THE DOE CSGF ANNUAL

DEIXIS

2014

DEPARTMENT OF ENERGY COMPUTATIONAL SCIENCE GRADUATE FELLOWSHIP

SUGARY SUCCESS

Heather Mayes Tracks Molecular Twists to Advance Biomass-based Fuels Research

PAGE 5

THE DOE CSGF ANNUAL

DEIXIS

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 contract DE-FG02-97ER25308.

For additional information about the DOE CSGF program, the Krell Institute or topics covered in this publication, please contact:

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

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.

The DOE CSGF is funded by the Department of Energy Office of Science and the National Nuclear Security Administration.

ANNUAL PROGRAM REVIEW

The Annual Program Review – a gathering of fellows, alumni and national laboratory staff to share research and make contacts – is a unique benefit of the Department of Energy Computational Science Graduate Fellowship (DOE CSGF). A summary of this year's review begins on page 30.

Editor
Shelly Olsan

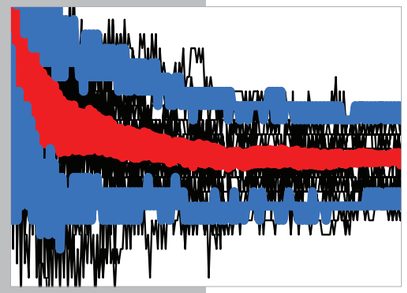
Senior Science Writer
Thomas R. O'Donnell

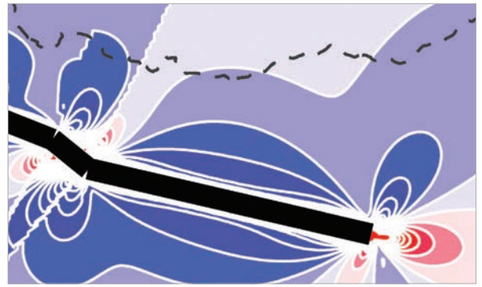
Copy Editor
Ron Winther

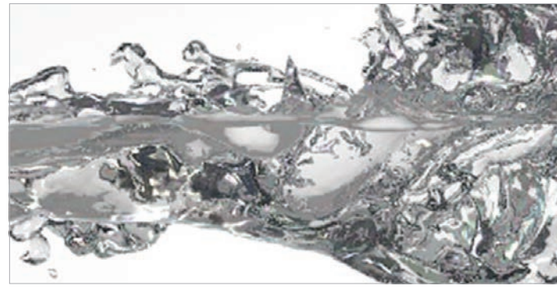
Design Coordinator
Buffy Clatt

Design
julsdesign, inc.

TABLE OF CONTENTS







THE SIGNIFICANT SABBATICAL

PRACTICUMS PROVIDE INSIGHT AND PRODUCE ADVANCES

Left to right: Miles Lopes, Jason Bender, Heather Mayes and Phoebe DeVries

EVERY YEAR, usually in summer, a subset of Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipients fans out across the country, headed to practicums at national laboratories.

The fellows test their skills at new research challenges that may be somewhat related to their doctoral projects but frequently are completely different. The students gain new perspectives while experiencing national lab life and making career connections.

The student-lab advisor collaborations often produce significant advances. For example, Miles Lopes' statistical training provided insights into a significant problem in machine learning at Sandia National Laboratories in California. And Heather Mayes' project at Colorado's National Renewable Energy Laboratory produced a kind of map for shape-shifting sugars.

At Lawrence Berkeley National Laboratory, Phoebe DeVries developed a new code to analyze the stresses underlying earthquakes. Meanwhile, Jason Bender stepped outside aerospace engineering to tackle a problem in computational chemistry at Argonne National Laboratory near Chicago.

Practicums may end, but the work often continues as fellows and their lab advisors continue collaborating. For both, short sabbaticals can have long-lasting effects.

The Department of Energy Computational Science Graduate Fellowship supports the nation's brightest science and engineering students, allowing them to concentrate on learning and research. The work of more than 325 DOE CSGF alumni has helped the United States remain competitive in a global economy.

~~~~~

Credit: National Renewable Energy Laboratory

THE SWEET SCIENCE OF SUGAR PUCKERING

HEATHER MAYES
Northwestern University
National Renewable Energy Laboratory

HEATHER MAYES WENT to her 2012 practicum thinking she'd research how nature creates something. At the end, she switched to the daunting task of deciphering how nature decomposes something.

The switch led Mayes, a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient, to significant insights into reactions behind one kind of biofuel production – and to a new thread for her chemical engineering doctoral research.

The material that ropes the projects together is cellulose, found in biomass like wood and plant stalks and leaves. It's a polysaccharide – sugar molecules linked by chemical connections called glycosidic bonds. Those strong bonds, Mayes says, are why you can take a cotton shirt, "wash it and beat it up over and over again and it's still a cotton shirt, not dust." But they also make it difficult to convert cellulose into biofuels – renewable energy for cars and other machines. Biomass competes less with food crops than do standard biofuel feedstocks like corn and soybeans.

With Chemical Engineering Professor Linda Broadbelt at Northwestern University near Chicago, Mayes investigates fast pyrolysis – heating biomass to convert cellulose directly to a liquid for processing into fuel. It's the sledgehammer approach, she says: "I'm just going to add so much heat, so much thermal energy that these bonds are going to be disrupted."

For her practicum at the National Renewable Energy Laboratory (NREL) in Colorado, Mayes planned to investigate how plants build lignin, another biomass component. But ultimately, Mayes and her advisor, NREL Principal Scientist Mark Nimlos, realized the task would take more time than Mayes had in Colorado. The studies yielded interesting results, but for her last three weeks Mayes turned to another project.

Mayes is accustomed to changing course. She found science and math classes at her suburban Chicago high school dry and uninspiring, and gravitated toward literature and art. She moved on to Harvard University but left after two years.

Back home, Mayes found a job with Teresa Woodruff, a Northwestern reproductive biology professor. Watching her boss, "I saw that science could be something to be passionate and excited about – which is not something I had previously witnessed."
~~~~~

BARNS, BIOFUELS AND NOBELS

A trip to Germany to meet with Nobel Prize winners wasn't all ballrooms and panel discussions for Heather Mayes. It also included a day on a farm, discussing alternative energy prospects.

Mayes was chosen for the 63rd Lindau Meeting of Nobel Prize Laureates, which unites science Nobel winners with young graduate researchers to share ideas and insights. As part of the conference, the journal *Nature* selected her and two other graduate students to appear in one of a series of videotaped discussions on selected topics. They talked with Steven Chu, winner of the 1997 physics prize and a former Department of Energy secretary, and Hartmut Michel, winner of the 1988 chemistry prize. Amid flowers, insects and mooing cows, they debated the merits and science of biofuels and other alternative energy sources. (*See the video at http://tinyurl.com/lobvj4s.*)

Mayes praised Chu's support for a variety of alternative energies. The other participants, she said, were involved in basic research, giving her an opening to discuss practical challenges associated with their ideas. "I was really happy to have the opportunity to bring an engineering perspective."

MEETINGS OF NOBEL LAUREATES LINDAU ®

There's no doubting Mayes' enthusiasm now. It comes through in a Web video, as she discusses renewable energy with Nobel laureates and fellow graduate students on a German farm. (See sidebar, "Barns, Biofuels and Nobels.") Biomass conversion is "an important problem and it's a good problem to answer right now" as supercomputers increase in power, Mayes says.

Broadbelt's group has long examined polymer decomposition but was specifically drawn to cellulose pyrolysis via the National Advanced Biofuels Consortium. The consortium evaluates competing technologies to choose the most promising ones for further research and piloting.

The partnership includes Gregg Beckham, an NREL senior engineer Mayes worked with on her second project: enzymatic hydrolysis, in which water and enzymes called glycosidic hydrolases (GHs) break cellulose into sugars with little or no heating. Scientists want to learn from these powerful enzymes, many of which bacteria and fungi produce, and improve their use in biofuel production.

It's fascinating, Mayes says. "It seems so elegant that (GHs) can do this really, really tough process at ambient temperature. That's really what hooked me," turning a three-week project into a continuing collaboration. GHs now are a major part of Mayes' doctoral thesis, and Beckham co-chairs her committee.

BACK TO SCHOOL

Pursuing a Ph.D. is a leap for Mayes, who was reluctant to try college again after Harvard. With support from her husband, Chris, Mayes earned a chemical engineering degree from the University of Illinois at Chicago. After three years in industry, she decided to pursue a doctoral and landed in Broadbelt's group, attacking cellulose.

The sugar molecules in cellulose and other complex carbohydrates are primarily based on rings of five carbon atoms plus a single oxygen. Attached to them are exocyclic (outside the ring) groups of hydrogen and oxygen atoms that influence the molecule's stability. As their name suggests, glycosidic hydrolases break the glycosidic bonds that join these sugars into carbohydrates.

The bonds are strong, Beckham says, but the sugar rings in them "are incredibly flexible molecules." GHs capitalize on that pliability to "pucker" the sugar rings, distorting them from stable shapes like a child twisting a Tinkertoy model.

To change geometry, molecules must have free energy – thermodynamic capacity to do work. Each sugar follows a kinetic path or landscape based on their free energies, physically changing between 38 puckering conformations. As sugar molecules follow kinetic paths to less-stable geometries, they often pass through intermediate conformations called transition structures.

The number and location of exocyclic groups also help determine energy barriers and affect the kinetic landscape, Mayes writes in a paper with Beckham and Broadbelt. To complicate matters, groups can rotate around their connection to the sugar ring, settling in positions called rotamers. Between puckers and rotamers, each kind of sugar can have thousands of conformations.

Although scientists believe puckering positions molecules so the catalyst can break glycosidic bonds, they're not sure how. Particular GHs find particular conformations more favorable for catalysis than others, but researchers aren't sure why. The question is, "are there kinetically favorable pathways to go from a (stable) conformation to one that's observed in enzyme-active sites?" Beckham says. "And can kinetics be part of the puzzle to ascertain why given enzymes will pucker toward one conformation versus another?"

Computation has an edge over experiments in answering these questions, Mayes says. Real-world tests can detect a molecule's change from one conformation to the other, but transition structures are elusive. Computers can model those changes and alternative kinetic paths. Unlike most previous research, Mayes' project also accounted for rotamers, taking advantage of increasingly powerful supercomputers and codes that incorporate the strange effects of quantum physics.

At first, Mayes planned to merely compare computational techniques, seeking those requiring the least work while still accurately figuring puckering energies and kinetic paths. To do that, however, she needed highly accurate but demanding calculations for comparison. "We realized there was a rich data set coming out of these 'gold standard'" calculations, leading her to pursue the project more deeply.

> "It seems so elegant that they can do this really, really tough process at ambient temperature. That's really what hooked me."
>
> ~~~~~

Mayes mapped the multiple "itineraries," or steps, six sugars can follow from one conformation to another, including the internal energy of each and where atoms lay in relation to each other. She generated atomic arrangements for 38 conformations for each sugar and a set of rotamers for each conformation. For most sugars, that led to thousands of possible conformations.

To screen them, Mayes used different mathematical methods, each more demanding and accurate than the previous one. Each time, she discarded conformations that clearly didn't fit. That way, she says, "I'm only spending the hard-core computational expense on conformations of the most interest."

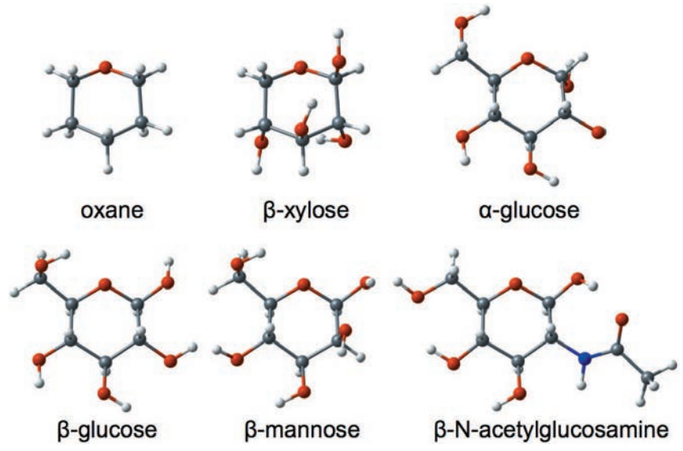
The expense came from computing the electronic structure – the positions and interactions of electrons – in conformations. For these calculations, Mayes used the coupled cluster for singles, doubles and perturbative triples method, or CCSD(T). It describes interactions in terms of excitations that push electrons into higher-energy orbitals. "That's the data I trust most. CCSD(T) is considered the gold standard."

MAPPING THE RINGS

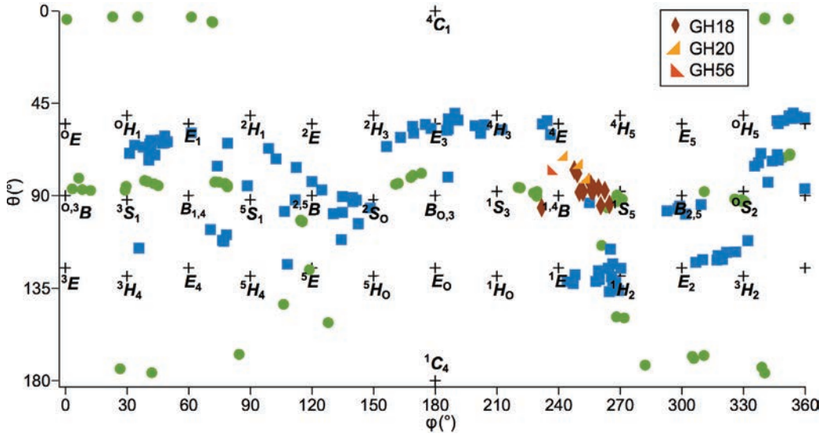
The calculations were conducted on several high-performance computers, including Gordon and Trestles at the San Diego Supercomputer Center (part of the National Science Foundation's XSEDE network), RedRock and RedMesa at NREL, and Carver at the National Energy Research Scientific Computing Center.

