The research highlighted in this book is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications.

Visit https://bluewaters.ncsa.illinois.edu/science-teams for the latest on Blue Waters-enabled science and to watch the 2017 Blue Waters Symposium presentations.
A MESSAGE FROM THE NC SA DIRECTOR

I continue to be amazed by the vast range of creative, limit-pushing research that scientists submit to this publication year after year. With the support of the National Science Foundation (NSF) and the University of Illinois, the National Center for Supercomputing Applications’ (NCSA) Blue Waters Project continues to empower scientists to make discoveries that have immense impact in a diverse range of fields, spark new understanding of our world, and open new avenues for future research.

An example of a once in a generation discovery is gravitational wave detection. Even prior to NCSA’s status as an official member of the Laser Interferometer Gravitational-wave Observatory (LIGO) consortium, the Blue Waters supercomputer was used by NCSA’s astronomy and gravity groups, as well as many other researchers, to simulate gravitational wave sources. But verifying Einstein’s Theory of Relativity is not the only cosmological work Blue Waters is performing. For example, Tiziana Di Matteo of Carnegie Mellon University successfully carried out a cosmological simulation that used all of Blue Waters that yielded information “crucial to understanding the formation of the first quasars and galaxies from the smallest to the rarest and most luminous.”

There are also the discoveries that directly impact people’s lives. Discoveries like high-resolution, high-precision elevation maps of the Arctic and Antarctic. These are but a few of the remarkable discoveries you’ll find in the pages that follow. In the past few years, the Blue Waters Project has enabled remarkable work in biology, chemistry, physics, geosciences, cosmology and astrophysics, atmospheric science, and many other fields such as economics and social sciences. Even more remarkable breakthroughs will be forthcoming as NCSA continues to partner with scientists around the nation to change the world as we know it.

Dr. William “Bill” Gropp, Director
National Center for Supercomputing Applications
SCIENCE NOT POSSIBLE ANYWHERE ELSE

Every day, the Blue Waters Project makes previously impossible investigations and insights possible for researchers across the United States. Every day, the combination of massive computing power and the intellectual might of pioneering scientists and engineers creates opportunities for us to better understand and shape our world. While we are more than four years into full-service operations for Blue Waters and its associated support, training, and education efforts, time hasn’t dulled the sense of wonder and delight I feel when I learn of another door that we opened to discovery.

This past year, a sample of 31 science teams that have used Blue Waters were surveyed and interviewed as part of a report meant to judge the effectiveness and productivity of Blue Waters. Using information gathered in the surveys, the report’s authors at International Data Corporation’s HPC Division, now known as Hyperion Research, ranked the impact of each team’s findings into an “innovation index”—using a methodology they developed to analyze the effectiveness of 700 plus scientific projects, including international HPC projects. The IDC/Hyperion Research analysts noted in the report that “NSCA did an unusually thorough job of preparing [science teams] for Blue Waters.” In fact, a notably higher percentage of scientific innovations enabled by using Blue Waters qualified for IDC’s highest impact levels than was true for IDC’s global database of supercomputer-enabled scientific innovations. You can find the entire IDC report on the Blue Waters portal (bluewaters.ncsa.illinois.edu).

The Blue Waters Project has always been about much more than just providing computing cycles, though that will still be more memory than any other system in the open research space and storage bytes. While such high performance resources are necessary for success, they are not sufficient for effectiveness. In other sections of this report, you will read about our expanded Petascale Application Improvement Discovery (PAID) program, where the Project provided millions of dollars to science teams and computational and data experts to improve the performance of applications (in a measurable manner). Likewise, the Project is proud to play a role educating and developing the next generation extreme-scale workforce, through our workshops, symposium, graduate fellowships, undergraduate internships, the Virtual School for Computational Science and Engineering, our funding for the HPC University and our workshops, and training allocations.

As you read through the reports on the following pages, you will notice badges on reports to signify use of some of the systems’ special features that make Blue Waters the National Science Foundation’s most powerful system for computation and data analysis. These capabilities include: data-intensive problems that use large numbers of files, large disk space/ bandwidth, or automated workflows/off-site transfers; GPU-accelerated (runs faster on XK nodes than on XE nodes); thousand-node (or “highly scalable”), i.e. scales to at least 1,000 nodes for production science input; memory intensive (used at least 50 percent of available memory on 1,000-node runs); multi-physics/multi-scale - spans multiple length/time scales or physical/ chemical processes; machine learning (employs deep learning or other techniques, including "big data"); multi-physics/multi-scale - requires high bandwidth/ low-latency interconnect for frequent tightly coupled messaging); and industrial application (private sector collaborators or results directly applicable to industry).

And of course, many reports carry the badge indicating “Only on Blue Waters” to signify that this research would not have been possible on other currently deployed U.S. open-science resources.

You will also notice this report, our fourth, contains 18 percent more high-impact result summaries than the 2016 report, which itself was a 20 percent increase over the 2015 report, which was a 25 percent increase over the 2014 report. The fact that in four years of providing one of most effective leadership-class systems in the world, Blue Waters is making such a remarkable difference to fields from physics, biomedicine, geo and space science, biology, economics, social and political science, and big data analysis continues to be a tribute to the vision of NSF, to our exceptional science and research teams, and to the dedicated Blue Waters Project staff.

While the scientific insights gained with support from Blue Waters are perhaps the most exciting way to judge the Project’s impact, there are ways to consider its unduplicated role in the open-science community. To date, the Project has provided over 20 billion core-hour equivalents. Looking through another lens, Blue Waters is supporting approximately $300 billion of funded research—projects supported not just by the National Science Foundation but also by NIH, NASA, DOE, NOAA and other agencies. As previously noted, these funded investigations might not even be possible (and certainly would take much longer) without Blue Waters.

One of these pioneering scientists reliant upon Blue Waters’ massive compute power was Klaus Schulten, who passed away very unexpectedly October 31, 2016. Dr. Schulten blazed the trail in using what he called the “computational microscope,” devoting over 40 years to “establishing the physical mechanisms underlying processes and organization in living systems from the atomic to the organism scale.”

Not only was Dr. Schulten a leader in computational biophysics, he was a passionate supporter of the National Center for Supercomputing Applications (NCSA), Blue Waters, and the need for a robust national program of leadership class computing for academic, even with Blue Waters he did not have enough computing power to realize his research dreams. While Dr. Schulten is dearly missed by all of the Blue Waters Project staff, his computational microscope lives on as his work transitions to other principal investigators.

I take great pride in the Blue Waters Project and in the outstanding science and Project teams that make Blue Waters an exceptional resource. Talented, dedicated individuals at NCSA and our partner institutions strive night and day to maintain the system, support our science and engineering partners in achieving breakthrough research, improve tools and processes, and train the next-generation of computational researchers. This publication celebrates their accomplishments as well as those of the researchers, educators, and students they support.

