interested in."

As part of this mold-breaking, Wood splices what he calls the two major camps of materials modelers: the atomistic, who explicitly simulate atoms and molecules; and the continuum crowd, who prefer to simulate fields, gradients and concentrations. He calls the mashup of these approaches "mesoscale modeling."

Wood uses two atomistic codes: Q box, developed at Livermore and the University of California, Davis; and Quantum ESPRESSO, developed in Europe. On the continuum side, he uses codes developed and adapted by his group and by others at Livermore.

Wood's interests span a range of developing technologies for futuristic transportation, each with major materials science challenges. President George W. Bush jumpstarted such research with his 2003 hydrogen initiative to develop fuel-cell vehicles powered with electricity generated in a chemical reaction between hydrogen and oxygen and releasing mainly water.

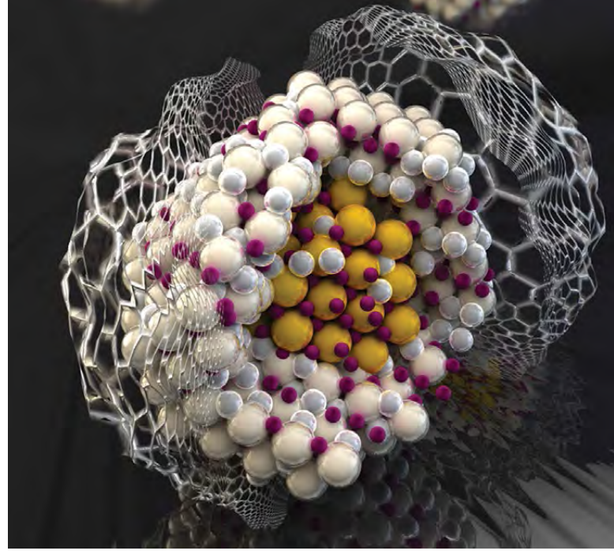
The hydrogen economy idea was initially oversold, Wood says, for want of adequate ways to produce, store and dispense hydrogen as a fuel. This, plus other political and economic factors, led auto companies to switch to battery electric propulsion. Now that battery technology is maturing, companies are reinvesting in hydrogen "but more as a longer-term strategy," Wood says.

His group's models are finding their way into all these realms. In 2010 Wood became principal investigator of a DOE Energy Efficiency and Renewable Energy-funded search for materials that can compactly store and deliver hydrogen inside special compounds without expensive chilling and pressurization. This led to his involvement in DOE's new Hydrogen Storage Materials – Advanced Research Consortium, for which he now directs theory activities across multiple national laboratories.

Using more than 80 million processor hours on Livermore's Sierra supercomputer and its more powerful replacement, called Quartz, his group explores the storage prospects of candidates such as magnesium borohydrate, a "very difficult" and "extremely complex" material. "That's actually a good thing," Wood adds, because trying to engineer such complexity "gives you more knobs to turn. The question is which knobs."

Based on experience, he's introducing kinetics – reaction rates – into models because fuel cell vehicles must be filled with fresh hydrogen relatively quickly. Reaction rates were too slow in earlier, rejected storage materials.

In 2013, Wood also received 60 million processor hours on Oak Ridge National Laboratory's Titan supercomputer under DOE's Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program to investigate solid-state materials for use as electrolytes in lithium ion batteries.

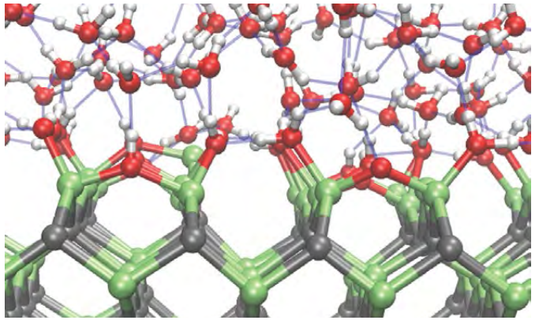


This model shows an indium phosphide semiconductor in contact with water. The material is a candidate to split water molecules with sunlight for hydrogen production. *Credit: Brandon Wood.*

Solids would be more stable and thus safer as electrolytes than the usual liquids, but moving ions through a solid is much harder, he says. After computationally reviewing thousands of candidates, his group found some potentially interesting materials. More important, it found new ways of forecasting their potential for shuttling ions.