The result, Mayes and her colleagues say, is the most complete ring interconversion map yet for these sugars. It represents "several hundred thousand, at minimum, calculations that were done at very high levels of theory," Beckham says. "The comprehensive nature really sets this study apart. All of that was due to Heather."

Still, the results don't provide clear answers to why an enzyme will pucker a sugar toward one conformation rather than another, Mayes says. "It wasn't like there was one simple rule that held for all of the sugars." Enzymes "have all these different variables and they're trying to find a good combination."



Monosaccharides – single-sugar units – such as glucose, xylose and mannose are ubiquitous in biology, and understanding their reactions is crucial for many fields, from health care to renewable energy. In these portrayals, carbon atoms are gray, oxygens are red, hydrogens are white and nitrogens are blue. *Reprinted with permission from Mayes, H.B.; Broadbelt, L.J.; Beckham, G.T.* J. Am. Chem. Soc. *2014, 136, 1008-1022. Copyright 2013* American Chemical Society.



A two-dimensional projection of the Cremer-Pople sphere used to map how the central ring in monosaccharides can take on different geometries. Using computational modeling, Mayes calculated thousands of ways enzymes can "pucker" the ring in N-acetylglucosamine, a monosaccharide important in biomass and human health. Green circles indicate possible low-energy pucker geometries, blue squares designate intermediate geometries adopted to pass from one low-energy geometry to another, and the red and orange markers show geometries that enzymes employ in catalysis. The high-level electronic structure calculations provide a better understanding of why these particular geometries are advantageous for catalysis. *Reprinted with permission from Mayes, H.B.; Broadbelt, L.J.; Beckham, G.T.* J. Am. Chem. Soc. *2014, 136, 1008-1022. Copyright 2013* American Chemical Society.

The practicum taught Mayes methods she'll use for her doctoral research, Broadbelt says. It's been "a fantastic collaboration and all credit to Heather. She's the one who has really made it work."

Mayes recently published the puckering calculations paper in the *Journal of the American Chemical Society*. At the 2012 Pan-American Advanced Studies Institute meeting on Molecular-based Multiscale Modeling and Simulation in Uruguay, she won a best poster award for a report on her early practicum work comparing conformation-calculating methods.

Next, Mayes will simulate puckering in GH catalysts. But first she must finish her pyrolysis work.

To a layperson, enzymatic processing seems to have a big advantage over fast pyrolysis: GHs work at room temperature, while pyrolysis consumes energy to heat cellulose. But "there are reasons to love pyrolysis," Mayes says. For example, "whatever you give me, I can handle it," whereas particular enzymes only work on particular polysaccharides. Enzymes also can be expensive and need precise temperature controls.

Broadbelt wants computer modeling to help predict pyrolysis products and quantities, what conditions are best to produce desired products, and more. As with catalysis, reactions in pyrolysis experiments happen too fast to precisely observe them.

In a 2012 paper, Mayes and Broadbelt describe a model for forming levoglucosan, a possible fuel comprising much of the liquid cellulose pyrolysis produces. They modeled pyrolysis reactions that break glycosidic bonds to convert a cellulose component, methyl-cellobiose, into levoglucosan. Researchers thought the reactions happened in one of two ways, each involving at least two steps to the final product.

Mayes and Broadbelt found a third possibility: a single-step reaction that forms levoglucosan as glycosidic bonds break. Their predictions agree well with experiments, raising the possibility that models can help tune pyrolysis for optimal results.

It's the kind of work Mayes hopes to continue after graduation, due sometime in 2015, although she's uncertain whether she'll be at a national laboratory or in academia. "I like to say I have a nonlinear career trajectory," Mayes says, and that's likely to continue.

Broadbelt is just glad Mayes is in the field. When she reflects on how Mayes nearly didn't finish college or go into chemical engineering, "I think, my gosh, what a loss to science and engineering that would have been."

ASSEMBLING ENSEMBLES FOR PEAK ACCURACY

MILES LOPES
University of California, Berkeley
Sandia National Laboratories, California

"Computer science and statistics and applied math are all converging on the same thing: data."

~~~~~

Credit: Sandia National Laboratories

THE RESEARCH MILES LOPES and Philip Kegelmeyer conduct delves into hostile data – information that resists scrutiny.

"Quite often the data you're analyzing weren't collected for the purposes of making your analysis easy. They were collected for something else," says Kegelmeyer, a senior scientist at Sandia National Laboratories in California and Lopes' supervisor for his 2012 Department of Energy Computational Science Graduate Fellowship (DOE CSGF) practicum there.

The data set may be overwhelmingly large. It may be noisy, with lots of irrelevant information, or sparse, with pertinent material comprising a small part. There may even be "sets where somebody is deliberately trying to muck with your data in order to make you do poorly," such as email spammers.

Tools to interrogate hostile data include ensemble methods for classification, a machine-learning technique that uses the characteristics of known examples to categorize new data. For instance, junk email filters "learn" the features of unwanted messages, such as words or formats, then identify and sequester new messages that have similar qualities.

These methods create a collection of predictions – an ensemble – of how a new piece of information should be classified. The predictions are tallied, with a majority vote making the final determination. Ensemble methods handle hostile data well and run efficiently on high-performance computers. They can be used to analyze gene expression data and for other applications.

What's often unclear, however, is how big the ensemble should be. It took Lopes, a statistician, to find a way to answer that.

But don't call Lopes a statistician just yet. It may best describe what he does, but Lopes wants to transcend disciplinary boundaries. "It's a weird time now to pigeonhole yourself," he says, because "computer science and statistics and applied math are all converging on the same thing: data." In that atmosphere, disciplines matter less and "I can play a unique role in making links between computation and statistics."

In his work with Kegelmeyer, for example, Lopes studied a randomized classification algorithm – particularly how it arrives at its final level of accuracy. Historically, computer scientists, rather than statisticians, tackle problems of determining how quickly an algorithm converges on its best level of accuracy, Lopes says. Yet "due to the randomness in the algorithm, I was able to make at least a small advance because I could apply statistical techniques computer scientists are usually not familiar with."

To hear Kegelmeyer tell it, the advance was more than small. To understand why, it helps to know a bit more about ensemble methods.
~~~~~

Take a thousand data points, for example, somehow divided between two classes. Kegelmeyer says you could create a single "sage" algorithm and "tune the heck out of it to get the parameters just right" for classifying new examples based on those thousand points.

Another approach, called bagging (for bootstrap aggregating), takes random samples of the points to create remixed data sets. Inevitably, many of these remixed sets will overlap with each other and contain some replicates of original data points.

IDIOTS BEAT THE SAGE

A single classifier built on sampled data will be less accurate than the sage built on the full set, Kegelmeyer says. But the majority vote of an ensemble built from random data sets "will do no worse than the sage would have done, and in practice you usually do much better. This sounds silly – it sounds like the opinion of a bunch of idiots put together is better than the opinion of your one sage." But it works.

The tough question is how many votes are needed for a reasonably accurate classification. Generally, the larger the ensemble, the higher the accuracy. But each classifier burns computer resources and time. Run too few, and accuracy suffers. Run too many – a number beyond which accuracy improves insignificantly – and it wastes computer cycles.

"It's really important to let the (ensemble) grow until it plateaus, partially so you can do good science, but also partially so you don't leave performance on the table," Kegelmeyer says. "But it's hard to do that – to know how big to grow them in a computationally reliable fashion."

A sensible approach is experimenting with datasets to get an idea of how big the ensembles should be and increase that by 10 percent. Then "you cross your fingers and validate every so often," Kegelmeyer says. It's an intuitive approach that seems to work, Lopes adds, but there's no theory to back it. His job, after boning up on ensemble methods, was to find a technique underpinned by data and theory.

Researchers want to know how quickly the accuracy of the ensemble converges to a stable limit to accuracy, beyond which there's little to gain from adding classifiers. "That's what I was able to figure out in a precise way: the rate of convergence," Lopes says. "You can explicitly write a formula for it and use that to quantitatively decide how many classification functions you should train."

Adds Kegelmeyer, "To my knowledge, Miles is the first person to not only put an actual equation on that, but to prove it's correct."

Here's the concept: Suppose predictions are either 0 or 1, meaning you can view output from randomized classifiers as a sequence of coin flips. If the flips are statistically independent, finding the convergence rate is straightforward. But "each voter is making its predictions based on the same overall data," Lopes says, so the classifiers are not entirely independent. That complicates matters.

There is something that's almost as simple as independence, Lopes says: exchangeability, which allows for a controllable degree of correlation between classifiers. Lopes realized that exchangeability is a natural part of Random Forests, Kegelmeyer's favorite ensemble method, and used it to create a solution.

With an infinite number of classifiers, accuracy may settle to perhaps 80 percent. "Now," Lopes asks, "what if you could get to 79 percent using only a hundred classifiers? Well, that's pretty good, right?" His approach lets researchers choose how close they want to get to the accuracy limit and determine how many classifiers they need to do it. Simulations show it works as expected.

STAT LESSONS

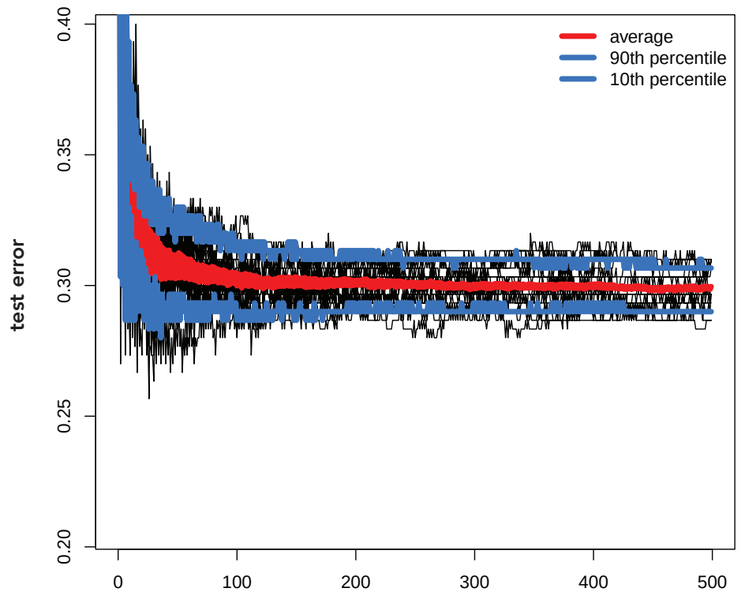
Kegelmeyer says Lopes helped him gain a deeper grasp of statistics, particularly a branch called nonparametric statistics. "There's a whole body of theoretical mathematics and theoretical statistics Miles is amazingly comfortable with and conversant in."

Lopes wrote a report on the project and delivered a talk at Sandia before leaving, prompting a researcher there to change his approach to a cybersecurity problem. Lopes also presented the work at the 2013 Joint Statistical Meetings in Montreal, and there are plans for a paper.

Since returning to Berkeley, he's worked on calculating the fluctuation in accuracy between multiple ensemble runs. Kegelmeyer explains: "Even though you know the convergence rate, there's some wiggle room" around it. "Miles' math will say, 'OK, if you only go to an ensemble of size 100, I can give you only a 73 percent confidence that you really hit your plateau of accuracy,'" Kegelmeyer says. "That's useful. You can trade off. You can say, 'How important is it that we be certain we're getting all the accuracy we can?'"

"That's what I was able to figure out in a precise way: the rate of convergence. …You can explicitly write a formula for it."

~~~~~



This plot reflects the performance of a collection of about 50 ensembles, each containing 500 classifiers, in a simulated classification problem. Even when all the ensembles have grown to a large size, the plot shows that their error rates may differ by a few percent, e.g. 28 to 32 percent. In practice, we never know the true test error of an ensemble, but we can be reasonably certain the ensemble is large enough so the error rate is stable. At the same time, we want to keep the ensemble size to a minimum in order to reduce computational costs. Lopes' work derived analytic formulae for the red and blue curves, offering practitioners a guide to knowing when an ensemble is "just large enough."

Kegelmeyer wants to implement the method for problems he's researching and to verify research demonstrating that some methods, when run on large enough ensembles, are no better than bagging. He'd also like to extend the work to ensembles with more than two classification categories.

Lopes' doctoral research at the University of California, Berkeley, also focuses on machine learning and targets the problem of taming giant scientific data sets. Big data, as it's known, is not only about more information but also more complicated information, says statistics professor Peter Bickel, Lopes' doctoral supervisor. Whether from genomics, astrophysics or other arenas, data have many dimensions – multiple features for each element, such as thousands of gene-expression results from a relative handful of cell lines.

To find patterns in these information clouds, researchers seek low-dimensional substructures such as sparse representations. In an astronomical image, for instance, only a few pixels out of millions show stars, Lopes says. "That means the image's effective dimension is much lower. The image is sparse, in the sense that if you put zero for black and one for white, the image is mostly zeros."

But algorithms often make assumptions about sparsity without validation, with potentially negative effects on the quality of their inferences and on computation.

In a paper published for the 2013 International Conference on Machine Learning, Lopes addressed this challenge in the context of compressed sensing, a framework for acquiring and reconstructing sparse signals. The paper posits a more robust definition of sparsity, making it possible to estimate the effective dimension of signals that are compressible, but have
~~~~~

many components that aren't perfectly zero. Lopes' approach also makes no assumptions about sparsity.