With great respect and thanks to be working with everyone,

Dr. William T.C. Kramer
Blue Waters Project Director and Principal Investigator
Research Professor of Computing Science
workload analysis reveals how blue waters is actually used

As the National Science Foundation’s flagship “leadership-class” supercomputer, Blue Waters was designed to enable the national scientific and research community to solve “grand challenge” problems that are orders of magnitude more complex than can be done on other systems. The great majority of Blue Waters machine time is awarded directly by the NSF in large allocations of a million node-hours or more, but the machine serves all disciplines through a diverse community of 800 to 1,000 users and 120-130 projects from 47 states. Given the important and unique role that Blue Waters plays in the U.S. research portfolio, it is important to have a detailed understanding of its actual usage by the scientific community as revealed by objective measurement and analysis. NCSA staff recently collaborated with the Center for Computational Research of the University at Buffalo, SUNY to produce a 106-page workload analysis of Blue Waters, available at https://arxiv.org/ftp/arxiv/papers/1703/1703.00924.pdf

Blue Waters is Instrumented on Many Levels

Analysis requires data, the collection of which has been designed into all of the components of Blue Waters from the beginning. Great care has been taken to avoid any impact of performance and utilization monitoring on the performance or stability of the science applications running on the machine. Also, the design of the system monitoring infrastructure on Blue Waters focuses on acquiring a great deal of detail about performance at every level in ways that do not require the application teams to do any work or add any instrumentation or performance profiling tools to the codes.

An important data collection and management tool used to do this analysis is Open XDMoD (xdmod.ncsa.illinois.edu) [1], which provided comprehensive statistics on the number and type of computational jobs run, resources (computation, memory, disk, network, etc.) consumed, job wait times, quality of service, and, of particular importance for this work, detailed job-level performance data aggregated from a variety of sources.

Job information is provided by the Torque resource manager, which operates in coordination with the Cray ALPS resource manager and the MOAB job scheduler. Log entries are created when a job passes through various states of execution, including submission, changes of scheduler status (queued, held, deleted, eligible to run, start/running, etc.), and completion. Job completion is accompanied by exit code status that indicates whether the job completed normally.

The Lightweight Distributed Metric Service (LDMS, https://orvis.ca.sandia.gov) [2] is used to collect metrics on load average, memory usage, limited filesystem data transfers, network utilization with one-minute granularity at the compute node level and Gemini NIC/router level. For the XK nodes, GPU utilization and GPU memory utilization are also collected. LDMS also collects machine-specific register (MSR) data from the CPUs such as the number clock ticks, the number of instructions retired and the number of floating point operations performed.

Darshan [http://www.mcs.anl.gov/research/projects/darshan/] [3] captures application MPI I/O behavior, including properties such as patterns of access within files, with minimum overhead. Darshan data summarized for each job includes total bytes written and read, the number of opened files, the time spent in file system I/O operations, and the number of opened files in HDF5 and NetCDF formats.

APRUN logs provide the primary means to determine what application(s) was executed during a job, including the executable path, the number of nodes requested, and the layout of the tasks on the cores of the nodes. This information is used to map the job to an application algorithm, and, in conjunction with the MSR data, to determine concurrency/parallelism.

Finally, the ALTD and XALT [4] tools track which statically and dynamically linked libraries are used by each job.

All-in-all about eight billion data points and event records are collected each day about how Blue Waters is used, making not only the most productive system in the world, but also possibly the most measured system.

Workload Analysis Requires a Supercomputer

The workload analysis itself was a challenging computational problem – requiring more than 35,000 node hours (over 1.1 million core hours) on Blue Waters to analyze roughly 95 TB of input data from over 4.5M jobs that ran on Blue Waters during the recent past. In the process, approximately 250 TB of data across 100M files was generated. This data was subsequently entered into MongoDB and a MySQL data warehouse to allow rapid searching, analysis and display in Open XDMoD. A workflow pipeline was established so that data from all future Blue Waters jobs will be automatically ingested into the Open XDMoD datawarehouse, making future analyses much easier.

Physics and Biology Dominate, but Long Tail Grows

The Mathematical and Physical Sciences and Biological Sciences NSF directorates are the leading consumers of cycles on Blue Waters, typically accounting for more than 2/3 of all node hours used. However, the number of fields of science represented in the Blue Waters portfolio has increased in each year of its operation – more than doubling since its first year of operation – and the historical trend for all NSF directorates is toward increasing use of Blue Waters, albeit at much different overall scales of utilization. The applications run on Blue Waters represent an increasingly diverse mix of disciplines, ranging from broadly-used community codes to specialized codes for specific scientific sub-disciplines. Common algorithms, as characterized by Phil Colella’s original “seven dwarfs” algorithm classification, are roughly equally represented within the applications run on Blue Waters.
from unstructured grids and Monte Carlo methods, which exhibit a smaller fraction of use. Optimized numerical libraries, essential for high performance, are used throughout areas in which such libraries are applicable. The top ten applications consume about 2/3 of all node hours, with the top five applications (NAMD, CHROMA, MILC, AMBER, and CACTUS) consuming about 1/2.

THROUGHPUT AND CAPABILITY COMPUTING CO-EXIST

Blue Waters supports a diverse mix of job sizes from single node jobs to jobs that use in excess of 20,000 nodes in a single application. The patterns of usage differ between the XE (CPU-only) and XK (GPU-accelerated) nodes. Single node jobs, some of which may be attributable to high throughput computing, represent less than 2% of the total node hours consumed on Blue Waters.

For XE node jobs, all of the major science areas using more than 1 million node hours run a mix of job sizes and all have very large jobs (> 4,096 nodes/110,072 core equivalents). The relative proportions of job size vary between different parent science areas. The job size distribution weighted by node hours consumed peaks at 1,025 – 2,048 nodes for XE jobs. Impressively, the largest 3% of the jobs (by node hours used) account for 90% of the total node-hours consumed.

The majority of XE node hours on the machine are spent running parallel jobs that use some form of message passing for inter-process communication. At least 25% of the workload uses some form of threading, however the larger jobs (> 4,096 nodes) mostly use message passing with no threading. There is no obvious trend in the variation of thread usage over time the data was available.

USE OF GPU ACCELERATION INCREASES OVER TIME

The XK (GPU) nodes are as heavily utilized as the XE nodes, and the number of science teams using them has steadily increased, but remain a relatively small number of applications can use them well. Because the XE and XK usage is charged the same rate, and we now can measure GPU usage, we know the teams that are running on the XK nodes used the GPUs and also decide they have a performance improvement over the XE nodes. The parent sciences Molecular Biosciences, Chemistry, and Physics are the largest users with NAMD and AMBER the two most prevalent applications.

The job size distribution weighted by XK node hours consumed peaks at 65 – 128 nodes for the XK jobs. Similar to the XE nodes, the largest 7% of the jobs (by node-hour) account for 90% of the node-hours consumed on the XK nodes.