Wood's group also is part of multi-lab efforts to find energy-efficient and productive ways to generate hydrogen by splitting water.

Trying to engineer such complexity 'gives you more knobs to turn. The question is which knobs.'



In this model, lithium nitride nanoparticles confined in a carbon matrix are charged with hydrogen atoms. The tiny beads are a possible hydrogen storage medium. *Credit: Brandon Wood.*

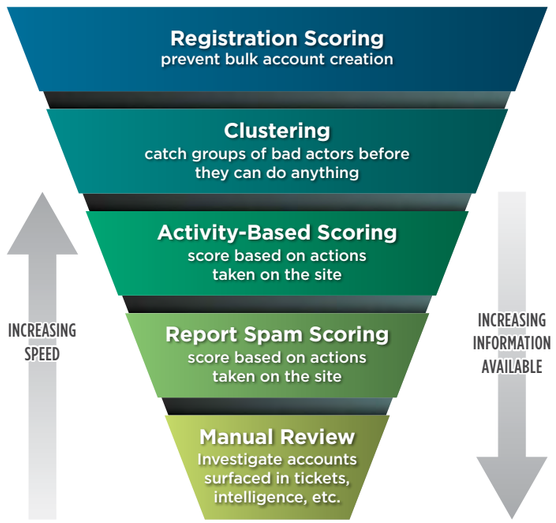
Wood received a bachelor's degree in physics at Stanford in 2001, but switched to materials science and engineering for his 2007 MIT Ph.D. because "I wanted something that was more tangibly connected to our everyday experience."

Doctoral training can be confining, he says. "You basically know the one thing you become the world's expert in." He credits the DOE CSGF for saving him from that bubble by exposing him to a spectrum of research and to the HPC community.

Not least, he met people at Lawrence Livermore, leading to his postdoctoral post there. "I can say honestly that I wouldn't be here today if it wasn't for the fellowship."

SCAM FILTER

At LinkedIn, alumna Jenelle Bray applies data techniques to block would-be fraudsters, scammers and scrapers.



Methods to block fake LinkedIn accounts make tradeoffs between how fast a decision can be made to accept a member and the amount of information available to make a ruling. *Credit: Jenelle Bray.*

By Thomas R. O'Donnell

Jenelle Bray's career has moved from predicting protein behavior to predicting human behavior. As a California Institute of Technology student and Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient from 2006 to 2009, Bray developed algorithms foretelling the structure of cell membrane proteins. As a Stanford University postdoctoral researcher, she used machine learning – teaching computers to act without explicit instructions – to predict where molecules would bind to proteins.

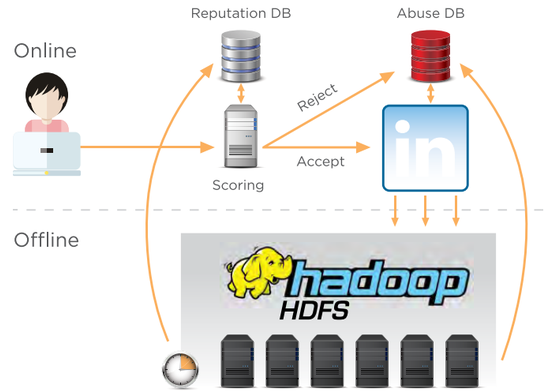
Now Bray applies similar techniques to stop misbehavior on LinkedIn, the career-oriented social media site. Like her, many of those on her team are trained in biology, physics and other fields outside computer science. They're well-suited for the job, says Bray, an engineering manager for LinkedIn's Anti-Abuse Relevance Group, because experimental results often contain multiple errors and omissions, especially in biology. "People who have dealt with that are very good at dealing with any kind of data."

At LinkedIn's Sunnyvale, California, headquarters, Bray's team seeks ways to predict when a user is likely to harm the site and its members. The group wants to block fraud, halt spam and stop data scraping, in which automated programs pull information from user profiles – information that LinkedIn charges recruiters and others to access as part of its premium product offerings. Scrapers could undercut LinkedIn with the data and their theft also can erode member trust, Bray says.