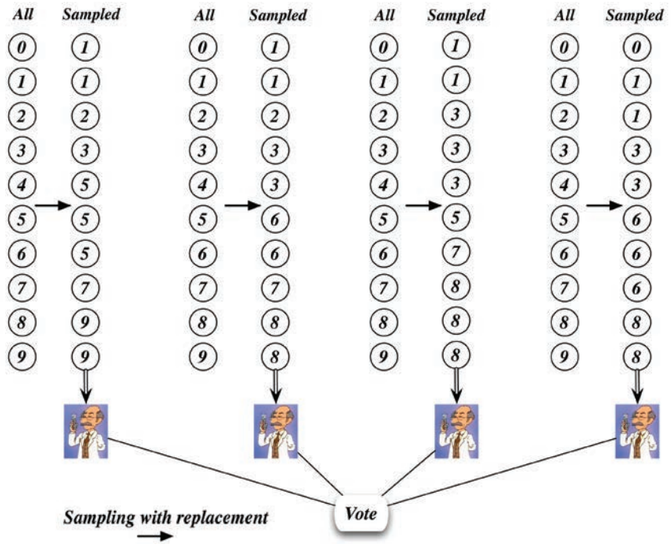
The algorithm makes it possible to adapt the measurement process to a signal's true sparsity level and to determine how many measurements should be collected for successful reconstruction. It uses random measurements to estimate sparsity and calculates how much confidence to place in that estimate. Simulations confirm the algorithm works as expected, regardless of how sparse the underlying signal may be.

One of the main themes of Lopes' research now is uncertainty quantification: putting a number on the error inherent in methods to seek signals in complex, high-dimensional, sparse data. Others have worked on similar problems, Bickel says, "but it was all in the context of moderately sized data sets and, more significantly, relatively low-dimensional data. That's not what's happening now."

That gives Lopes an opening, which he's sure to find intriguing. "Miles is an extremely independent guy," Bickel says. "He basically goes off and finds interesting questions he likes to work on, solves them and then comes and tells me about it. That's an ideal student."

Lopes expects to graduate in May 2015, but beyond that is uncertain what's next. "I want to pursue a research career, but whether that's at a national lab or a research university or a company, those are things I'm sorting out" – like finding a signal buried deep in data.

Bagging is a method for generating a randomized ensemble of classifiers. In this example we have a data set of 10 points labeled 0,1, ..., 9. The first step of the bagging process involves taking 10 random draws from this set (with replacement), producing a scrambled version of the original data set. A classifier is trained on the sampled points and then the sampling and training steps are repeated as many times as desired, creating a large ensemble of classifiers. Because the scrambled datasets contain different subsets of the original data, the resulting classifiers will have different accuracies in predicting labels for future examples.



"It's changed the trajectory of my Ph.D. It's really been fun."

~~~~~

Credit: Lawrence Berkeley National Laboratory

THE 'EARTHQUAKE MACHINE'

PHOEBE DEVRIES
Harvard University
Lawrence Berkeley National Laboratory

PHOEBE DEVRIES FIGURED her fall 2012 stint at Lawrence Berkeley National Laboratory (LBNL) was a chance to step outside her Harvard University doctoral research. "I thought of this as a three-month, kind of risky, not-knowing-what-I-was-doing practicum," says DeVries, a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient.

The project took a turn, however, shattering her research plan for a doctorate in earthquake science. Since returning to Harvard, DeVries has continued the practicum work, which could lead to improved models of forces that feed quakes. "It's changed the trajectory of my Ph.D. It's really been fun," she says.

Such course corrections aren't unusual for DeVries. As a Harvard mathematics undergraduate, she was drawn to Earth science when a friend majoring in the field invited her to a research talk. DeVries took a few classes in the subject and did her senior thesis with Brendan Meade, an Earth and planetary sciences professor.

After graduation, DeVries left Harvard's home in Cambridge, Massachusetts, for a fellowship at the other Cambridge – the famed English university – where she dove into a subject totally unrelated to her undergraduate major, earning a master of philosophy degree in polar studies.

DeVries returned to Harvard and joined Meade's group, which focuses less on what happens during earthquakes and more on what happens between them – particularly the signatures or signals that may say something about future temblors. The researchers want to use computer models and Earth measurements to characterize fault systems' present and past states well enough to explain and – perhaps someday – predict quakes.

Roland Burgmann, an LBNL researcher and geophysics professor at the University of California, Berkeley, says his group also investigates the entire earthquake cycle or "earthquake machine."

"When we have a large earthquake, what does that tell us about the mechanics of the system, but then ultimately, also, how does that translate into seismic hazard in surrounding regions?" Burgmann says.

Geophysicists once thought a large earthquake meant "in that particular area we're essentially done – we've relieved all the stress," Burgmann says. But studies have found stresses are redistributed in complex ways, both in time and space. "Understanding that redistribution can help us better understand how other faults might be pushed closer or further away from failure."

When DeVries joined Burgmann's group for her practicum, she planned to study three-dimensional viscoelastic relaxation codes for modeling earthquake cycles. Viscoelastic models combine two ways to calculate stresses on faults and subterranean strata: viscously, or how rock redistributes and relaxes stresses through its fluid-like movement over faults
~~~~~

and other rock over years or decades; and elastically, or how structures seek a return to their original condition.

The elastic model system, Burgmann says, "is a spring: You pull on it, it deforms immediately" in proportion to the force applied and "if you let it go it jumps immediately back." Time dependency – how stresses change throughout the long earthquake cycle – is inherent in viscoelastic models. "We're hitting our rocks with one sudden change in stress: the earthquake," then seeing how they flow to redistribute that stress over years.

TAKING TIME

But DeVries says the time dimension makes viscoelastic computational models so slow most can only simulate one earthquake and its effects, not long periods between temblors. Accelerating codes and making them more efficient is crucial, Burgmann says, because researchers want to run models hundreds or thousands of times to investigate how parameter changes affect the outcome. "We want to make sure that within that parameter space the solutions are accurate at all scales and all times."

To help reach that goal, DeVries planned to implement one or more earthquake models for parallel processing, which divides calculations and distributes pieces to multiple computer processors. The processors, working simultaneously, reach a solution more quickly than a single one would. DeVries hoped parallelizing the codes would make it easier to compare them and use them for more complex simulations.

But just days after starting, DeVries learned something that, again, altered her course. At the Southern California Earthquake Center Annual Meeting, she saw a poster describing a project benchmarking five earthquake cycle models. "It looked like they didn't agree very well, or they agreed but not as closely as you would expect," DeVries says. "That puzzled me."

Instead of "blindly just trying to run (the codes), compare them and parallelize them, I thought I should understand them" by writing one. She chose to implement a viscoelastic stress transfer solution developed by Mitsuhiro Matsu'ura, a famed Japanese seismologist who described the math in papers published over decades. "There's no other model that's that well documented," DeVries says. She devoted her practicum to understanding "the nitty-gritty of what was happening" in the algorithms, with the second goal of making them run faster.

DeVries first wrote the code for standard sequential processing, using the MATLAB language geoscientists prefer. Tests, however, found it bogged down on a key calculation: Fourier transforms, which decompose waveforms into their simpler regular components. The code requires calculating the inverse Fourier transform – reconstructing a signal from harmonic components – in cylindrical coordinates. For the viscoelastic earthquake problem, that's done with Bessel functions – thousands of them. And the Bessel function calculation "just takes forever," she says.

As a side project, DeVries devised a kernel that sends Bessel functions to a graphics processing unit (GPU), a chip originally made for computer games. GPUs, now used to accelerate operations in the world's biggest computers, take small amounts of information and perform myriad operations on them. The result: Bessel function calculations ran 40 times faster on a GPU than on a standard processor. The exercise was useful, DeVries says, but she probably won't integrate the kernel into the main program.

Back at Harvard, DeVries has implemented the code, dubbed Spectre, in the more versatile Fortran language and to run in parallel using the OpenMP interface. DeVries also has expanded Spectre to include Earth layers with different elastic qualities and to locate faults in any layer.

The exciting part, she says, is that a parallel three-dimensional earthquake cycle code "will allow us and has allowed us to have much more tectonically complex fault systems in our model." A 3-D code can account for effects that 2-D models omit, generating more accurate estimates of slip rate, the speed with which one side of a fault moves with respect to the other side.

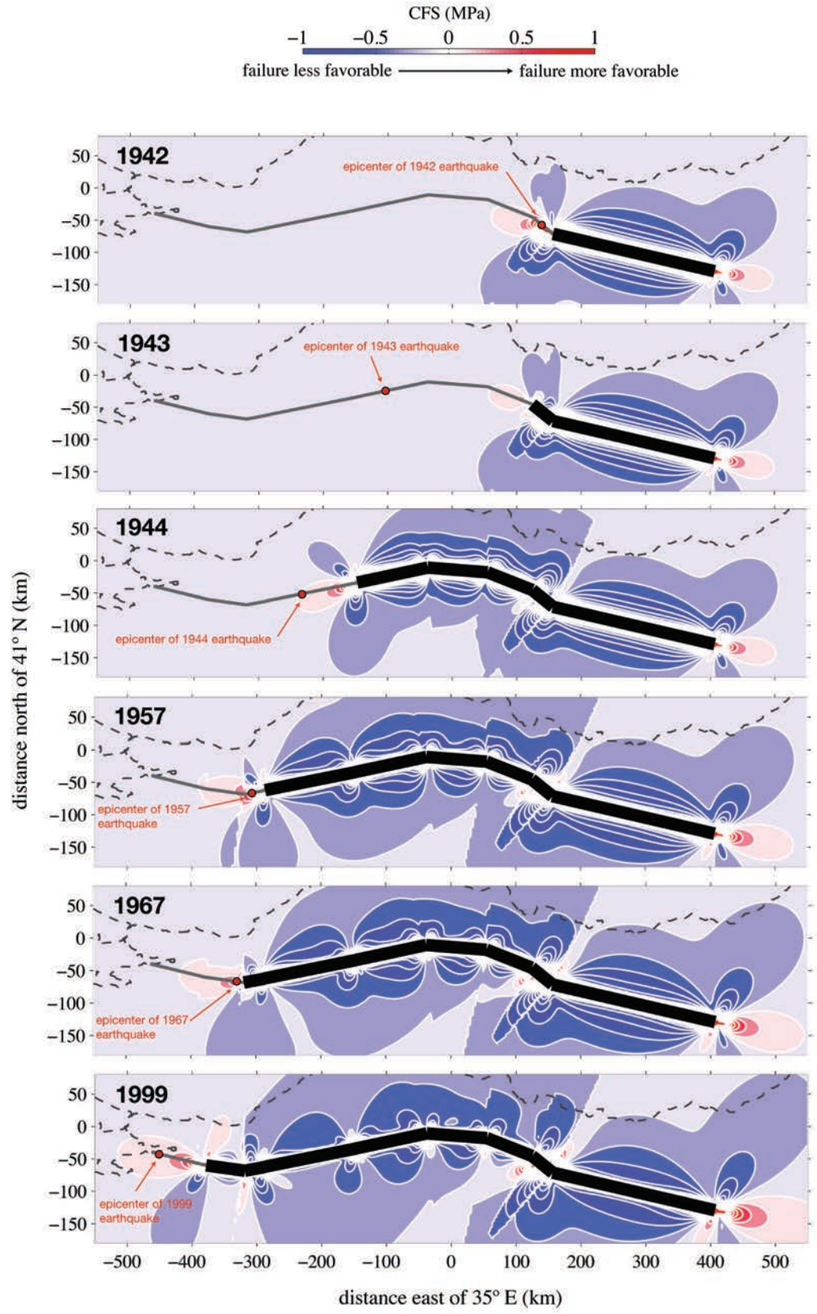
In one project, DeVries used Spectre to model activity on the North Anatolian Fault (NAF) stretching across northern Turkey. The fault is famed for the strong earthquakes that have occurred along it over the past 60 years, with most striking west of the previous one. "For earthquake scientists, it's way more of a regular pattern than anyone's ever seen," she says. "It seems like this is the place to look if we're ever going to understand earthquake triggering."

FAULTY BEHAVIOR

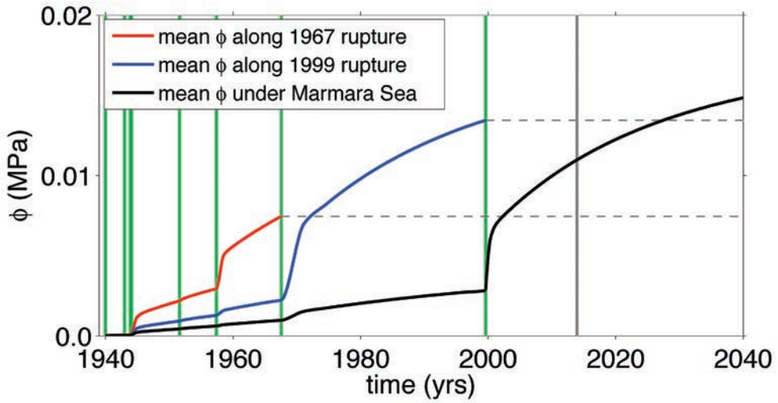
DeVries computed how the NAF earthquakes affected Coulomb failure stress (CFS), a calculation of stress a quake transfers along a fault or to nearby faults. Scientists hypothesize that CFS says something about how likely quakes are at locations close to previous tremors. Although controversial, DeVries adds, "it's really the only thing we have as a metric for that."

She set out to see if viscoelastic codes like Spectre could better explain the NAF earthquakes. Elastic models calculate stress changes that may contribute to the next quake at a specific location, but they can't explain why years or decades pass between events. If Earth's layers were only elastic, a second earthquake would occur almost instantly.