The aggregate GPU utilization (efficiency) varies significantly by application, with MELD achieving over 90% utilization of the GPU and GROMACS, NAMD, and MILC averaging less than 30% GPU utilization. However, for each of the applications, the GPU utilization can vary significantly from job to job and within code steps. It is important to note the applications with lower GPU utilization are still getting performance and time to solution improvements and the application design and problems do not expect higher utilization.

Despite the increases in the number of applications using GPUs, it is important to understand that there are still many applications that do not, and maybe should not, use GPU technology without completely redesigning their methods – which could be a tremendous amount of work.

CPU-ONLY NODES ENABLE PETABYTE-MEMORY CALCULATIONS

Most jobs that run on the XE nodes use less than 50% of the memory available on the node. However, the distribution of memory use has a substantial tail to higher memory usage, as shown in Figure 1, representing capability-class jobs exploiting the 1.38 PB of aggregate memory available on the XE nodes. Most jobs that run on the XK nodes use less than 25% of the available memory in the node with a short tail to higher usage. GPU memory usage is very small with few jobs using more than 1GB per GPU. The XE and XK nodes show no historical differences in memory use from year to year. For almost all applications and parent fields of science, memory usage has not changed over time under study.

GPU-only nodes enable petabyte-memory calculations. Most jobs that run on the XE nodes use less than 50% of the memory available on the node. However, the distribution of memory use has a substantial tail to higher memory usage, as shown in Figure 1, representing capability-class jobs exploiting the 1.38 PB of aggregate memory available on the XE nodes. Most jobs that run on the XK nodes use less than 25% of the available memory in the node with a short tail to higher usage. GPU memory usage is very small with few jobs using more than 1GB per GPU. The XE and XK nodes show no historical differences in memory use from year to year. For almost all applications and parent fields of science, memory usage has not changed over time under study.

Figure 1: 2D binned scatter plot showing the total peak memory usage for XE node jobs. The color coding shows the total wall-hours for the jobs in each bin using a log scale. The diagonal line indicates the points corresponding to 64 GB per node. The points that are to the right of 11,100 nodes and above 800 TB indicate jobs that could not be executed on any other open system in the US or world.

STORAGE SYSTEM HANDLES DIVERSE LOAD

On average, Blue Waters’ three filesystems (home, projects, and scratch) have a balanced reads/writes ratio with large fluctuations. The volume of traffic on the largest filesystem (scratch) peaks at 10PB per month.

Partners’ jobs exhibit a wide range of I/O patterns. Overall there is a tendency to use a very large number of small files, which is a challenge for many parallel file systems and different form the projections that file size increases and system computational capability grows. Read and write rates stay significantly below possible peak filesystem performance, which actually is good and desired.

Partner jobs spend a very small fraction of time in filesystem I/O operations (0.04% of runtime for 90% of jobs). Many jobs utilize specialized libraries for their I/O operations (about 20% use MPI-IO, HDF5, or NetCDF).

HARD DATA GUIDES OPERATIONAL DECISIONS

Beyond enabling a comprehensive and evidence-based report on the workload and utilization of the NSF’s largest supercomputing resource, the greater legacy of the Blue Waters monitoring and data analysis effort will be improved scientific productivity for users of current and future leadership-class supercomputers. The Blue Waters team at NCSA has real-time and historical data available at their fingertips, allowing system operators and SEAS support staff to identify and adapt to unanticipated new usage patterns and to identify opportunities to work with science teams to improve the performance and throughput of their jobs. Lessons learned from the workload study, such as demand for aggregate and per-node memory, filesystem bandwidth, and GPU accelerators, will inform the design of the Blue Waters follow-on systems called for in the 2017 NSF solicitation “Towards a Leadership-Class Computing Facility – Phase 1.”
INTRODUCTION

The PAID program began by first establishing a team of experts in various technology solutions, who were referred to as Improvement Method Enablers (IMEs). The technology areas were established by reviewing the results of a less structured and more self-directed program that funded work on improvements to existing applications by NSF PRAC teams. The five general categories selected from that first-year opportunity were: task mapping and load balancing, scalable I/O and hierarchical data formats (HDF), Fourier transforms (FFT's), programming models best practices, and exploitation of graphics processing unit (GPU) accelerators. Eight expert teams were formed and led by researchers from five different universities. NVIDIA supported participation in PAID by providing one of their lead application analysts to work with one of the GPU improvement teams. Each IME team developed an individual statement of work that documented the expected number of PRAC teams they would work with, participation in Blue Waters workshops and other training events, and any resulting products from the program such as libraries or best practices. Each IME was assigned a Blue Waters’ staff member as their Point of Contact (PoC) to facilitate tracking progress, reviewing milestones and deliverables, and addressing any support-related issues.

The participating application teams were NSF-selected PRAC teams with an award end date far enough in the future to allow for adequate time left in their allocation period so they could reap the benefits from any improvements to their application while still running their production science on Blue Waters. In total, 15 PRAC teams from seven fields of science engaged with the eight IME teams; some application teams worked with up to three different IME teams. Both the IME’s and PRAC teams received funding from the Blue Waters Project and also as assistance from the Blue Waters Science and Engineering Applications Support team. Each PRAC team–IME team pairing developed individual work plans or statements of work with the assistance of a Blue Waters PoC. All statements of work required baseline performance as a first deliverable to enable quantitatively assessing the impact of the PAID program. This baseline was used to ascertain and quantify improvements in subsequent progress milestones and associated deliverables (usually reports). In order to assure progress, the Blue Waters Project hosted monthly progress review meetings for all PRACs and IMEs. The PRAC teams also presented a public summary of their PAID findings at the annual Blue Waters symposium in order to inform other Blue Waters participants of potential approaches they could adopt in their own applications. The PoC for each science team was responsible for tracking progress, reviewing milestones and deliverables, and addressing any support-related issues. The PAID report in the 2016 Blue Waters Annual Report documented in more detail work done, work in progress, or to be done by the IMEs and the science teams.

The entire PAID program was continuously assessed by an independent team lead by Dr. Lianne DeStefano. These assessments were used to make adjustments in the program as it progress and also to do summative assessments of the impact of the program.

RESULTS

The ultimate results of each partnership of a PRAC team with an IME team is a documented level of time to solution performance improvement that clearly translates to improved science throughput for the PRAC team. In most cases, the baseline performance indicates the performance of a code that was being used by the science teams for their major science projects and the PAID program allowed for potentially substantial gains in performance. In some situations, such as for established community codes, baseline performance designates an application that is already sufficiently optimized, such as a Compute Unified Device Architecture (CUDA)-accelerated application already tuned for Fermi GPUs (the version of GPU prior to Kepler) that runs on the newer K20x GPUs on Blue Waters. Percentage performance improvements should be expected to be lower compared to the upside potential for a code being ported to the GPU for the first time. Additionally, performance may be the final result but enabling new functionality is as important as traditional performance for existing methods for the teams benefiting from the new methods enabled in the code.