Scammers use software robots to open fake LinkedIn accounts. These phony members establish connections with as many LinkedIn members as possible, then gather their data.

To stop these ersatz users, LinkedIn workers formerly watched for unusual activity spikes, like a surge in new accounts, Bray says. They then would examine the data for common characteristics, like a single IP address, similar email addresses or identical country of origin, that may indicate illegitimate activity. Her team "wanted to automate that and say, if we group on IP and day they're registered, and we get groups of accounts, can we create a model to look at which groups are good and which are bad?" Bray says. The catch is that similar activity could be legitimate, such as groups of students creating profiles in a job-hunting exercise.

With a computer cluster running Hadoop data analysis software, the LinkedIn team feeds its machine-learning algorithms a diet of data: examples of real and fake profiles and how they behave. The algorithm learns identifiers for each and predicts whether a new profile is fake.



LinkedIn has developed an anti-abuse infrastructure with online and offline systems. Offline, a Hadoop cluster uses machine-learning algorithms to analyze data on real and fake profiles and learn identifiers for each. That's fed to the online system, which scores requests against characteristics of good and bad accounts. The online system takes action against the bad accounts or requests more information to make better decisions. This information is fed back into the offline models. *Credit: Jenelle Bray.*

Some of the models Bray's team has created run in real time. Others – especially ones that flag fake accounts – are so computationally demanding they run only offline. "It would be a horrible member experience if every time you requested a page we'd say you need to wait five minutes until we decide if you're good or bad," Bray says. But running offline means malicious users can cause damage before they're removed. "A big push for us now is to figure out how to get more things on line or at least what we call near line: We wouldn't block your request, but hopefully within a few seconds we could decide" whether it's legitimate.

> ***'We want to make their job so hard and so expensive that maybe they want to give up.'***

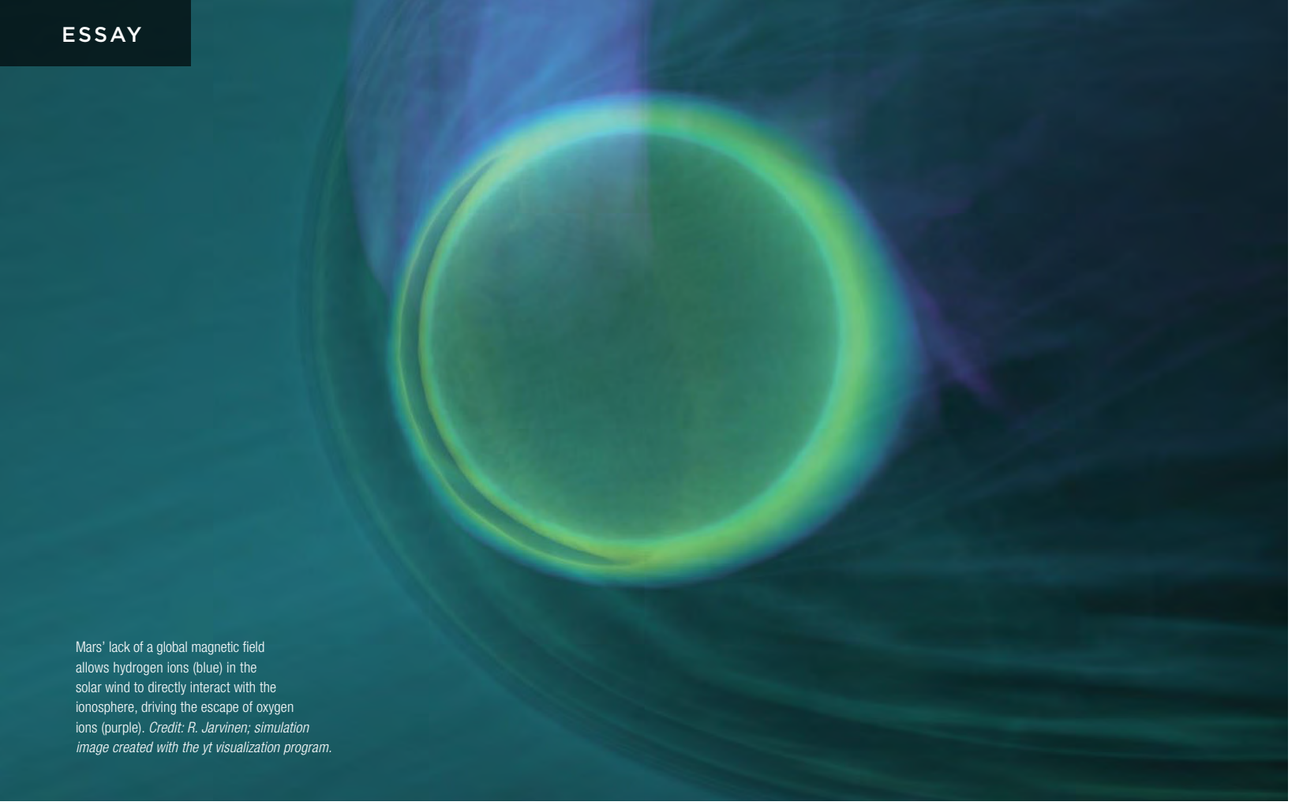
But, Bray says, "you can't just feed in all the data. The fun part is figuring out what signals would show these accounts are good or bad, and then writing the code to collect those."