The Spectre model indicated that viscoelastic behavior could indeed explain the delay between earthquakes through



Coulomb failure stress calculations from a two-layer Burger's model at a depth of 10 km. Model results are shown at six times, just before major historical earthquakes on the North Anatolian Fault. The dotted gray line is the coastline of Turkey, the solid gray line is the surface trace of the North Anatolian Fault, and the bold black lines highlight the segments of the North Anatolian Fault that already have ruptured in the earthquake sequence.



This graph shows preliminary calculations of the time-dependent component of Coulomb failure stress (ϕ) along 1967 and 1999 fault planes prior to earthquakes. The vertical green lines show the times of the earthquakes in the sequence. The black curve shows the modeled mean value of ϕ over time along an eastern extension of the North Anatolian Fault under the Sea of Marmara, toward Istanbul. The vertical gray line marks the year 2014.

slowly increasing CFS at each epicenter. In a presentation at the 2013 American Geophysical Union Fall Meeting, DeVries exhibited a graph of calculations indicating the time-dependent, or viscous, CFS jumped at points west of earthquakes occurring in 1957 and 1967; quakes occurred later at those western points, in 1967 and 1999.

A third line tracks time-dependent CFS along a fault under the Sea of Marmara in western Turkey, showing a similar jump and slow increase after the 1999 earthquake. "That doesn't bode well for Istanbul," a city of more than 13 million on the sea's northern shore. The model has limits, but "if you follow this progression you would expect a really large earthquake" there.

DeVries next plans to implement Spectre in the ubiquitous Message Passing Interface (MPI), which runs on systems in which memory is either distributed or shared among processors. OpenMP runs in shared memory only. Using MPI will enable higher-resolution calculations and models of more complicated and realistic fault geometries.

The goal, Meade says, isn't to create the most detailed model. It's to develop one capable of showing whether a model can explain complex fault behavior. Since it's impossible to precisely quantify what's happening underground, "we want to ask, is there a simple model that has the properties that explain those sorts of things?" As a well-documented, open earthquake cycle code, Spectre also could be a tool to benchmark other models against.

Burgmann says DeVries, with her expertise in applied math, programming and geophysics, is prepared for the job. Although he mainly focuses on using fast, accurate and capable models, DeVries' interests are "both in the methodology and, hopefully, still in what we can learn from those tools. I think that's where we want our students to be."

If all goes well, DeVries could graduate in 2015 – or maybe 2016. She enjoys the academic life, so she may follow some previous students in Meade's group and finish in six years.

With parents who are professors, it's also no surprise DeVries' first career choice is university research and teaching. The flexibility and freedom to pursue research are attractive, she says, but there are downsides, so "I'm open to anything."

"I very much like being an ambassador between different fields and bridging a gap."

Credit: NASA/JPL-Caltech/Malin Space Science Systems

CONNECTING CHEMISTRY, AEROSPACE AND COMPUTERS

JASON BENDER
University of Minnesota
Argonne National Laboratory

JASON BENDER OFTEN MENTIONS breaking out of his comfort zone. The principle that has taken him on wilderness hikes and canoe trips across the country. He's on the board of an environmental conservation nonprofit. He left a challenging job in spacecraft design to pursue fundamental research.

And now he's ventured from aerospace engineering into the unfamiliar world of computational chemistry. "It's been challenging, but it's been rewarding," says Bender, a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient. "I very much like being an ambassador between different fields and bridging a gap."

With Graham Candler, a University of Minnesota professor of aerospace engineering and mechanics, Bender researches improved mathematical models of chemical reactions in fast-flowing gases. The combination of computational fluid dynamics (CFD) and chemistry can help scientists understand how air behaves around hypersonic aircraft and spacecraft – ones moving at many times the speed of sound.

Bender's boundary-crossing quest includes his summer 2012 practicum with Stephen Klippenstein in Argonne National Laboratory's Gas Phase Chemical Dynamics Group. As in Bender's doctoral research, the subject was fundamental chemical reactions but of a different flavor.

Klippenstein and his group calculate combustion reaction rates, especially for biologically based fuels like ethanol. Like Candler's group, the Argonne researchers study potential energy surfaces (PESs) to understand reactions. A PES is like a topographical map for a molecular system. The north-south and east-west dimensions correspond to individual atoms' positions within the system. The height of peaks and valleys represent the energy associated with those different arrangements.

"You can think of a reaction as always going from some reactant to some product" along an energy pathway, says Klippenstein, an Argonne distinguished fellow. Different paths have different energies. A PES portrays how a system's energy varies as molecules interact and their orientations change.

Tracking reaction pathways on a PES is like following boulders rolling along a hillside, Bender says. The path of a single point traveling along a PES describes a collision between two molecules – the type of event that can lead to a reaction. "Looking at the result of that event can tell you something about how the chemical reaction happens."

Computational chemists, including University of Minnesota chemistry professor Donald Truhlar, build PESs by calculating the electronic structure – roughly, the location of electrons – associated with different arrangements of atoms. Bender and Candler analyze reactions by using the quasiclassical trajectory (QCT) method to follow a large number of paths, called trajectories, along the surfaces.

"I like that fundamental aspect of this research – that we're really advancing the methods and the theories and the means of understanding these phenomena."

~~~~~

Klippenstein uses another technique: transition state theory. Rather than focusing on the entire path, it concentrates on a specific part that largely determines how the reaction proceeds.

To continue the analogy, imagine two lakes in a mountain range: one for reactant molecules, another for products. There's a barrier between the two – a bottleneck representing energy needed for the reaction, Klippenstein says. "The thing that determines how fast the reaction happens is how often it goes through that dividing surface." Transition state theory is based on understanding the reaction by focusing on the positions and energy of atoms in just that bottleneck region.

MOLECULES IN A TWIST

At Argonne, Bender sought a more accurate transition state approach for treating torsional vibrational modes – essentially twists in a molecule's three-dimensional shape. The modes contribute to the transition state partition function, a key quantity incorporating a sum over all possible energies.

As a molecule vibrates and stretches, hydrogen bonds connecting to oxygen atoms can form or break – and almost all biofuels have oxygen, Klippenstein says. "Treating the torsional partition functions for biofuel-related species is an interesting and important problem."

But calculating torsional vibrational modes' role in the partition function is tricky. Most solutions deal poorly with anharmonicity – when a stretching mode deviates from the harmonic behavior of a simple spring. Anharmonicity is particularly significant in torsional modes.

Bender says he and Klippenstein knew it was an untraditional project for an aerospace engineer, but it played to his interdisciplinary interests. It also was an ideal opportunity "to really put some concepts into practice and get my hands dirty." For example, the Argonne project was Bender's first parallel-processing code written from scratch.

Working with Computational Chemistry Specialist Yuri Georgievski and Argonne Distinguished Fellow Lawrence Harding, Bender used the standard programs Molpro and Gaussian to compute molecules' electronic structures. The codes can directly calculate torsional energies across a large sample of molecular shapes. But Molpro, Klippenstein says, isn't effectively implemented to process the randomly sampled geometries in parallel at large scales and can struggle with calculations to evaluate vibrational frequencies. Molpro computes energies for hundreds of molecular coordinates, but "if any one of them fails, the whole job fails."

Bender's code, called p3calc for "Properties of Potential energy surfaces Parallel calculator," performs normal mode calculations simultaneously across many processors, then gathers the results. To cope with the faults that halt Molpro, p3calc sets aside failed calculations for later treatment with another electronic structure solver. P3calc also can use Monte Carlo sampling to compute partition functions associated with anharmonic torsional modes.

Whether Klippenstein or another scientist is asking, the questions go beyond biofuels. Bender's code could help analyze any reaction involving torsional vibrational modes. "I like that fundamental aspect of this research – that we're really advancing the methods and the theories and the means of understanding these phenomena," Bender says.

That fundamental research interest is rooted in Bender's Cornell University undergraduate studies and his work as an associate engineer at NASA's Jet Propulsion Laboratory. He was on the team that put the Curiosity rover on Mars in August 2012, and his duties included helping develop actuators to drive its wheels and other parts. "I remember more than a few long nights in the lab, monitoring them during tests," Bender says.

The job got Bender interested in hypersonic vehicles. Whether on Earth or a planet like Mars, spacecraft moving that fast through the atmosphere encounter temperatures as high as 35,000 degrees Fahrenheit and pressures greater than 100 times that of Earth's atmosphere at the surface.

Under those conditions, nitrogen and oxygen molecules in air react in strange ways, Candler says. "They dissociate and recombine and ionize. All sorts of things happen that we just don't understand very well."

CHEMISTRY PLUS HYPERSONICS

Candler is a leader in using CFD for hypersonic flows. With a grant from the Air Force Office of Scientific Research, the Minnesota group wants to "plug the chemistry community into hypersonics," Bender says. "It was that interdisciplinary work that drew me to Minnesota."

The CFD codes engineers use to design high-speed vehicles like the space shuttle account for gas reactions based on data collected as far back as NASA's Apollo program. To trim the computer time and power needed to run them, the codes also make simplifying assumptions about gas chemistry.

To compensate for this uncertainty, spacecraft designers build in big safety margins. "That can get you a workable spacecraft, but it might be very far from the optimal design," Bender says.

Candler says that's why potential energy surfaces are important. "We're going to be able to understand what's really going on instead of having to guess about it. I very much doubt that the guesses of models we've made in the past are actually valid."

Bender's practicum gave him an advantage, Candler says, providing a computational chemistry foundation early in his graduate program. "He's really jumped into both areas fearlessly," he adds. Bender "has a good aerospace background but he's really picked up a lot on the chemistry side and now is conversant in their language as well as ours."

The Minnesota team is building PESs for reactions such as dissociation and recombination of oxygen and nitrogen. They start with a database of thousands of Molpro electronic structure calculations Truhlar and his fellow computational chemists have compiled for different arrangements of atoms.

Bender has focused mainly on using analytic fitting functions to construct PESs from this database. It's similar to drawing a trend line through measurements plotted on a graph. The team uses the fitting functions to model reactive trajectories with QCT, providing data to improve CFD models.

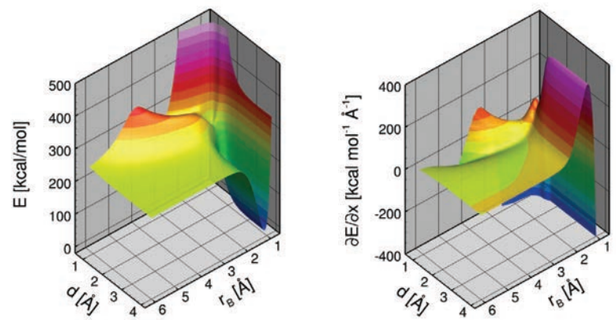
Determining the fitting functions, however, is tougher than just plotting a line on a graph. Each reaction the group studies involves at least four atoms, so the PES must account for six dimensions, one for each of the six internal degrees of freedom in a four-atom system. The multiple dimensions and related issues make fitting a PES for such systems a formidable challenge and a prominent theoretical research subject.

One technique, described in a paper published this year in the *Journal of Chemical Physics*, constructs a PES from a large number of local fits – meaning only the nearest data influence the fitted energy at any spot on the surface. Not every local fit is necessary for every calculation, however. To increase efficiency, the method statistically excludes some data points. Tests of the researchers' proposed method on a system of four nitrogen atoms found it did well in accurately predicting energies and smoothly representing the PES.

Bender and postdoctoral researcher Sriram Doraiswamy also have written a parallel processing implementation of the Truhlar group's ANT (Adiabatic and Nonadiabatic Trajectories) QCT chemistry code. "We can now run on our big cluster and do billions of interactions very efficiently," Candler says. "If all my students were as great to work with as Jason, I would live a charmed life – not that I don't have other excellent students. Jason is just off the scale."

Bender sees his work as directly tied to spacecraft on the launch pad. "That's an insight few people get to have. I can look at them and I can, in my mind, zoom in on what's actually going on between individual molecules crashing into the surface."

Bender plans to continue fundamental research after graduation, expected in late 2015. His dream job, at present, would be with a national laboratory. "Scientists have an absolutely vital and often underappreciated role to play in tackling some of the biggest challenges we're going to face in the next few decades, and I see the DOE playing a very important part in that."



This visualization shows a portion of the potential energy surface (PES) for a system of four nitrogen atoms, constructed using a new fitting method Jason Bender and colleagues designed and using electronic structure data from Donald Truhlar's group. The figures depict the energy (left) and one component of the energy gradient (right) as functions of two characteristic dimensions in an arrangement of the four nitrogen atoms. The full PES is six dimensional, corresponding to the four-atom system's six internal degrees of freedom, so the figures illustrate only a small subset of the full PES. *Reprinted with permission from J. D. Bender, S. Doraiswamy, D. G. Truhlar, and G. V. Candler,* J. Chem. Phys. *140, 054302 (2014).*
~~~~~

FORMER FELLOWS FIND SUCCESS

Doing Virtual Experiments to Tune Materials

Carolyn Phillips
Argonne National Laboratory

IT TOOK NEAR-DISASTER – twice – for Carolyn Phillips to conclude she was better equipped for computational science than engineering.

After earning a mathematics bachelor's at the Massachusetts Institute of Technology, Phillips stayed on for a master's in mechanical engineering. It let her move from pure theory, with its formulae and doing proofs, toward designing components and working in a machine shop.

But Phillips also found she lacked the type of attentiveness experimental research requires. She got distracted and let a container of liquid nitrogen go dry, creating a problem that eventually would have caused her experimental device to explode. Phillips and her advisor made repairs, but she later let it happen again. This time Phillips fixed it herself.