“The PAID program gave us access to computing specialists in GPU and HDF5 optimizations . . . which . . . provided us with rapid insights into our code performance. The insights provided by the PAID technical experts led to faster resolutions and improved outcomes for important code improvements that are needed to support the high-frequency simulations by our science community.”

- Thomas Jordan, University of Southern California

PERFORMANCE IMPROVEMENTS

Blue Waters provides the ideal platform for porting production codes to the NVIDIA GPUs while continuing to conduct scientific campaigns with its balanced mix of CPU nodes and GPU nodes. Several of the applications were ported to OpenACC or CUDA, or had already-existing OpenACC or CUDA code tuned to make better use of the NVIDIA K20x GPUs. Many projects teamed up with Wen-mei Hwu’s IME to enable or improve the use of the NVIDIA GPUs on Blue Waters. The IME team worked on the CyberShake code AWP-ODC of the Thomas Jordan PRAC team to get a 1.1x speedup from a message passing interface (MPI) code that was already ported to CUDA with some additional porting of code not running on the GPU. The Hwu IME worked on the ChNaGq astrophysics code with Tom Quinn’s PRAC team to get an impressive time to solution 1.6x speedup from Charm++.
enabled code tuned for the Fermi generation of NVIDIA GPUs at scales of several thousand nodes. An OpenACC example of the success of this work is the 2.8x improvement in time-to-solution for the MS-FLUKSS code of the Nikolai Pogorelov PRAC team obtained by working with the Programming Model team led by Bill Tang for a single Blue Waters GPU node (XK) over a single CPU node (XE). Another OpenACC example is the 3.9x speedup obtained with the 3D-FDTD code of Jamesinha Simpson for a single Blue Waters GPU node (XK) over a single CPU node (XE). Both use MPI and the work to implement the new programming model into the distributed code remains to be completed. The results obtained from the OpenACC work is in line with earlier results showing that MPI codes using XK (GPU) nodes have typically 2x to 3x faster time-to-solution than their XE (CPU) counterparts, node-for-node, for complete applications.

The Blue Waters high-speed network uses a highly scalable communication, 3D torus topology interconnect subsystem, but some applications needed improvements to ensure their communication pattern does not cause congestion due to poor MPI task placement. The DME team led by Sanjay Kale worked with PRAC several teams to understand application communication patterns and develop a new utility to automatically provide a better MPI task placement strategy using a library the team developed that is discussed below. The library collects the necessary information during a shorter instrumented run and then produces an MPI rank reorder file that works with the Cray MPI library to place the tasks optimally for on-node communication and for reduced off-node congestion. A 2.2x and a 1.2x speedup in time-to-solution for several applications that used the library range from 1.2x for P.K. Yeung’s 3D pseudo-spectral DNS code to a 2x speedup for Bob Sugar’s MILC-based su3_rhmc application without having to know detailed knowledge of the communication graph. TopoMapping is available on Blue Waters and is documented on the Blue Waters portal at https://bluewaters.ncsa.illinois.edu/topology-considerations/topomapping.

In addition, a new I/O library called meshio (available at https://github.com/ukoshkos/meshio) was developed by the I/O IME team led by Bill Gropp. The library improved I/O performance by more than 20x with a MILC I/O pattern and by 19x for the Plascomcm application, reducing I/O time from 95% of the runtime to 5%. Finally, the SpiralGen FFT library range from 1.2x for P.K. Yeung’s 3D pseudo-spectral DNS code to a 2x speedup for Bob Sugar’s MILC-based su3_rhmc application without having to know detailed knowledge of the communication graph. TopoMapping is available on Blue Waters and is documented on the Blue Waters portal at https://bluewaters.ncsa.illinois.edu/topology-considerations/topomapping.
Blue Waters education allocations are available to the national community to support the development of a national workforce with expertise in petascale computing. Blue Waters has committed one percent (the equivalent of 60 million core-hours) of the system’s computational capacity each year for educational projects, including seminars, course-ware development, courses, workshops, institutes, internships, and fellowships.

Education and training projects utilizing these allocations have ranged from one-day workshops, to week long institutes, to semester courses, as well as support for internships and fellowships. Applications for education allocations are accepted throughout the year. Applicants are asked to apply at least a month in advance of when the allocations are needed, to allow time to process the request including setting up accounts on the system.

But the Blue Waters Educational, Outreach and Training allocations are more than just computational time and storage. A project can also get special scheduling and reservations so they can do in-class, interactive exercises and assignments. Blue Waters also provides technical support to instructors, and Blue Waters staff that serve as "Points of Contact" for these projects.

To date, there have been more than 200 approved education, outreach, and training projects from organizations across the country. These allocations have directly benefitted over 3,700 individuals learning about different aspects of computational and data-enabled science and engineering at more than 160 institutions, including 41 institutions in EPSCoR jurisdictions and at 14 Minority Serving Institutions.

A sampling of the projects and programs that have benefitted include: Blue Waters Graduate Fellows; Blue Waters Interns; Parallel Computing Summer Research Internships sponsored by Los Alamos National Laboratory; Southern California Earthquake Center (SCEC) undergraduate studies for understanding seismic hazards; NASA’s Oklahoma EPSCoR Summer Research Program; University of Illinois CS420/CSE402/ECE492 Parallel Programming course for advanced undergraduates and early graduate students; and the recent Scaling to Petascale Institute that engaged over 500 people participating from over 23 locations and via YouTube Live.

Each project is requested to complete a final report documenting the impact of the education allocation. Below are examples of the benefits of these allocations to the participants as reported to us by the project PIs:

• “The Blue Waters system gave the students access to a system with much larger scaling capabilities than they typically have access to. In particular, one team was able to scale their code to hundreds of nodes, which would be impossible at their current institutions.”

• “Blue Waters in particular supported our efforts to replicate at-scale production environments and the opportunities and challenges these present to users.”

• “The sheer scale of the resource is making it possible for us to do combined analysis on more data sets than has been possible on any other resource. This makes us especially anxious to implement parallelism to see how much speed-up can be accomplished with improved approaches.”

• “Students had a context where they could learn about the significance of parallelism and that there are problems in nature such as coupled n-body problems that are ideally suited for parallel programming.”

• “This project introduced many of the students to high performance computing for the first time—setting up batch scripts, requesting interactive nodes, running remote simulations, and using gpu-accelerated programs.”

• “The computing resources provided by Blue Waters enabled many of the students to carry out their projects for the course in which they used simulations to elucidate a biophysical problem related to their own research.”

• “The use of Blue Waters enabled students to understand issues of programming at scale (up to 128 nodes [4,096 core equivalents]) and to tune their codes in a dedicated environment where performance is roughly repeatable.”