If a new registration looks suspicious to its algorithms, LinkedIn could require extra steps, like entering a CAPTCHA code to prove the user isn't a robot. Existing profiles identified as fake could be removed entirely, with provisions for appeal in case of false positives.

Bray says she and her team know they can never completely stop scammers. "We want to make their job so hard and so expensive that maybe they want to give up." That's tough, because the hackers quickly find ways around new safeguards, and with more than 467 million users, it's demanding to classify every action. "Speed and scale are big technical challenges."

Technical challenges like these drew Bray away from the academic career she always thought she'd pursue.

"Working against bad actors is interesting. There are always new technological challenges, and I work with a lot of smart people" to tackle them.



Mars' lack of a global magnetic field allows hydrogen ions (blue) in the solar wind to directly interact with the ionosphere, driving the escape of oxygen ions (purple). *Credit: R. Jarvinen; simulation image created with the yt visualization program.*

TO BOLDLY GO

(AND SURVIVE WHEN WE GET THERE)

By Hilary Egan

"Star Wars." "Avatar." "Interstellar." Aside from being fun movies, these all have something important in common: imagined planets outside of our solar system. From the deserts of Tatooine to the rainforest of Pandora, it's easy to imagine a future in which we can explore another world with the same ease as exploring another country.

Studying planets is exciting right now because we're constantly finding ones orbiting stars besides our sun. The closest star to us is Alpha Centauri and scientists just found three rocky planets around it. Before you pack your bags for the next spaceship heading out, however, we have to ask ourselves if the planet is able to support life or, as scientists say, whether it's habitable.

One key requirement for habitability is liquid water. This means the planet can't be too hot or too cold; otherwise all the water would boil off or turn to ice. If the liquid is on the surface, like at Earth, the planet must also have an atmosphere. Without one, the water would just evaporate into space.

Even within our solar system neighborhood, there's a big variation in planetary atmospheres. Venus is a hot, poisonous greenhouse and Mars is a dry wasteland with next to no atmosphere. In our vicinity, only Earth is the Goldilocks of planets: just right.

It's relatively easy to look at planets within our own solar system and figure out what their atmospheres are like. We can send satellites and rovers and take pictures from relatively nearby. Studying planets outside our solar system is much harder because they are so far away. The most common way astronomers discover them is by looking at shadows the planets make as they cross in front of their star during an orbit. This makes any direct detection of an atmosphere difficult, particularly for small, Earth-sized planets. I'm working on figuring out how we can tell if one of these exoplanets, as they're called, is potentially habitable in the absence of pictures, and specifically whether or not it can have an atmosphere.

Just because Earth is the only habitable planet nearby right now doesn't mean it was always this way. Planets can change dramatically over time. For example, the surface of Mars is covered in dried-up river beds and lake basins, indicating that it used to have liquid surface water and an atmosphere. Since this no longer is the case, we know Mars has significantly evolved in the past 4 billion years.