"I started realizing that when I'm doing computational work I was both intensely focused and able to play with the data, letting me find solutions in ways I couldn't in an experimental context," she says. Computational work "was better suited to my strengths and more tolerant of my weaknesses." Now Phillips, the Aneesur Rahman Postdoctoral Fellow at Argonne National Laboratory near Chicago, conducts her experiments on high-performance computers.

NAVY RESEARCH

Phillips attended MIT on a Navy Reserve Officers' Training Corps scholarship and eventually landed in Washington, where she managed Navy research programs at several government labs.

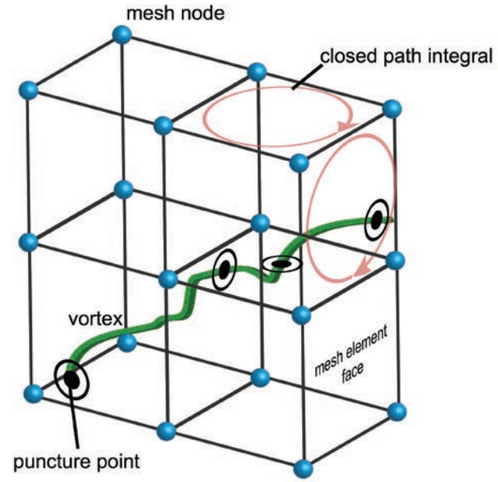
"I always felt [I was] on the wrong end of the phone," Phillips says. When talking with researchers, "I'd be excited about what they were doing and I would tell them the ideas I had, but ultimately it was not my job to do the research." She left to attend the University of Michigan as a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient from 2006 to 2010, earning an applied physics and scientific computing doctorate.

Phillips uses computers to model problems where the simple interactions of systems – atoms, molecules or materials – cause emergent behaviors. "The system almost becomes an experiment – like an experiment you can perform inside a computer." By changing parameters, Phillips can alter the systems' gross-scale behavior. The goal is to understand those changes and to manipulate them for desirable qualities in things like nanomaterials and superconductors. Phillips' research also draws her into the burgeoning "big data" arena.

In one project, Phillips and Argonne colleagues Tom Peterka, Dmitry Karpeyev and Andreas Glatz modeled a Type-II superconductor, characterized by magnetic vortices that inhibit conduction. Researchers want to understand and control the vortices to improve performance.

The superconductor model will produce prodigious data when it runs on high-performance computers, Phillips says, but "there's no way that you can store those data. You need to find the interesting features as they're being generated."

Phillips and her colleagues used a more precise, but also more computationally demanding, mathematical definition for the vortices. It let the team discard unnecessary information and focus solely on the vortex center, so "all the uninteresting parts have been made transparent, so to speak."



This illustrates a magnetic vortex line weaving through computational mesh elements. The bull's-eyes indicate the four points at which the vortex line punctures a mesh element face.

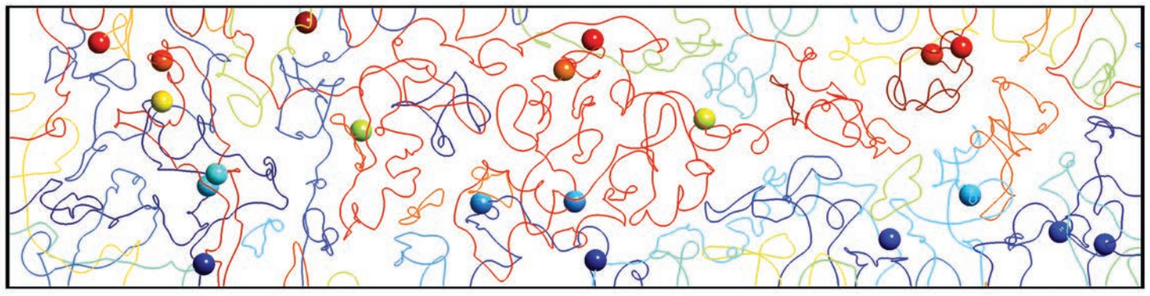
MORE SUPER SUPERCONDUCTORS

The research is part of OSCon (optimizing superconductor transport properties through large-scale simulation), an Argonne project under DOE's Scientific Discovery through Advanced Computing program. Argonne computational mathematician Stefan Wild, another OSCon participant, is a DOE CSGF alumnus who mentors Phillips.

Another algorithm Phillips developed discovers and maps materials' identities. The technique, described in the May 2014 *Journal of Computational Physics*, distributes data points throughout a model space. The algorithm asks what the material is at each one. Each query produces a label. The goal is to quickly map regions where the material's identity stays fixed and discover new identities with a minimum of data.

In regions where queries find multiple labels, the algorithm adds more points. Successive iterations quickly map the contiguous regions.

The technique will become important as computational models better predict material behaviors and emergent structures, Phillips says. But it already can help remove humans from directing data analysis – an important consideration as data rushes in. Her algorithm is "trying to keep up with the speed at which data can stream at us so we can focus on the interesting parts faster."



This image from a visualization shows a view down the x-axis of a simulated superconducting material containing inclusions (spheres) and tangled magnetic vortices. The magnetic field and current along the x-axis cause the vortices to twist and writhe.

Mario Trujillo
University of Wisconsin-Madison

Getting a Bead on Multiphase Flows

MULTIPHASE FLOWS – gas and liquid mixes that include bubbles, droplets and other forms – are everywhere, Mario Trujillo says.

They're in the oceans and rain showers, says Trujillo, a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) recipient from 1997 to 2000. They're in engines and furnaces and in supercomputers, where liquid cooling is a growing trend. Understanding multiphase flows can help improve these devices.

Researchers have studied these omnipresent phenomena for decades, but their dual nature also makes them difficult subjects because gases and liquids have such different densities from one another.

"In order to capture the interplay between momentum, for instance, or energy between those phases, you really must know where that interface is located, and that's not an easy problem," says Trujillo, now an associate professor at the University of Wisconsin-Madison.

Trujillo's path to chilly Wisconsin flowed from tropical El Salvador. In 1981 his father, a hospital administrator, and mother, a professor, moved the family to the United States to escape civil war.

DEFYING GRAVITY

As a student in Bloomington, Indiana, Trujillo discovered an aptitude for mathematics, but says "what I wanted to do with it wasn't clear. I remember thinking about gravity – antigravity machines – and I thought maybe that's what I want to do." It was an improbable idea, but "when you're that young, you don't know what you don't know. It's all about dreams."

Trujillo earned a mechanical engineering degree from the University of Minnesota. In graduate school at the University of Illinois at Urbana-Champaign, he turned to computational fluid dynamics (CFD).

Understanding the physics underlying multiphase flows has been a theme throughout Trujillo's career, enabled by the growing power of algorithms and computers. Trujillo and colleagues simulate the forces involved to find which ones affect flows the most.

"That's huge, because it allows you to simplify the equations" to focus on the main drivers, creating a versatile tool for further research. A picture may be worth a thousand words, Trujillo says, but "the right equation, that's worth a thousand pictures." With "the right mathematical expression for the physics you're looking at, you're capturing a wealth of information."

Focusing on the dominant physics while discarding others also cuts computational cost, making complex simulations feasible without long runs on huge machines. That's important for industry, where engineers run multiple simulations under variable conditions.

Trujillo studies fuel injection and vaporization at UW-Madison's Engine Research Center. His group also conducts basic science for the Office of Naval Research and other agencies and has a growing interest in developing numerical methods for CFD.

This visualization shows a simulated liquid jet striking a quiet pool from a low angle. Resistance from the pool's surface creates a stagnant point, redirecting the jet into a wave that collapses on itself, capturing an air cavity.

INTO THE POOL

In a 2013 *Physics of Fluids* paper, Trujillo and doctoral student Suraj Deshpande simulated a liquid jet striking a quiet pool. Experiments found jets at a shallow angle to the water's surface, as low as 10 degrees, capture large air cavities and push them below the surface. "We decided to home in and investigate it computationally," Trujillo says.

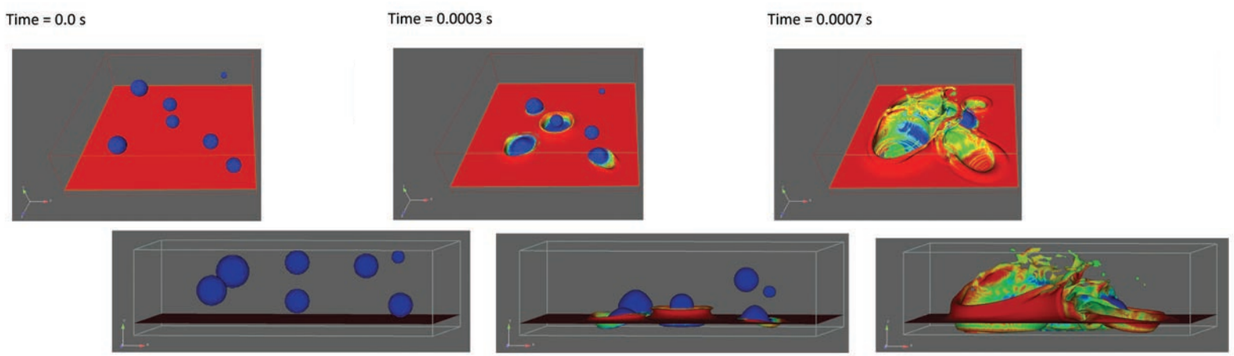
Simulations found the jet captures air cavities in a repeating process that's nearly identical to what happens in the initial impact. The researchers found the jet encounters resistance from the pool surface, creating a stagnant point and redirecting the jet into a wave that collapses on itself, capturing air.

The Navy is interested because bubbles are a torpedo signature. Other applications may include metal casting and blood transfusions, Trujillo says.

In a 2012 *Physics of Fluids* paper, Trujillo and master's student Steven Lewis simulated droplets striking a heated, wetted surface, a situation found in fuel injection and microelectronics cooling. They wanted to understand heat transfer in the thermal boundary layer, a thin liquid sheet that forms between the heated surface and the fluid. Generally, a thin boundary layer promotes heat transfer; a thicker one slows it.

The simulations found factors like droplet velocity and spacing greatly affected the boundary layer, while surface tension and other aspects had little effect. That means modelers can simplify the algorithms while preserving accuracy.

The research fits a statement Trujillo once heard – that there are three types of people: ones who make things happen, ones who watch things happen and ones who ask, "What happened?" With their simulations, he and his colleagues are, in a sense, doing all three.



These views from above and from the side show a simulation of multiple droplets over time landing in a shallow, hot liquid film at various angles, resulting in splashing and heat transfer. The color represents temperature in degrees Kelvin.

Steven Parker
NVIDIA Corporation

Following Photons from HPC to Hollywood

GROWING UP IN THE 1980s, Steven Parker didn't have the cash to master video games his pals played at local arcades.

He had an early console, the Atari 2600, but "you'd go to an arcade for anything fancy," Parker says. "Those games, you had to devote a lot of quarters to get good."

Parker was more intrigued by how the games – and computers – worked. While his buddies played Defender and Gauntlet, he accessed the mainframe his father, a physics professor, used. "I was always interested in both the electronics as well as the computer science," Parker says, "and so a lot of my career has been involved with the border between hardware and software."

Working on the border led him to a Department of Energy Computational Science Graduate Fellowship (DOE CSGF) from 1994 to 1997. At the University of Utah, Parker earned a doctorate with research combining high-performance computing (HPC) and the emerging field of visualization – creating images to help interpret data.

Parker later joined the faculty in Utah's School of Computing and its Scientific Computing and Imaging Institute, where his research included ray tracing, a key to making realistic visualizations for science, but also for video games and computer animation. Ray tracing simulates light physics, following virtual photons as they interact with virtual materials and eventually comprise images.

"The problem with ray tracing has always been that it produces high-quality images, but it's not fast," Parker says. Calculating photon paths gobbles so many cycles that most video games omit it.

GRAPHIC ACHIEVEMENT

Parker and his Utah colleagues spent years accelerating ray tracing with a combination of algorithms and programming models. In 2007, he founded a company, Rayscale, to commercialize the work. Just a few months later, computer graphics giant NVIDIA bought it, moving Parker out of academia and into industry.

Rayscale and NVIDIA technology combined to make NVIDIA OptiX, a general-purpose ray-tracing engine. OptiX lets developers easily write ray-tracing kernels that it compiles into instructions for graphics processing units (GPUs), workhorse computer graphics chips. NVIDIA developed the first GPUs for games, but they've since moved into some of the world's largest computers.

OptiX, Parker says, isn't typically used to build the main rendering tool. It gives designers a preview of how lighting changes will look in the final product. "An artist can move a light source around or move a character around and very quickly see the results of those changes, rather than waiting sometimes minutes to hours for a picture."

As OptiX was incorporated into NVIDIA software, Parker followed, moving out of research and into product management. He's now vice president and chief technology officer for the company's professional graphics business, overseeing the technical side of multiple software and hardware products.

Because his DOE CSGF experience blends visualization and HPC, Parker also develops applications for the Tesla, an NVIDIA GPU product designed for supercomputing. He helped develop Titan, a Cray XK7 at Oak Ridge National Laboratory that incorporates thousands of Tesla chips, and works with national laboratories on HPCs yet to come.

EXPANSIVE INTERESTS

"I keep track of a lot of different things, including how we're doing in media and entertainment, like video editing," as well as computer-assisted design and other arenas. "It's quite fun because it's a broad set of people and a broad set of applications focused on using advanced graphics to solve important problems. Plus, I work with an amazing team."