If you have undergraduates or professionals who are learning about high performance computing, or experienced graduate students and researchers who want to expand their petascale knowledge and skills, the Blue Waters project welcomes requests for education allocations for preparing current and future generations to advance discovery in all fields of study.

Additional information about education allocations is available at https://bluewaters.ncsa.illinois.edu/education-allocations. If you have questions please feel free to contact the Blue Waters staff at bw-eot@ncsa.uiuc.edu.
Simulating Two-Fluid MHD Turbulence in Star-Forming Molecular Clouds on the Blue Waters System

Supermassive Black Holes at the Cosmic Frontier

Theoretical Astrophysics and Data Analysis

Detection of Gravitational Wave Sources in Dense Stellar Environments

Deep Neural Networks to Enable Real-Time Multi-messenger Astrophysics

Toward Robust Magnetohydrodynamic Simulations of Galaxy Cluster Formation

Exploring the Nature of Exploding Massive Stars with High Resolution

Modeling Plasma Flows with Kinetic Approaches using Hybrid CPU-GPU Computing

Three-Dimensional Nature of Collisionless Magnetic Reconnection at Earth’s Magnetopause

Transformative Petascale Particle-in-Cell Simulations

Mini-disk Dynamics about Supermassive Black Holes Binaries

Realistic Simulations of the Intergalactic Medium: The Search for Missing Physics

Simulating Galaxy Formation Across Cosmic Time

3D General-Relativistic Radiation-Hydrodynamic Simulations of Core-Collapse Supernovae

Modeling Physical Processes in the Solar Wind and Local Interstellar Medium with a Multiscale Fluid-Kinetic Simulation Suite

Unified Modeling of Galaxy Populations in Clusters

Kinetic Simulations of Large-Scale Plasma Turbulence

GPU-Accelerated Adaptive Mesh Refinement

Magnetorotational Collapse of Supermassive Stars, Black Hole Formation, Gravitational Waves, and Jets

GPU-Accelerated Simulations: Black Holes, Spaghettified Stars, and Tilted Disks

Advanced Space Weather Modeling

3D Simulations of i-Process Nucleosynthesis
EXECUTIVE SUMMARY

We are at the threshold of a new data-rich and simulation-rich era in star-formation studies. The question of how stars form is fascinating in itself and has a great impact on several other areas of astrophysics. There is a general consensus that a predominant amount of star formation in our galaxy takes place in molecular clouds, and specifically in giant molecular clouds (GMC). The molecular cloud material consists of a highly turbulent, very weakly ionized, strongly magnetized plasma. It is within such a plasma that we observe protostellar cores—the progenitors of the stars that will eventually form. The present project is designed to simulate the two-fluid turbulence that regulates star formation. This turbulence, and the nonideal physics it gives rise to, is fundamentally different from single-fluid magnetohydrodynamic (MHD) turbulence. The project aims to study these differences.

METHODS & CODES

The core MHD algorithms in our RIEMANN code are based on higher-order Godunov schemes. Balsara and his collaborators have been on the forefront of the effort to develop high-accuracy schemes for computational astrophysics in general and computational MHD in particular. Two-fluid methods have been described in [6–8].

RESULTS & IMPACT

This project is newly funded and is in its initial stages where large-scale simulations have been planned and are ongoing on Blue Waters. Several papers have been published by our group using lower-resolution simulations [1–3]. The new work will be a substantial improvement on our previous work in terms of resolution as well as in the details of input physics and accuracy of simulation code.

WHY BLUE WATERS

Balsara’s group has also simulated at petascale on Blue Waters via the Great Lakes Consortium for Petascale Computation. We are, therefore, extremely familiar with the Blue Waters system. The PIs are also funded via NSF grant DMS-1622457 to develop computational capabilities for turbulent simulations in computational astrophysics. This newly funded proposal will provide us with the impetus for developing petascale-ready simulation tools for astrophysical turbulence and making them freely available to the greater astrophysics community.

PUBLICATIONS AND DATA SETS

SUPERMASSIVE BLACK HOLES AT THE COSMIC FRONTIER

EXECUTIVE SUMMARY

The BlueTides (BT) cosmological simulation is unique because it directly probes the range of scales (masses and epochs) of galaxies that are likely to be discovered in the near future (for example, by the James Webb Space Telescope (JWST), the successor to the Hubble). In this project, we have extended the BT simulation to cover the evolution of the first billion years of cosmic history (BlueTidesII). At this epoch, there are additional, already-existing observations that tell us that the first billion-solar-mass black holes (also known as quasars) had already formed. It is important to make contact with observations of quasars, which have not been discovered at these early stages of cosmic history. The extreme early growth depends on the early interplay of high cosmic gas densities and low tidal fields that shape the mode of accretion and allow this first generation of massive black holes to form at the most rapid pace. We have tracked the descendents of the first supermassive black holes in today’s universe by running a new BT MassTracer simulation that has run on the full Blue Waters, the BlueTides II cosmological simulation. The run was made possible through our new code, Massively Parallel (MP)-Gadget. The simulation has proven to be crucial to understanding the formation of the first quasars and galaxies from the smallest to the rarest and most luminous, and their effect to reionization. The simulation is being used to make predictions about what the WFIRST and JWST (expected to launch in 2018) will see. The largest next-generation telescopes (ALMA, JWST, WFIRST) currently planned aim to study the “end of the dark ages,” an epoch in the early universe when the first galaxies and quasars form and reionization of the universe takes place. The fundamental challenge to understanding this epoch is that extremely large volumes need to be simulated while at the same time extremely high resolution is required. Our BT run was made possible through our new Cosmological Cosmological Simulation of the High-Redshift Universe (BlueTidesII). The goal is to increase significantly the scientific impact of the BT run on Blue Waters is the first and only cosmological simulation of structure formation that has run on the full Blue Waters machine.

In the last year, we have extended the BT simulation with an unprecedented volume and resolution to cover the evolution of the first billion years of cosmic history (BlueTidesII). The goal is to increase significantly the scientific impact of the BT run on Blue Waters is the first and only cosmological simulation of structure formation that has run on the full Blue Waters machine.

RESEARCH CHALLENGE

Computational cosmology, simulating the entire universe, represents one of the most challenging applications for the era of petascale and beyond computing resources. We have successfully carried out a full-machine run on Blue Waters, the BlueTides II cosmological simulation. The run was made possible through our new code, Massively Parallel (MP)-Gadget. Recent radical updates to the code efficiency, and also to the smoothed-particle hydrodynamics formulation and star formation modeling, mean that we are able to meet the challenge of simulating the next-generation space telescope fields and effectively use the full Blue Waters machine. We have extended the BT run, which has an unprecedented volume and resolution, to cover the evolution of the first billion years of cosmic history.