Wind made up of electrically charged particles from a star can slowly (or quickly) erode the atmosphere over the course of a planet's life. Keeping it is like trying to hold onto an armful of feathers while a strong fan is pointed at you. Feathers are light and easier to blow away than you are. Similarly, a star can blow away a planet's atmosphere while keeping the planet intact. Lots of factors affect how fast this atmospheric escape happens, from the star's proximity to the planet, to the planet's size, to whether the planet has a magnetic field.

Thinking about all these factors at once is difficult, so I run computer simulations to keep of track them – and more. To run a simulation, we start by setting up a planet with an atmosphere and then turning on a strong wind. As the simulation goes on, particles in the wind and the atmosphere will move around based on their interactions with the planet's gravity, each other, and the magnetic field. After the simulation runs for a while, we can count how many atmospheric particles have escaped the planet and use that rate to figure out how long it could maintain an atmosphere.

I'm most interested in exploring how magnetic field strength affects how fast the atmosphere escapes. Some planets, like Earth, have a strong magnetic field, while others, like Venus, have no magnetic field at all. Mars used to have a magnetic field, but it turned off billions of years ago. It can affect a planet's atmosphere because charged particles are forced to travel in the direction of a magnetic field, like little trains on a magnetic-field track. Both the intense wind from the star and the planet's atmosphere contain these charged particles.

One way a magnetic field affects atmospheric escape is by acting as a sort of giant shield to protect the planet's

atmosphere. For example, the magnetic field is pointed sideways on the sun-facing part of Earth. This means that as particles from the sun approach the Earth, the magnetic field forces them to go around. If the sun is a giant fan trying to blow away your feathers, a magnetic field can be like a giant umbrella you can use to defend yourself from the breeze.

On the other hand, at the poles of a planet like Earth, the magnetic field is pointed outward. This means that charged particles from Earth's atmosphere will have a direct path to escape, possibly increasing the rate at which the atmosphere leaves. This means there are two competing ways the magnetic field can affect atmospheric escape: one that helps and one that hurts.

By running computer simulations, we can evaluate when each of these effects is dominant. Then we can start to answer questions like "If we find a planet with no magnetic field, is it possible that it has an atmosphere?" And "Is there a minimum magnetic field that is required for keeping an atmosphere?" The answers to these questions are likely to be complicated and depend a lot on the size and location of the planet, as well as the star in question. But just like packing for our first trip to a far-off planet, it's better to find the answers before we leave.

The DOE CSGF stages the Communicate Your Science & Engineering contest to give fellows and alumni an opportunity to write about computation and computational science and engineering for a broad, non-technical audience. The author of this year's winning essay is a third-year fellow studying astrophysics at the University of Colorado.

CLASS OF 2017

Kathleen Alexander
Massachusetts Institute of Technology
Computational Materials Science
Advisor: Christopher Schuh
Practicum: Oak Ridge National Laboratory
Contact: katcalex@gmail.com

Nicholas Frontiere
University of Chicago
Physics
Advisor: David Reid
Practicum: Lawrence Livermore National Laboratory
Contact: nfrontiere@gmail.com

Chelsea Harris
University of California, Berkeley
Astrophysics
Advisor: Peter Nugent
Practicum: National Renewable Energy Laboratory
Contact: chelseaharris@berkeley.edu

Thomas Holoien
The Ohio State University
Astronomy
Advisor: Krzysztof Stanek
Practicum: SLAC National Accelerator Laboratory
Contact: tholoien@astronomy.ohio-state.edu

David Ozog
University of Oregon
Computer Science
Advisor: Allen Malony
Practicum: Argonne National Laboratory
Contact: ozog@cs.uoregon.edu

David Plotkin
University of Chicago
Earth Sciences
Advisor: Dorian Abbot
Practicum: Argonne National Laboratory
Contact: dplotkin@uchicago.edu

Daniel Rey
University of California, San Diego
Biophysics
Advisor: Henry Abarbanel
Practicum: Lawrence Livermore National Laboratory
Contact: nadrey@gmail.com