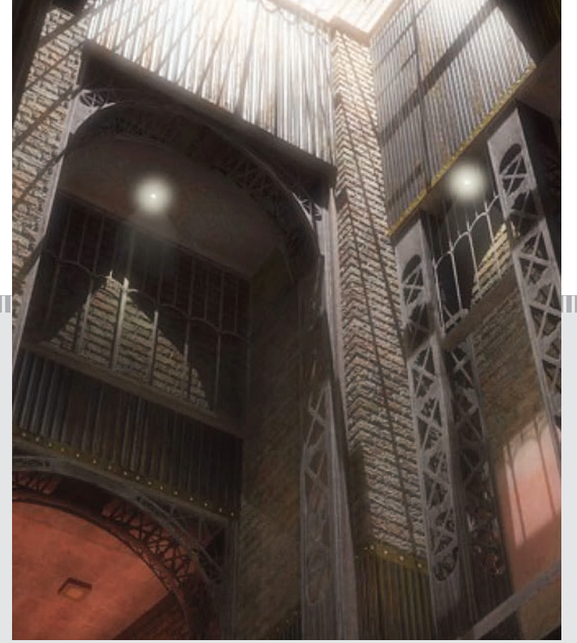
Hardware Parker oversees includes the Quadro, a GPU for professional workstations running high-end graphics software; GRID, a server board that makes NVIDIA GPUs accessible via cloud computing; and GRID VCA, which assembles GRID boards into an appliance serving multiple users of virtual graphics.

Software in Parker's management portfolio includes Iray and mental ray, image-rendering programs for professional users, like computer animation designers. Both include elements of the OptiX technology.

Hollywood studios frequently use NVIDIA products and Parker visits many of them, including Pixar, Disney and others, to develop collaborations.

It's changed the way he sees movies and television. Parker watched "Avatar," the 2009 blockbuster blending computer animation with live action, twice – once for the effects and once for the story.

"It's the same mindset as the video games," Parker says. "You wonder, 'How did they do that one? How could I make that one look more realistic? How can it be done faster?'"



This image demonstrates Image Space Photon Mapping, a real-time rendering algorithm published by Morgan McGuire of Williams College and David Luebke of NVIDIA Corp. in 2009 and implemented in the OptiX engine.

OptiX, a general-purpose ray-tracing system for many-core computing architectures, is part of the application used to create this image demonstrating physically based light transport through ray tracing. *Both images are from "OptiX: A General Purpose Ray Tracing Engine," published in the* ACM Transactions on Graphics (SIGGRAPH 2010 Proceedings), *August 2010.*

WINNING CYSE ESSAY

COMPETITION BOOSTS SCIENCE COMMUNICATION SKILLS

PHOEBE DEVRIES

EXCELLENCE IN COMMUNICATION

The DOE CSGF launched a writing contest in 2005 to give current and former fellows an opportunity to write about their work with a broader, non-technical audience in mind. The competition encourages better communication of computational science and engineering and its value to society.

In addition to recognition and a cash prize, winners receive the opportunity to work with a professional science writer to critique and copy-edit their entries. The latest winner is Phoebe DeVries, a fourth-year fellow at Harvard University.

For more information on the Communicate Your Science and Engineering Contest, visit **www.krellinst.org/csgf/outreach/cyse-contest**.

WHEN?

THE WINDOWS IN MY OFFICE rattled for a few seconds, but I didn't look up from my computer – it was probably just kids running around in the museum upstairs. Then my phone lit up with a text from my old friend Neal: "Phoebe! Was that an earthquake?"

I chuckled. Typical, dramatic Neal. Having studied earthquakes for a few years, I texted back confidently: "We're in Massachusetts. I bet a truck just drove by and shook your building!" and went back to work.

A few minutes later, I opened my web browser and the New York Times popped up: "EARTHQUAKE RATTLES EAST COAST."

In some ways, the earthquake we felt in Boston was representative of ones worldwide: It was unexpected and relatively small. In fact, most are even smaller than the one that shook my windows – too small to be felt at all.

Large earthquakes are rare, but just as sudden and much more damaging. The 2011 Japan earthquake resulted in a 15-meter-high tsunami, killed 18,000 people, triggered the meltdown of three reactors at the Fukushima-Daiichi nuclear power plant, and paralyzed the country's economy. The better we understand these big earthquakes, the more precisely we can estimate when and where they will occur and the more effectively countries can prepare for them.

For decades, researchers in active tectonics (the study of plate tectonics – how sections of the Earth's surface move – as it relates to human society today) have made progress in understanding seismic hazard. In fact, large-scale computer models can now predict where big earthquakes will occur. Scientists have known for decades that most take place on the boundaries, or faults, between tectonic plates. But large plates are actually comprised of many smaller ones, all bounded by faults that could cause earthquakes.

To understand which faults are most dangerous, researchers have developed three-dimensional computer models of systems in California, Japan, Tibet, and Turkey. These models are based on movement of the Earth's surface, tracked with thousands of Global Positioning System (GPS) stations – high-precision versions of the same GPS devices in your car and cellphone. Using this information, the models can estimate where stress is accumulating the fastest along fault systems, and therefore where large earthquakes are most likely to occur.

While scientists now understand the *where* relatively well, predicting *when* these huge ruptures will occur is much more difficult. But we have evidence that could help.

In the months following a large earthquake, nearby GPS stations record surface motions that are up to 10 times faster than those before it struck. This rapid post-earthquake motion (or "postseismic deformation") typically dies down after about a year.

One explanation for postseismic deformation is that big earthquakes stress the Earth's crust and rocks flow and deform in response. These stresses are greatest right after the event, driving the rocks' fast postseismic movement. As the stresses gradually relax, the shifting slows. If this is the case, and stresses due to earthquakes change over time, then past seismic events must be taken into account to understand the risk on nearby faults today.

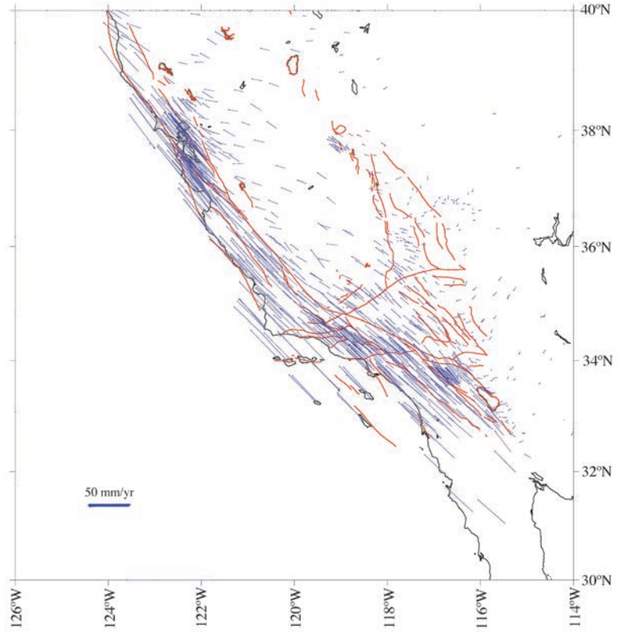
Based on these ideas, for two years my research group has been building a three-dimensional viscoelastic model of the Earth's crust and upper mantle. "Viscoelastic" means rocks in the model deform elastically – eventually returning to their previous condition – if hit instantaneously with large stresses, like those from an earthquake. But if a steady stress is applied over a long time period, the rocks can flow (viscously) and deform permanently. In other words, the Earth acts like Silly Putty: If you drop it on the floor, it bounces back, but if you knead it slowly, it flows.

With our viscoelastic model, we can input information about all previous earthquakes in a region and watch how stresses evolve. Solving equations for these stresses requires millions of calculations that would take years to run on a laptop. But breaking the problem into pieces and simultaneously running them on thousands of processors on Harvard University's Odyssey computer cluster allows us to complete these programs in days to weeks, and to understand the effects of geometrically complex faults and thousands of years of earthquake history.

Our models calculate stresses due to previous earthquakes. To understand when and where future serious events may occur, we compute a quantity called Coulomb failure stress. Earthquake scientists define this as a ratio of shear stress to normal stress, but it's intuitive: Think about an Oreo. To unstick the cookies on either side of the cream filling, you have to pull them apart (decrease the normal stress, or "unclamp" them) and twist them a little (increase the shear stress). Just the same way, faults are more likely to rupture if the normal stress is low and the shear stress on the fault plane is high.

By mapping the ratio of normal stress to shear stress in the Earth's crust that we calculate in our models, we can estimate where earthquakes are likely to occur. Furthermore, because our models allow rocks to flow under stress over tens to thousands of years, we can watch the Coulomb failure stresses change and try to understand when, as well as where, large earthquakes will happen. All of this is possible only with the help of high-performance computing.

Like my friend Neal, most people assume that because I study active tectonics I know something about when seismic events will occur. But the 2011 East Coast earthquake I felt in Boston was just as unexpected for me as it was for him. Trying to answer the key question – "when?" – is no easy feat, but it's the most important piece of earthquake-related information for the billions of people who live near faults that have ruptured before and will rupture again.



On this map of Southern California, blue arrows show GPS velocities, black lines are state boundaries and red lines are faults. GPS velocities like the ones shown are used to estimate where stress is accumulating the fastest within fault systems. *Data are from McClusky et al., 2001; Shen et al., 2003; Hammond and Thatcher, 2005; Williams et al., 2006; McCaffrey et al., 2007; and the Plate Boundary Observatory network velocity field, http://pboweb.unavco.org.*

The Frederick A. Howes Scholar in Computational Science award was established in 2001 to honor the late Frederick Anthony Howes, who was a champion for computational science education.

NO UNCERTAIN ACCOMPLISHMENT

STRIPLING HONORED FOR RESEARCH EXCELLENCE AND LEADERSHIP

David Brown, right, director of the Computational Research Division at Lawrence Berkeley National Laboratory, presents the 2014 Howes award to Hayes Stripling IV.

Stripling presents at the 2014 DOE CSGF Annual Program Review.

Hayes Stripling IV has spent much of his research career studying uncertainty, but he has no doubts that the Department of Energy Computational Science Graduate Fellowship (DOE CSGF) led him down unexpected and beneficial paths.

Stripling, a fellow from 2009 to 2013 and now a research engineer at ExxonMobil Corp. in Houston, is the 2014 Frederick A. Howes Scholar in Computational Science, recognizing his leadership and research achievements. He received his honorarium and award in July at the fellowship's annual program review in Arlington, Virginia.

"My work led to many different facets of computational science," Stripling said in a talk after accepting the prize. The DOE CSGF, he added, opened many of those doors.

The award honors Howes, who supported the fellowship as manager of DOE's Applied Mathematical Sciences Program and died unexpectedly in 1999 at age 51. Recipients are chosen not only for their outstanding technical achievements, but also for their exceptional leadership, integrity and character – qualities that reflect the prize's namesake.

Stripling, who earned a 2013 doctorate in nuclear engineering from Texas A&M University, says he was humbled to be chosen. Previous honorees, he says, "have been some of the most high-performing fellows – the fellows who are leaders" not only in their disciplines "but also as giving back to this community and being part of this fellowship."

After learning he was chosen, Stripling looked into Howes' background and found a prescient mathematics researcher and computational science advocate. "Somebody said Howes was always scheming to get mathematicians to talk to the applications people," Stripling says. "That's still hard now, and he was trying to make it happen 20 years ago." Howes' dedication, Stripling told his program review audience, "is a big reason we're in this room right now."

The committee of DOE CSGF alumni and friends that chose Stripling noted his major contributions to uncertainty quantification (UQ), which puts a number on how much a simulation's output can be trusted, and to algorithm development for high-performance computing (HPC). His approach to UQ for predictive simulations of advanced nuclear reactors "required the mastery of a breadth of topics from statistics, applied mathematics, nuclear physics, and computer science, which he achieved while maintaining a perfect scholastic record," the committee wrote.

Some of that mastery, Stripling says, arose from the fellowship's most unusual features: the program of study (POS) combining courses in applied mathematics, computer science and an application discipline (in his case, nuclear engineering); and the DOE laboratory practicum.

Stripling did two practicums, at Lawrence Livermore and Argonne national laboratories. The experiences provided him with useful new skills in statistics, an area in which he had little experience. His POS classes, meanwhile, taught him invaluable new approaches to problems.

New fellows have similar opportunities, Stripling says. The practicum, in particular, "will inject new ideas, new ways of thinking, new problems, new solutions into their research and really broaden their horizons."

Stripling's doctoral project identified and modified an obscure numerical method to solve equations arising from UQ. He developed a technique to evaluate how sensitive a nuclear reactor simulation is to changes in the parameters that govern it and implemented his method with novel algorithms for HPC. Stripling also developed computational techniques to analyze tradeoffs between computation, data storage, and data movement.

UQ is vital to the nuclear power industry, Stripling says. Licensing a reactor requires multiple tests and computer models to predict its behavior. More importantly, utilities must be able to assess the models' accuracy. For example, "It's not only how hot is the reactor going to be. It's how hot is it going to be and how sure are you about that answer."

Most of the computational and engineering tools used to design today's reactors were developed in the 1960s. The next generation of reactors will run hotter, must be safer, and must have longer lives to make them economical. They're more complex and the models must catch up, Stripling says.

Geophysics, the discipline he works in now at ExxonMobil, is poles apart from nuclear engineering, Stripling says. "But I still walked away from the CSGF program with a tremendous number of tools in my toolbox and all of those are still very handy," he adds. His UQ background also gives him a healthy skepticism of simulation results. "It's, 'Yes, I have this answer, but how much do I trust it?' That's an important question in any kind of computational science effort."

ABOUT FRED HOWES

In the 13 years since it was first conferred, the Frederick A. Howes Scholar in Computational Science award has become emblematic of 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.