RESULTS & IMPACT

We have used BT to model the environment of the most massive black hole and host galaxy (top row) compared to Figure 1: The environment of the most massive black hole and host galaxy (top row) compared to the most massive black hole and host galaxy (bottom row) at different scales. The orange ellipse and arrow illustrate the tidal field components in the two cases. Tidal fields stretch material along one direction and compress material along the perpendicular.

We have run BT forward toward the epoch of observed quasars. The growth of the most massive black holes in the early universe, consistent with the detection of highly luminous quasars when the universe is less than a billion years old, implies sustained, critical accretion of material to grow and power them. It is still uncertain which conditions in the early universe allow the fastest black hole growth. Large-scale hydrodynamical cosmological simulations of structure formation allow us to explore the conditions conducive to the growth of the earliest supermassive black holes.

We used BlueTides to follow the earliest phases of black hole critical growth. At a few hundred thousand years after the Big Bang, we have shown that the most massive black holes approach 400 million solar masses. Examining the large-scale environment of hosts, we find that the initial tidal field is more important than overdensity in setting the conditions for early black hole growth. In regions of low tidal fields gas accretes “cold” onto the black hole and falls along thin, radial filaments straight into the center, forming the most compact galaxies and most massive black holes at earliest times. We have pushed the evolution of this large volume some 50 million years further and we have seen that this initial exponential growth of the first supermassive black holes is quenched even at redshift 7. This is exciting: we are starting to see the evidence of so-called feedback, as there are already hints in observations that this process is in action even in the first quasars.

To explore the question of where the most massive early-forming quasars are today, we have run a dark matter-only cosmological simulation that has the same volume and initial conditions of the BTMassTracer. The lower resolution allowed us to run all the way to the present universe. The results are extremely interesting: the most massive black holes that form in the early universe are not the most massive ones today and are not found in the most massive galaxy clusters. With BTMassTracer, we have traced the descendents of this first generation of supermassive black holes and found that they are in moderately massive galaxies today. In some way, the first black holes have a very fast initial growth phase, but other black holes that assemble later are likely to grow in higher-density regions. The first massive black holes are not in very special places in the overall density field of the universe.

WHY BLUE WATERS

From the simulation/theoretical perspective, large-scale uniform volume hydrodynamical simulation of the high-redshift universe is a problem that is perfectly suited to the largest modern petascale facilities such as Blue Waters. This is because it is now feasible to run memory-limited simulations with the resources that computer time allocation panels are able to grant, and so reach unprecedented volumes and resolutions.

PUBLICATIONS AND DATA SETS


THEORETICAL ASTROPHYSICS AND DATA ANALYSIS

EXECUTIVE SUMMARY

We have been investigating different avenues for mapping dark matter in the universe, on both the very largest scales (extending to the cosmological horizon) and the smallest scales (<1,000x smaller than the typical size of galaxies).

We have produced the highest-fidelity map of dark matter on cosmological scales using data from the South Pole Telescope, a 10-meter (m) telescope at the Geographic South Pole that is mapping the cosmic background microwave temperature and polarization fluctuations.

We have also started a search for clumps of dark matter on small scales in galaxies that are acting as gravitational lenses of background star-forming galaxies. We have previously detected one such clump in a galaxy and we are now extending this search to a larger sample of galaxies.

RESEARCH CHALLENGE

Roughly 80% of the mass in the universe is in the form of dark matter, which appears to be a particle beyond the standard model of particle physics. It barely (if at all) interacts with particles that make up the visible matter in the universe, on both the very largest scales (extending to the cosmological horizon) and the smallest scales (>1,000x smaller than the typical size of galaxies).

If dark matter is not actually “matter” at all but is instead a first crack in the theory of general relativity, it could be expected that the very largest scales may show anomalous behavior, while interesting self-interactions or other possible properties of dark matter could be expected to affect how it clumps together on the smallest scales.

METHODS & CODES

We used the gravitational deflection of light to map out where the mass is in two different regimes, using either the cosmic microwave background fluctuations as the light source or star-forming galaxies observed at millimeter (mm)-wavelengths.

Using the cosmic microwave background, which is the most distant light we can measure in the universe, we make mass maps on extremely large scales by looking for subtle systematic correlations in the temperature and polarization fluctuations [1]. The South Pole Telescope, a 10-m mm-wave experiment at the Amundsen–Scott South Pole Station is mapping the cosmic microwave background at high resolution and sensitivity providing the experimental data for this work. Gravitational lensing changes the local statistical properties of the maps in a way that is sensitive to the local mass density, so we use custom-made second-order statistical estimators to make a map of the mass density.

We used the gravitational deflection of light to map out where the mass is in two different regimes, using either the cosmic microwave background fluctuations as the light source or star-forming galaxies observed at millimeter (mm)-wavelengths.

RESULTS & IMPACT

Using Blue Waters, we were able to run Monte Carlo simulations of the South Pole Telescope data analysis pipeline at a level that allowed us to make the map shown in Fig. 2. Over 2,500 square degrees, we have a map of how much mass there is at every point on the sky. “Lensing convergence” is a weighted average of the mass density along the line of sight to the source, and since the source (the cosmic microwave background) is at the edge of the observable universe, we get a complete census of the mass in the universe. This map is being used to compare with galaxy surveys to better understand the relationship between visible mass and dark matter. Going forward, the South Pole Telescope has a new camera that is roughly 10x faster at mapping the sky, which will enable both higher precision and higher resolution in this mass map within a year or two.

The search for small-scale substructure in the dark matter is ongoing. With the data currently being analyzed, we expect to be able to confirm or rule out dark matter candidates that have either a large thermal velocity (“warm dark matter”) or dark matter composed of ultra-light particles that have a quantum mechanical fuzziness that extends to macroscopic scale (“fuzzy dark matter”).

WHY BLUE WATERS

The combination of computing power and support staff for Blue Waters made this the premier option for performing this research.
DETECTION OF GRAVITATIONAL WAVE SOURCES IN DENSE STELLAR ENVIRONMENTS

EXECUTIVE SUMMARY

The direct detection by the twin LIGO (Laser Interferometer Gravitational-Wave Observatory) detectors of gravitational waves from merging black holes has ushered in a revolution in astrophysics [1,2]. LIGO's detection campaigns are enabling the construction of catalogs of gravitational wave sources to enable accurate censuses of the mass and angular momentum distribution of black holes and neutron stars. These studies will provide new and detailed information about the formation, evolution, and environments in which compact objects reside. In particular, the detection of gravitational waves from eccentric compact binaries will provide the cleanest signature of compact object populations in dense stellar environments [3].

To detect these events, we introduce the first waveform model in the literature that reproduces the features of eccentric compact binary coalescence throughout merger. To validate this model, we used numerical relativity simulations performed with the Einstein Toolkit on the Blue Waters supercomputer.