Adam Richie-Halford
University of Washington
Physics
Advisor: Aurel Bulgac
Practicum: Brookhaven National Laboratory
Contact: richford@uw.edu

Alexander Turner
Harvard University
Atmospheric Science
Advisor: Daniel Jacob
Practicum: Lawrence Berkeley National Laboratory
Contact: aturner@fas.harvard.edu

BY THE NUMBERS: DOE CSGF IMPACT

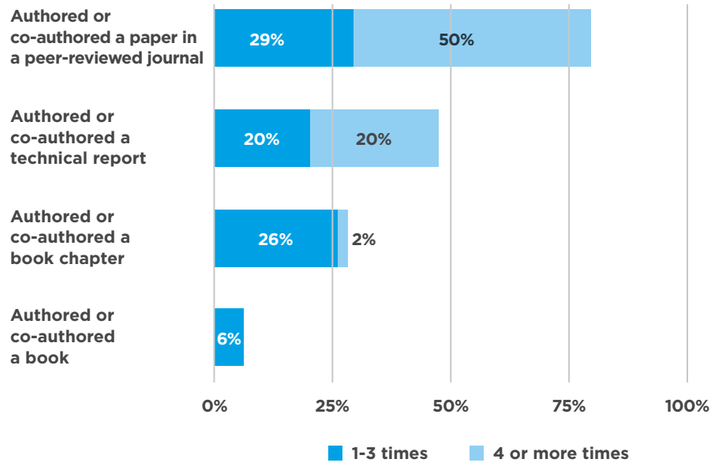
Now numbering more than 400, Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipients exert great influence in government laboratories, academic institutions and private companies. Alumni are developing new technologies and applying high-performance computing to important national and scientific problems. A recent longitudinal study quantified this impact, collecting survey responses from nearly 300 graduates and current fellows and compiling statistics from curricula vitae. It found the DOE CSGF corps has produced thousands of publications and presentations to advance science in myriad subjects. The following graphs provide further detail on the program's impact. For more information, go to: www.krellinst.org/csgf/about-doe-csgf/2017-longitudinal-study.

Number of publications alumni produced, by type

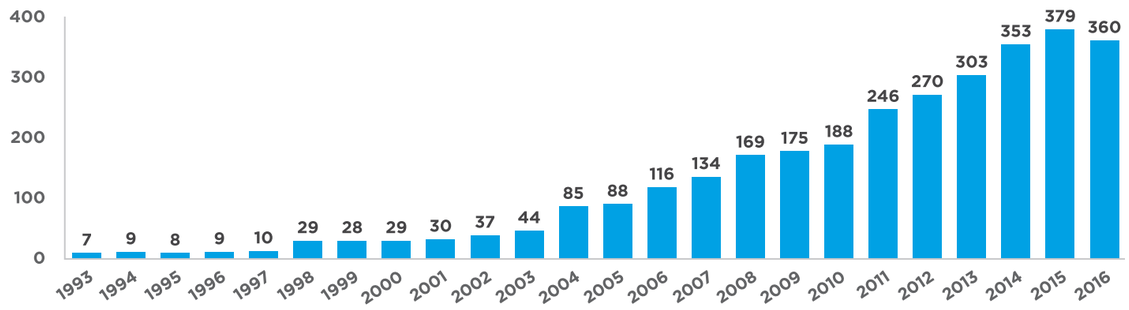
Product type	Unique products
Journal article	3,252
Conference presentations, papers, or proceedings	1,808
Presentation (e.g. invited talk, lecture, or keynote)[1]	981
Book chapter	112
Book	13
Other	206
Total	**6,372**

[1]Products which listed two or more alumni as authors were counted only once as unique products. While some alumni included multiple instances of presentations with the same title, the count of presentations refers to unique presentations listed in a given curriculum vitae. Unique presentations appeared in CVs between one and 41 times. Of the 981 presentations, 60 percent were listed three or fewer times.