In 2000, colleagues formed an informal committee to honor Howes. They chose the DOE CSGF as the vehicle and gathered donations, including a generous contribution from Howes' family, to endow an award in his name.

PAST HOWES SCHOLARS

2013	Ashlee Ford Versypt
2012	Carolyn Phillips and Matthew Reuter
2011	Alejandro Rodriguez
2010	Julianne Chung
2009	David Potere
2008	Mala Radhakrishnan
2007	Jaydeep Bardhan and Kristen Grauman
2006	Kevin Chu and Matthew Wolinsky
2005	Ryan Elliott and Judith Hill
2004	Collin Wick
2003	Oliver Fringer and Jon Wilkening
2001	Jeffrey Hittinger and Mayya Tokman

Fellow Evan Gawlik delivers his talk on universal meshes for problems with moving boundaries. Fellows' talks are a program review staple.

ANNUAL PROGRAM REVIEW

FELLOWS AND ALUMNI GAIN INSIGHTS AND BUILD BONDS

For students just entering the fellowship, the annual program review is a unique introduction to the DOE CSGF's benefits, to new colleagues and their diverse research, and to the computational science community.

For returning fellows, it's a chance to examine unexplored interests, renew acquaintances, make new connections and learn about practicums, computational methods and more.

At the gathering, fellows hear from top Department of Energy officials. They meet national laboratory personnel who are recruiting for practicums and, possibly, full-time jobs. Alumni working at labs, in academia or in industry also are on hand. Many are available to mentor fellows.

Fourth-year fellows give talks about their research, gaining valuable lecture skills and sharing their discoveries. Other fellows display posters about their latest projects, communicating their work while enhancing their presentation abilities.

"You're at the bleeding edge of computational science."

~~~~~

This year's program review featured a high-performance computing workshop designed to get fellows up and running on some of the world's biggest machines. A special session on poster design also helped them improve their communication and research skills. And prominent guest speakers provided insights into science and government.

New fellow Gerald Wang of the Massachusetts Institute of Technology said the program review helped him appreciate the sense of community the DOE CSGF fosters. He made friends and "we've talked about our problems and common problems and challenges we face in our research." They've already planned to reconnect at professional conferences in the coming academic year.

Here are highlights from the 2014 program review, held July 14-17 in Arlington, Virginia. Video of most presentations can be found at **www.krellinst.org/csgf/conf/2014/video.**

DOE SPEAKERS: *Fellowship Program is Vital*

High-ranking officials from the Department of Energy Office of Science and the National Nuclear Security Administration (NNSA) make time to welcome DOE CSGF recipients and alumni to each program review. This year was no different.

Patricia Dehmer, Office of Science deputy director for science programs, told fellows the DOE CSGF is a top priority because "high-performance computing touches everything. It touches science, it touches engineering, and it touches all aspects of our everyday life." The program, she added, is one of the country's most innovative fellowships and "the who's who of leaders now in applied math, computational sciences, disciplinary sciences, has come through this program."

Kathleen Alexander, assistant deputy administrator for the NNSA Office of Research, Development, Test, and Evaluation described the agency's mission and the role of computation in meeting it. Maintaining the safety and readiness of the nation's nuclear deterrent without testing requires simulation, and the NNSA must have high confidence in its models' predictive capabilities, she said.

"Training the next generation of computational scientists and engineers is more important than ever to us," Alexander added,

Fellow Carmeline Dsilva makes a point in her talk on techniques to mine data from chemical and biological systems.

The poster session prompts a discussion between fellows Sarah Middleton, Isha Jain and Hannah De Jong.

and the innovation high-performance computing enables is critical to national security and economic competitiveness.

"You're at the bleeding edge of computational science," especially at the extreme scales and conditions relevant to nuclear stockpile stewardship. "You're leaders in a race that we can't afford to lose."

FELLOWS' TALKS: *Summing Up Research*

At each DOE CSGF Annual Program Review, those finishing their time in the fellowship deliver lectures to share their research findings with colleagues, DOE laboratories and DOE sponsors.

In 2014, 18 fourth-year fellows gave talks. The topics demonstrate the diversity of computational science applications: genomics, methane emissions, nanotubes, machine learning, chemistry, volcanic flows, mouse behavior, fusion and more.

SPECIAL PRESENTATIONS: *Science and Policy*

Keynote: Sarah Richardson, DOE CSGF Fellow from 2007-2011

Hearing from graduates of the program is always a highlight for attendees. This year, alumna Sarah Richardson presented the keynote to open the program review.

Richardson, a distinguished postdoctoral fellow in genomics at the Joint Genome Institute at Lawrence Berkeley National Laboratory, said evolution has given bacteria and plants the tools
~~~~~

for myriad useful tasks. She helps find ways humans can harness those abilities to address problems in energy and the environment.

Richardson thinks of Earth as a massively parallel computation in which bacteria and plants have solved challenges to survival. "Chemically, computationally," humans "are not yet as good at figuring out this stuff as 3.6 billion years of this massively parallel computation has been," she added.

With her colleagues, Richardson studies genomic data, seeking ways to employ bacteria to fix nitrogen from the air for plants, reduce methane emissions, produce vitamins in the human gut, and perform other tasks.

"If I could grow these things, if I could domesticate them, if I could get them to work with us instead of just on us, what could we do?" she asked. "That's my dream. I want to be a bacteria wrangler."

Special Presentation: Tobin Smith — Vice President for Policy, Association of American Universities

Another regular feature of the program review is a presentation focused on science's role in society and government so participants better understand their position in the larger community.

This year's speaker, Tobin Smith, said scientists must know how the federal system for supporting research developed if they're to defend it.

The history of federal science is key to understanding many policy questions, Smith said, even some he deals with today as vice president for policy at the Association of American Universities. It began during World War II, when the government created science agencies and supported universities for war-related research like the Manhattan Project, Smith said. The United States later settled on today's largely peer-reviewed, merit-based research support system – one that can be contentious, but backs basic research.

The creation of mission-related research agencies like the Office of Naval Research and the DOE Office of Science also strengthens federal science efforts, Smith said. When combined with support for basic research, "accidentally we ended up with a system that supports science pluralistically," he added. "It's one of the reasons our system is the (world's) most successful" – and worth endorsing.

POSTER SESSIONS:
Science, Camaraderie and Connections

Fellows' Poster Session

The fellows' poster session is an important part of each year's program review, attendees say. Science is at center stage, with first- through third-year fellows explaining their latest projects via a technical display, but participants also form friendships and explore connections with others' research. Personnel from national laboratories and the Department of Energy also are on hand to visit with fellows.

Left: A discussion at the fellows' poster session prompts laughter from fellows Sherwood Richers (left), Thomas Holoien, Gerald Wang and Maxwell Hutchinson.

Alumna Sarah Richardson, right, talks with fellows Jay Stotsky, left, and Victor Minden at the Lawrence Berkeley National Laboratory booth.

Steve Binkley, associate director of Advanced Scientific Computing Research in the DOE Office of Science, welcomed the group. "The skills that are being cultivated in CSGF members are absolutely vital to the future health and vitality of computational science activities in the Department of Energy," he said, and the fellowship has advanced computational science and expertise across the entire agency.

The applications on display were as varied as fellows' interests. Posters covered fluid dynamics, cosmology, algorithm development, imaging, chemistry, and even online games as examples of complex systems analysis. Fellows also benefitted from peer review by alumni volunteers, who provided written poster critiques.

DOE Laboratories Poster Session

Department of Energy national laboratories know from experience that DOE CSGF recipients are among the nation's best students. That's why so many send staff – some of them fellowship alumni – to the program review's lab poster session. For fellows, it's a key entrée not only to practicum opportunities but also to collaborations and post-graduate employment.

This year, 15 laboratories participated. "It's very rare that you get representatives from all the national labs in one location," said Keith LeChien, director of the Office of Inertial Confinement Fusion for the National Nuclear Security Administration, as he opened the session. "This is an excellent opportunity to go out and talk with folks."

HPC WORKSHOP:
Bringing Fellows Up to Speed

High-performance computing workshops are a key benefit of attending the DOE CSGF Annual Program Review. For some fellows it's an introduction to the world of supercomputers. For others, it's a chance to deepen their knowledge of the fundamental art that distinguishes their discipline.

Participants at 2014's gathering chose between introductory topics or an advanced track intended for third- and fourth-year fellows. The session ended with a panel discussion on how fellows can tap into some of DOE's most powerful computers.

Rob Neely, associate division leader at the Lawrence Livermore National Laboratory (LLNL) Center for Applied Scientific Computing, led the first introductory workshop. Neely ran through a quick HPC history before explaining memory structure, scalability, parallel programming concepts and other aspects.

In the second session, DOE CSGF alumna Judith Hill told participants a successful computational science investigation should address applications, algorithms, computer architecture, and results analysis. "The measure of our success should be the impact we have on a scientific domain, computer science or mathematics," said Hill, the task lead for facility liaison support at Oak Ridge National Laboratory's Leadership Computing Facility.

In the first advanced session, Katie Antypas, Services Department head at Lawrence Berkeley National Laboratory's National Energy Research Scientific Computing Center, discussed preparing applications for many-core computer architectures.

Fred Streitz, chief computational scientist for LLNL's Physical and Life Sciences Directorate, told workshop participants they belong to a "magic generation" that will set standards for exascale, the next supercomputing milestone.

At an HPC workshop session, alumna Judith Hill of Oak Ridge National Laboratory ventures into the audience to encourage discussion.

ALUMNI: MAJOR DISCIPLINES

The Department of Energy Computational Science Graduate Fellowship (DOE CSGF) creates a community of scholars and researchers capable of using the best high-performance computers to advance science and innovation. The accompanying chart divides graduates' majors into four broad areas, but alumni work in a diverse range of occupations.

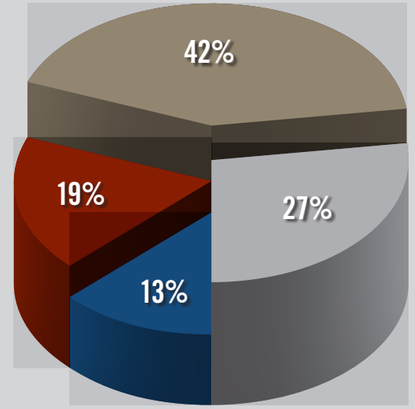
The DOE CSGF is a partnership between the DOE Office of Science and the National Nuclear Security Administration. Since its founding in 1991, the fellowship has graduated more than 325 doctoral students, seeding industry and academia with scientists and researchers who lead the way in helping the United States maintain and regain its competitive edge, generating jobs and income.

For a complete list of alumni (by last name, Ph.D. institution, fellowship start year, practicum location, current location, and area of study), go to **www.krellinst.org/csgf /alumni.**

DOE CSGF ALUMNI: DEGREES AWARDED BY DOCTORAL FIELD

- Engineering
- Physical Science
- Computer Science & Applied Mathematics
- Biological Sciences & Engineering

Information current as of September 2014



CLASS OF 2014

Mary Benage
Georgia Institute of Technology
Geophysics
Advisor: Josef Dufek
Practicum: Lawrence Livermore National Laboratory
Contact: mary.benage@eas.gatech.edu

Aleah Caulin
University of Pennsylvania
Genomics and Computational Biology
Advisor: Carlo Maley
Practicum: Lawrence Berkeley National Laboratory
Contact: alefox@mail.med.upenn.edu

Seth Davidovits
Princeton University
Plasma Physics
Advisor: Nathaniel Fisch
Practicum: Lawrence Berkeley National Laboratory
Contact: sdavidov@princeton.edu

Carmeline Dsilva
Princeton University
Chemical Engineering
Advisor: Ioannis Kevrekidis
Practicum: Argonne National Laboratory
Contact: cdsilva@princeton.edu

Christopher Eldred
Colorado State University
Climate Modeling
Advisor: David Randall
Practicum: Lawrence Berkeley National Laboratory
Contact: chris.eldred@gmail.com

Thomas Fai
New York University
Applied Mathematics
Advisor: Charles Peskin
Practicum: Lawrence Berkeley National Laboratory
Contact: tfai@cims.nyu.edu

Charles Frogner
Massachusetts Institute of Technology
Computational Biology
Advisor: Tomaso Poggio
Practicum: Lawrence Berkeley National Laboratory
Contact: frogner@mit.edu

Evan Gawlik
Stanford University
Applied Mathematics
Advisor: Adrian Lew
Practicum: Argonne National Laboratory
Contact: egawlik@stanford.edu

Christopher Ivey
Stanford University
Computational Fluid Dynamics
Advisor: Parviz Moin
Practicum: Argonne National Laboratory
Contact: civey@stanford.edu

Irene Kaplow
Stanford University
Computational Biology
Advisor: Hunter Fraser
Practicum: Lawrence Berkeley National Laboratory
Contact: ikaplow@stanford.edu

Miles Lopes
University of California, Berkeley
Machine Learning
Advisor: Peter Bickel
Practicum: Sandia National Laboratories, California
Contact: mlopes@stat.berkeley.edu

Peter Maginot
Texas A&M University
Nuclear Engineering
Advisor: Jim Morel
Practicum: Knolls Atomic Power Laboratory
Contact: pmaginot@neo.tamu.edu

Devin Matthews
University of Texas
Chemistry
Advisor: John Stanton
Practicum: Argonne National Laboratory
Contact: dmatthews@utexas.edu

Scot Miller
Harvard University
Atmospheric Sciences
Advisor: Steven Wofsy
Practicum: Lawrence Berkeley National Laboratory
Contact: scot.m.miller@gmail.com

Kenley Pelzer
University of Chicago
Theoretical Physical Chemistry
Advisor: Greg Engel
Practicum: Argonne National Laboratory
Contact: kpelzer@uchicago.edu