RESEARCH CHALLENGE

Gravitational wave (GW) observations provide insights into the astrophysical sources that generate them and can be used to map the structure of spacetime in the vicinity of compact binary systems [4]. To construct the merger phase of our model, we assume that this assumption covers the class of eCBC that aLIGO can detect using continuous waveform signals. Highly eccentric signals that do not fall into this category can be searched for in aLIGO data using burst search algorithms, which capture very well the imprint of eccentricity on GW signals [5].

In order to detect and characterize eCBC with aLIGO, we introduce an inspiral–merger–ringdown (IMR) waveform model that reproduces the dynamics of state-of-the-art, nonspinning, quasi-circular waveform models, and the dynamics of eccentric numerical relativity simulations.

METHODS & CODES

To construct our model, we derived higher-order post-Newtonian results for eCBC and then combined these results with state-of-the-art results from the self-force formalism and BHPT. This hybrid formalism is the first of its kind to reproduce eCBC accurately for compact binaries with mass-ratios between 1 and 10.

To construct the merger phase of our model, we assume that the system circularizes prior to the merger event. We have found that this assumption covers the class of eCBC that aLIGO can detect using continuous waveform signals. Highly eccentric signals that do not fall into this category can be searched for in aLIGO data using burst search algorithms, which capture very well the burst-like nature of highly eccentric eCBC.

RESULTS & IMPACT

We carried out the first systemic analysis of the effect of eccentricity on the first two GW transients detected by aLIGO. Fig. 1 indicates that the first two GW transients could have non-zero eccentricity at a GW frequency of 15Hz and still be misclassified as quasi-circular systems due to the restricted sensitivity of aLIGO below 25Hz. Thus, once aLIGO attains design sensitivity and is able to probe lower frequencies, we will be in a better position to accurately extract the signatures of eccentricity in GW transients. This study is the first of its kind to show that the effect of eccentricity in waveform signals cannot be mimicked by spin corrections. Indeed, the signal manifold described by GW's with eccentricity e < 0.1 at 15Hz is orthogonal to spin-aligned GW signal manifolds.

Another important result of this study is the validation of our waveform model using eccentric numerical relativity simulations. In Fig. 2 we show that our waveform model can accurately reproduce the true, accurate dynamics of these systems throughout the merger of the binary system. This is a noteworthy result because we did not use eccentric numerical simulations to calibrate our model. The exploitation of this model for upcoming GW detection campaigns with the aLIGO detectors will enable us to extend GW rule out the existence of compact binary populations that exist in dense stellar environments.

WHY BLUE WATERS

Blue Waters enabled us to create a large catalog of numerical relativity simulations, which required thousands of node hours that we ran in parallel to sample a deep region of parameter space. No other resource but Blue Waters can provide the required computational power to obtain a numerical relativity catalog of this nature in a timely manner. Furthermore, the Einstein Toolkit has been extensively used on Blue Waters since it began operations. NCSA Gravity Group members are key developers and maintainers of this software.

PUBLICATIONS AND DATA SETS

**EXECUTIVE SUMMARY**

The detection of gravitational waves has opened up a new spectrum of observation into the Universe. The LIGO flagship detection pipelines target a specific class of binary black holes that generate burst-like gravitational wave signals. In order to capitalize on the unique opportunities that gravitational wave astrophysics presents for new discoveries, it is necessary to extend the depth of gravitational wave searches to extract signals that currently go unnoticed with these pipelines. To address this issue, we introduce Deep Filtering: a new method that combines two deep convolutional neural networks for classification and regression to detect and characterize signals much weaker than background noise. We show that Deep Filtering significantly outperforms conventional machine learning techniques and enables the detection of new classes of gravitational wave signals that go unnoticed with existing detection algorithms.

**RESEARCH CHALLENGE**

Advanced LIGO (aLIGO) detection algorithms have confirmed the existence of a particular class of gravitational waves (GWs) using a 3D search: quasi-circular, spin-aligned binary black holes (BBHs). Extending these searches to target the full 6D parameter space of astrophysically motivated sources presents outstanding computational challenges [1, 2, 3]. Multimessenger searches of electromagnetic (EM) and astroparticle counterparts of GW transients rely on accurate and low-latency GW analyses, which at present take from days to months to finish. To overcome these limitations, we introduce Deep Filtering, a deep learning algorithm to directly process aLIGO data, which outperforms other machine-learning methods, and is many orders of magnitude more computationally efficient than matched filtering for both detection and parameter estimation.

**METHODS & CODES**

We consider a 2D parameter space that describes non-spinning BBHs on quasi-circular orbits, with masses between 5 and 75 solar masses, and mass-ratios of 1 to 10. We generate our data sets using BHs on quasi-circular orbits, with masses between 5 and 75 solar masses, and mass-ratios of 1 to 10. We generate our data sets using the surrogate waveform family introduced in [4]. The mass-ratio values of the BBH signals are between 1 and 10 in steps of 0.1 for training, and intermediate values for testing. We superimposed different realizations of Gaussian white noise on top of the signals over multiple iterations, thus amplifying the signal-to-noise ratio (SNR) contained about 100,000 time-series vectors produced from 4,000 templates of BBH signals by adding multiple batches of noise and shifting in time. The validation and testing sets at each SNR contained about 25,000 elements, produced from 586 clean templates and different noise realizations.

We designed simple deep neural networks (DNNs) from the ground up, since deep-learning alternatives to matched filtering had not been proposed before. We evaluated configurations of DNNs, and found that a design with three convolutional layers followed by two fully connected layers yielded the best results. We utilized the neural network functionality in the Wolfram Language, based on the MXNet framework, which utilizes the CUDA deep learning library for acceleration with NVIDIA GPUs. We used the ADAM method as our learning algorithm.

We developed a new strategy to improve the performance and reduce the training times of the DNNs. By starting with training inputs having a high SNR of less than 16, and then gradually increasing the noise in each subsequent training session until a final SNR of 0.06, we observed that the performance of prediction can be quickly maximized for low SNRs. Our algorithm can be applied to a continuous data stream using a 1-second sliding window with offsets of 0.2 seconds.

**RESULTS & IMPACT**

We trained our classifier to achieve 100% accuracy with zero false positives for signals with a SNR > 0.3. For comparison, we trained standard implementations of all commonly used machine learning classifiers along with the DNNs on a training set of 8,000 elements. Unlike DNNs, none of these algorithms was able to directly handle raw noisy data, as shown in Fig. 1. Our predictor successfully measured binary component masses of noisy GW signals that were not used for training, with an error of the same order as the spacing between templates for a SNR > 1. Although our initial goal was to create a pipeline for only non-spinning, quasi-circular BBH signals, we tested our DNNs using moderately eccentric simulations that we obtained using the Einstein Toolkit on the Blue Waters supercomputer. The classifier detected all these signals with nearly the same rate as the original test set (with over 99.7% mean accuracy for SNR > 0.36 and 100% accuracy at SNR > 0.5). The predictor was able to estimate the component masses of our eccentric simulations for SNR > 0.25 with a mean relative error less than 13%, 19%, 32%, and 34% for mass-ratio 1, 2, 3, and 4, respectively. This result is very encouraging, since these types of signals go unnoticed with all LIGO detection [5].