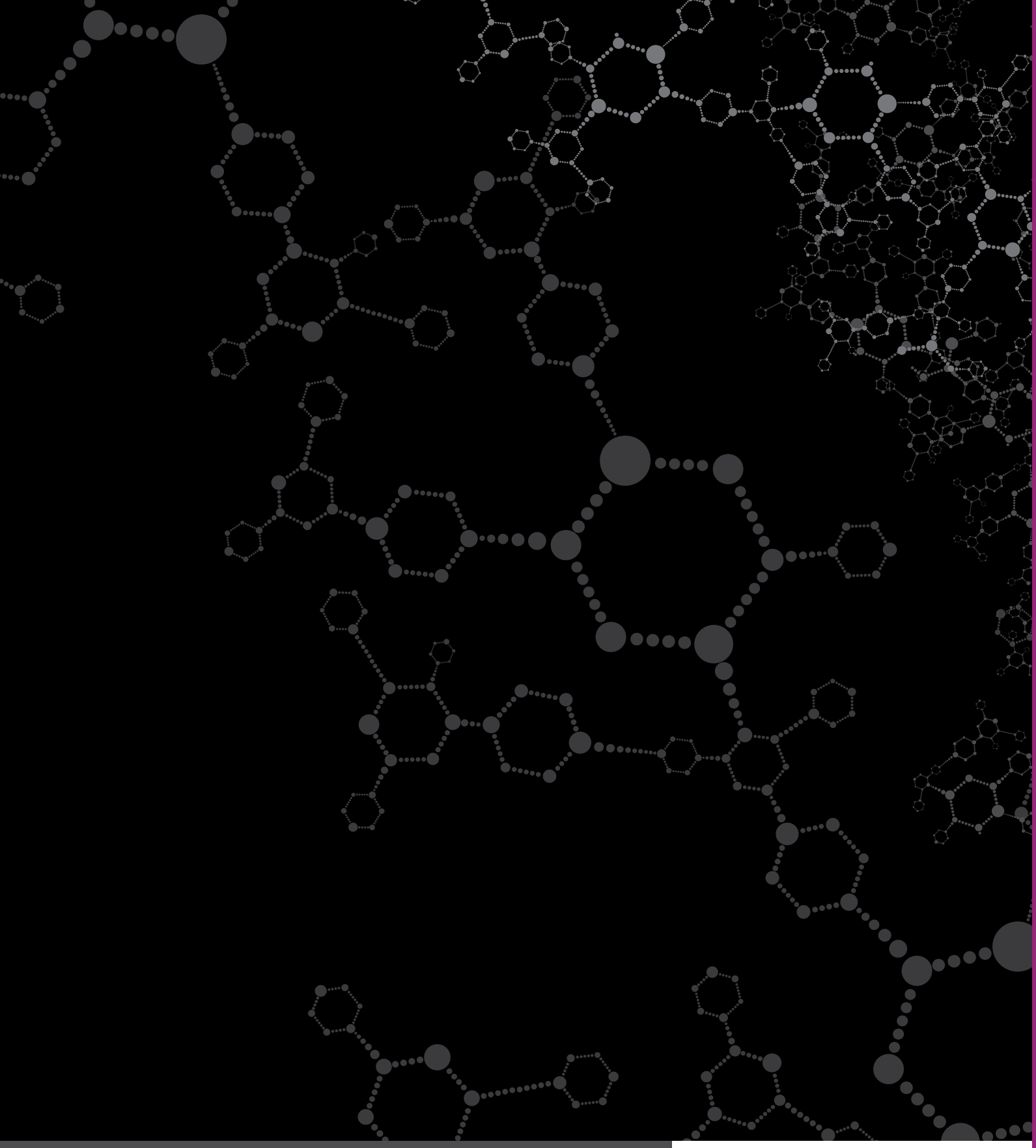
Percent of alumni reporting how often they had authored publications



Number of unique journal articles published by DOE CSGF alumni, by year



NOTE: Figure is limited to unique articles for which publication date information was listed in the CV. Data in this figure are based on information collected through 12/2/16.



The Krell Institute
1609 Golden Aspen Drive, Suite 101
Ames, IA 50010
(515) 956-3696
www.krellinst.org/csgf

Funded by the Department of Energy Office of Science
and the National Nuclear Security Administration.