Christopher Quinn
University of Illinois at Urbana-Champaign
Communications
Advisor: Negar Kiyavash
Practicum: Sandia National Laboratories, California
Contact: quinn7@illinois.edu

Aaron Sisto
Stanford University
Computational Chemistry
Advisor: Todd Martinez
Practicum: Sandia National Laboratories, California
Contact: asisto@stanford.edu

Edgar Solomonik
University of California, Berkeley
Computer Science
Advisor: James Demmel
Practicum: Argonne National Laboratory
Contact: solomonik@berkeley.edu

Zachary Ulissi
Massachusetts Institute of Technology
Chemical Engineering
Advisor: Michael Strano
Practicum: Argonne National Laboratory
Contact: zulissi@gmail.com

4TH YEAR FELLOWS

Front, left to right: Alexander Rattner, Heather Mayes, Michael Rosario, Maxwell Hutchinson and Rogelio Cardona-Rivera; **Middle, left to right:** Aurora Pribram-Jones, Curtis Lee, Hansi Singh, Jason Bender, Robert Parrish, Matthew Zahr and Chris Smillie; **Back, left to right:** Phoebe DeVries, Sarah Loos, Jarrod McClean, Joshua Vermaas, Daniel Dandurand* and Omar Hafez

*Withdrew in 2012

Jason Bender
University of Minnesota
Hypersonic Computational Fluid Dynamics
Advisor: Graham Candler
Practicum: Argonne National Laboratory
Contact: jbender73@gmail.com

Rogelio Cardona-Rivera
North Carolina State University
Artificial Intelligence
Advisor: R. Michael Young
Practicum: Sandia National Laboratories, New Mexico
Contact: recardon@ncsu.edu

Phoebe DeVries
Harvard University
Earth Science
Advisor: Brendan Meade
Practicum: Lawrence Berkeley National Laboratory
Contact: phoebemaherrobinson@gmail.com

Omar Hafez
University of California, Davis
Computational Solid Mechanics
Advisor: Mark Rashid
Practicum: Lawrence Livermore National Laboratory
Contact: omhafez@ucdavis.edu

Maxwell Hutchinson
University of Chicago
Physics
Advisor: Robert Rosner
Practicum: Lawrence Berkeley National Laboratory
Contact: maxhutch@gmail.com

Curtis Lee
Duke University
Computational Mechanics
Advisor: John Dolbow
Practicum: Lawrence Berkeley National Laboratory
Contact: calee181@gmail.com

Sarah Loos
Carnegie Mellon University
Verification of Hybrid Systems
Advisor: Andre Platzer
Practicum: Oak Ridge National Laboratory
Contact: sloos@cs.cmu.edu

Heather Mayes
Northwestern University
Chemical Engineering
Advisor: Linda Broadbelt
Practicum: National Renewable Energy Laboratory
Contact: hmayes@u.northwestern.edu

Jarrod McClean
Harvard University
Chemical Physics
Advisor: Alan Aspuru-Guzik
Practicum: Los Alamos National Laboratory
Contact: jmcclean@fas.harvard.edu

Robert Parrish
Georgia Institute of Technology
Theoretical Chemistry
Advisor: David Sherrill
Practicum: Lawrence Livermore National Laboratory
Contact: robparrish@gatech.edu

Aurora Pribram-Jones
University of California, Irvine
Theoretical Chemistry
Advisor: Kieron Burke
Practicum: Sandia National Laboratories, New Mexico
Contact: apribram@uci.edu

Alexander Rattner
Georgia Institute of Technology
Mechanical Engineering
Advisor: Srinivas Garimella
Practicum: Idaho National Laboratory
Contact: Alex.Rattner@gatech.edu

Michael Rosario
Duke University
Evolutionary Biomechanics
Advisor: Sheila Patek
Practicum: Sandia National Laboratories, California
Contact: mvr9@duke.edu

Hansi Singh
University of Washington
Atmosphere-Ocean Physics
Advisor: Cecilia Bitz
Practicum: Pacific Northwest National Laboratory
Contact: hansi@atmos.washington.edu

Chris Smillie
Massachusetts Institute of Technology
Biology, Computer Science and Bioengineering
Advisor: Eric Alm
Practicum: Lawrence Berkeley National Laboratory
Contact: csmillie@mit.edu

Joshua Vermaas
University of Illinois at Urbana-Champaign
Biophysics
Advisor: Emad Tajkhorshid
Practicum: National Renewable Energy Laboratory
Contact: vermaas2@illinois.edu

Matthew Zahr
Stanford University
Computational and Mathematical Engineering
Advisor: Charbel Farhat
Practicum: Lawrence Berkeley National Laboratory
Contact: mzahr@stanford.edu

3RD YEAR FELLOWS

Front, left to right: Miles Lubin, Sherwood Richers, Brian Powell*, Dragos Velicanu, Victor Minden and Andrew Till; **Middle, left to right:** Melissa Yeung, Jamie Smedsmo, Daniel Strouse, Andrew Stine, Brenhin Keller, Sarah Middleton, Justin Lee and Britni Crocker; **Back, left to right:** Eileen Martin, Eric Isaacs, Thomas Catanach, Samuel Blau, Derek Macklin, Andrew Stershic and Jesse Lopez

*Withdrew in 2013

Samuel Blau
Harvard University
Chemical Physics
Advisor: Alan Aspuru-Guzik
Practicum: Argonne National Laboratory
Contact: sblau@fas.harvard.edu

Thomas Catanach
California Institute of Technology
Applied and Computational Mathematics
Advisor: Jim Beck
Practicum: Los Alamos National Laboratory
Contact: picatanach@gmail.com

Britni Crocker
Massachusetts Institute of Technology
Computational Neuroscience
Advisor: Sydney Cash
Practicum: Los Alamos National Laboratory
Contact: intirb@hotmail.com

Eric Isaacs
Columbia University
Applied Physics
Advisor: Chris Marianetti
Practicum: Brookhaven National Laboratory
Contact: ebi2104@columbia.edu

Brenhin Keller
Princeton University
Geochemistry and Geochronology
Advisor: Blair Schoene
Practicum: Lawrence Berkeley National Laboratory
Contact: cbkeller@princeton.edu

Justin Lee
Massachusetts Institute of Technology
Computational Imaging/Biomedical Optics
Advisor: George Barbastathis
Practicum: Lawrence Berkeley National Laboratory
Contact: jlee08@gmail.com

Jesse Lopez
Oregon Health and Science University
Environmental Science and Engineering
Advisor: Antonio Baptista
Practicum: Argonne National Laboratory
Contact: lopezj@stccmop.org

Miles Lubin
Massachusetts Institute of Technology
Operations Research
Advisor: Juan Pablo Vielma
Practicum: Los Alamos National Laboratory
Contact: miles.lubin@gmail.com

Derek Macklin
Stanford University
Computational and Systems Biology
Advisor: Markus Covert
Practicum: Lawrence Berkeley National Laboratory
Contact: derek.krellinst.org@nrm.com

Eileen Martin
Stanford University
Computational and Mathematical Engineering
Advisor: Biondo Biondi
Practicum: Lawrence Livermore National Laboratory
Contact: ermartin@stanford.edu

Sarah Middleton
University of Pennsylvania
Genomics and Computational Biology
Advisor: Junhyong Kim
Practicum: Lawrence Berkeley National Laboratory
Contact: sarahmid@mail.med.upenn.edu

Victor Minden
Stanford University
Computational and Mathematical Engineering
Advisor: Lexing Ying
Practicum: Lawrence Berkeley National Laboratory
Contact: victorminden@gmail.com

Sherwood Richers
California Institute of Technology
Astrophysics
Advisor: Christian Ott
Practicum: Lawrence Berkeley National Laboratory
Contact: srichers@tapir.caltech.edu

Jamie Smedsmo
University of North Carolina
Environmental Modeling
Advisor: Marc Serre
Practicum: Oak Ridge National Laboratory
Contact: jsmedsmo@live.unc.edu

Andrew Stershic
Duke University
Civil Engineering/Computational Mechanics
Advisor: John Dolbow
Practicum: Oak Ridge National Laboratory
Contact: ajs84@duke.edu

Andrew Stine
Northwestern University
Chemical and Biological Engineering
Advisor: Linda Broadbelt
Practicum: Argonne National Laboratory
Contact: andrewstine2015@u.northwestern.edu

Daniel Strouse
Princeton University
Theoretical Neuroscience
Advisor: William Bialek
Practicum: Lawrence Berkeley National Laboratory
Contact: danieljstrouse@gmail.com

Andrew Till
Texas A&M University
Multiphysics Scientific Computational Nuclear Engineering
Advisor: Marvin Adams
Practicum: Los Alamos National Laboratory
Contact: attom@tamu.edu

Dragos Velicanu
Massachusetts Institute of Technology
High Energy Physics
Advisor: Gunther Roland
Practicum: Lawrence Berkeley National Laboratory
Contact: velicanu@mit.edu

Melissa Yeung
California Institute of Technology
Mathematics
Advisor: Mathieu Desbrun
Practicum: Lawrence Berkeley National Laboratory
Contact: myeung@caltech.edu

2ND YEAR FELLOWS

Front, left to right: Alexander Turner, Nick Frontiere, Adam Richie-Halford and Daniel Rey;
Back, left to right: Isha Jain, David Plotkin, David Ozog, Chelsea Harris and Kathleen Alexander
Not pictured: Will Fletcher

Kathleen Alexander
Massachusetts Institute of Technology
Computational Materials Science
Advisor: Christopher Schuh
Contact: katcalex@mit.edu

Will Fletcher
Stanford University
Biophysics
Advisor: Vijay Pande
Contact: will.r.fletcher@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
Contact: c.axen@berkeley.edu

Isha Jain
Harvard University/Massachusetts Institute of Technology
Computer Science and Systems Biology
Advisor: Vamsi Mootha
Contact: ijain@mit.edu

David Ozog
University of Oregon
Computational 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

1ST YEAR FELLOWS

Left to right: Alexander Kell, Ryan McKinnon, Julian Kates-Harbeck, Jonathan Gootenberg, Mukarram Tahir, Kathleen Weichman, Kyle Felker, Aditi Krishnapriyan, Hannah De Jong, Thomas Holoien, Alexander Williams, Adam Riesselman, Danielle Rager, Gerald Wang, Hilary Egan, Jay Stotsky, Thomas Anderson, Thomas Thompson, Morgan Hammer, Joy Yang, Alnur Ali, Adam Sealfon and Jordan Hoffmann

Alnur Ali
Carnegie Mellon University
Machine Learning
Advisor: Zico Kolter
Contact: alnurali@gmail.com

Thomas Anderson
California Institute of Technology
Applied and Computational Mathematics
Contact: tga3@njit.edu

Hannah De Jong
Stanford University
Genetics
Advisor: Euan Ashley
Contact: hnd7@cornell.edu

Hilary Egan
University of Colorado
Astrophysics
Advisor: Jack Burns
Contact: hilary.egan@colorado.edu

Kyle Felker
Princeton University
Applied and Computational Mathematics
Advisor: James Stone
Contact: kfelker@math.princeton.edu

Jonathan Gootenberg
Harvard University
Computational Biology
Advisor: Feng Zhang
Contact: goodband@gmail.com

Morgan Hammer
University of Illinois at Urbana-Champaign
Theoretical Chemistry
Advisor: Sharon Hammes-Schiffer
Contact: m-hammer@onu.edu

Jordan Hoffmann
Harvard University
Applied and Computational Mathematics
Advisor: Chris Rycroft
Contact: jhoffmann@g.harvard.edu

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

Julian Kates-Harbeck
Harvard University
Physics
Advisor: Mara Prentiss
Contact: juliankh@stanford.edu

Alexander Kell
Massachusetts Institute of Technology
Computational Neuroscience
Advisor: Nancy Kanwisher
Contact: alexkell@mit.edu

Aditi Krishnapriyan
Stanford University
Condensed Matter Physics and Materials Science
Advisor: Evan Reed
Contact: a1k2112@gmail.com

Ryan McKinnon
Massachusetts Institute of Technology
Physics
Advisor: Mark Vogelsberger
Contact: ryanmmckinnon@gmail.com

Danielle Rager
Carnegie Mellon University
Neural Computation
Advisor: Valerie Ventura
Contact: drager@andrew.cmu.edu

Adam Riesselman
Harvard University
Bioinformatics and Integrative Genomics
Advisor: Peter Park
Contact: adam.riesselman@drake.edu

Adam Sealfon
Massachusetts Institute of Technology
Computer Science
Advisor: Piotr Indyk
Contact: asealfon@mit.edu

Jay Stotsky
University of Colorado
Applied Mathematics
Advisor: David Bortz
Contact: jay.stotsky@colorado.edu

Mukarram Tahir
Massachusetts Institute of Technology
Materials Science and Engineering
Advisor: Alfredo Alexander-Katz
Contact: mtahir@mit.edu

Thomas Thompson
Harvard University
Geophysics
Advisor: Brendan Meade
Contact: tthompson@fas.harvard.edu

Gerald Wang
Massachusetts Institute of Technology
Mechanical Engineering
Advisor: Nicolas Hadjiconstantinou
Contact: jerry.wang@mit.edu

Kathleen Weichman
University of Texas
Physics
Advisor: Michael Downer
Contact: kweichman@utexas.edu

Alexander Williams
University of California, San Diego
Computational Neuroscience
Advisor: Timothy Gentner
Contact: alex.h.willia@gmail.com

Joy Yang
Massachusetts Institute of Technology
Computational and Systems Biology
Advisor: Eric Alm
Contact: yangjy@mit.edu

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.