Our DNNs are only 2MB in size each. The average time taken for evaluating them per input of 1 second duration is approximately 6.7 milliseconds and 135 microseconds using a single CPU or GPU, respectively. For comparison, we estimated the evaluation time for time-domain matched filtering with the template bank of clean signals used for training: the results are shown in Fig. 2. Our extremely fast inference rate indicates that real-time analysis can be carried out with a single computer even with DNNs that are significantly larger, and can be trained over much bigger template banks of signals.
TOWARD ROBUST MAGNETOHYDRODYNAMIC SIMULATIONS OF GALAXY CLUSTER FORMATION

EXECUTIVE SUMMARY

Our goal is to understand the laws of physics that control the dynamics of ordinary, baryonic matter during the formation of galaxy clusters. Most of that matter is very hot, very dilute plasma. A key dynamical component that results from cluster formation is turbulence in the plasma that generates magnetic fields that control the microphysical processes essential to the physical state of the plasma. Much of the important action takes place on very small physical scales, which are quite challenging to capture in simulations, especially magnetic field behaviors. The work reported here was a vital component in developing a novel, exceedingly high-performance and exceptionally scalable MHD (magnetohydrodynamics) cosmology code named “WOMBAT” that we will apply to the problem detailed above on coming exascale systems. Here, we demonstrated that WOMBAT meets our design objectives when scaled to more than 16,000 Blue Waters nodes. Initial test simulations examined behaviors of astrophysical, hypersonic plasma jets in cluster settings.

RESEARCH CHALLENGE

Galaxy clusters are knots in the cosmic web with masses that can exceed $10^{15}$ M$_{\odot}$ and sizes of several million light years. They are the largest and last structures to form by gravitational collapse from fluctuations in the Big Bang. Their formation is a critical diagnostic of cosmological theory. Galaxies are actually minor constituents of these clusters. Most of the matter is “dark matter,” whose nature remains unclear. Most of the “ordinary” baryonic matter is very hot, very dilute plasma filling the cluster. It’s properties trace the current cluster dynamical state and its history. A full understanding of the dynamics of this intracluster medium plasma, or ICM, is absolutely necessary to comprehend how clusters form. Much of the important ICM action, and especially action responsible for generation of the magnetic fields that control ICM physical properties, takes place on scales that are multiple orders of magnitude smaller than the size of a cluster. Until now, simulations that could capture those small-scale actions in cosmological-scale computations, especially including magnetic field generation (which depends on activating the small-scale, MHD dynamo), were beyond reach with existing software and high-performance computer systems. Even with petascale systems and high-performance computer systems. Even with petascale systems such as Blue Waters, no simulation codes existed until now with sufficient performance and scaling behavior to begin to address this research challenge. Our objective is to develop tools to resolve this problem on the coming generation of exascale systems, where the properties designed into WOMBAT will truly shine.

METHODS & CODES

To address this important astrophysics problem we have built simulations that approach needed levels of physical fidelity. Our primary objective on Blue Waters for this allocation was to test and tune the technologies built into WOMBAT in order to prepare to address the cosmological problems outlined above. We have succeeded very well in that. We demonstrated, for example, up to 75% weak scaling efficiency going from one Blue Waters XE6 node to 16,224 XE6 nodes with 16 floating point threads per node. As part of our team’s collaboration with Cray, the Cray Programming Environments group developed a new MPICH release, based in part on our Blue Waters results, with enhanced performance for hybrid parallel applications. This will be the default MPICH model going forward, and will considerably benefit both the Blue Waters user community and the broader high-performance community. The N-body dark matter routines in WOMBAT were still in testing at the end of the allocation period. Those tests are very promising, but we were unable to carry out planned cosmological simulations before the allocation expired. We did, however, carry out very high-resolution WOMBAT MHD simulations of hypersonic plasma jets propagating in ICM environments in order to explore their propagation behavior. The distribution of jet plasma from one such simulation is volume rendered in Fig. 1. Instabilities forming along the boundary of the light plasma cocoon are evident.

WHY BLUE WATERS

Blue Waters was essential for this project because of the importance of using a system capable of sustained petaflop performance with very fast inter-node communications to allow WOMBAT performance testing. Blue Waters offers the unique opportunity to both test our strategy and to begin to carry out simulations that approach needed levels of physical fidelity.

PUBLICATIONS AND DATA SETS

EXPLORING THE NATURE OF EXPLODING MASSIVE STARS WITH HIGH RESOLUTION

EXECUTIVE SUMMARY

Core-collapse supernovae (CCSNe) are the violent explosions of massive stars at the end of their evolutionary lives. CCSNe play an important role in the dynamics of galaxies—injecting energy, producing and expelling heavy elements, and triggering formation of new star systems. CCSN simulation is a nexus of exotic basic physics—neutrinos, extreme gravitation, and nuclear physics. We are using 3D simulations to explore the variation in explosions and ejecta that result from the known variations in properties of massive stars (initial mass, composition, rotation, etc.) and to understand the impact of resource limitations on those outcomes. Because CCSNe are major contributors to the elemental composition of planets (including Earth), we have computed a 3D model that captures that element production in detail. Another model, limited in time and geometry, explores extremely high resolution to gauge the impact of resolution limits imposed by current supercomputers.

RESEARCH CHALLENGE

In CCSNe, the explosion is achieved by heating material in the cavity between 1–200 km radius) accretion shock and the neutron star (NS) newly formed by the collapse of the star’s iron core. The heating comes from neutrons emitted from the NS, which release the gravitational binding energy of the NS and are absorbed in the cavity. This requires energy resolved (spectral) neutrino transport coupling the NS and the heating cavity. After an explosion begins (the last 80–90% of the simulation), the evolution of the nuclear isotopes in the ejecta require a reaction network. Larger networks can directly capture the evolution of more species and more accurately account for the input of energy into the expanding ejecta. In addition to the strength and composition of the explosion, CCSNe are important targets for observation by neutrino and gravitational wave detectors. The pre-collapse progenitors come in wide variety, driven largely by variations in the initial mass, rotation, and composition. To account for this variation in the CCSN outcomes and impacts on the evolution of heavy element abundances and compare to observed CCSN properties, wide ranging sets of models must be built. We are using a modest set of 3D models that span initial stellar parameters and exploring the impact of resource-limited resolution on current and prior CCSN simulations by all groups.

METHODS & CODES

We have developed the custom multiphysics code Chimera ([2] for CCSN. It couples the hydrodynamics of a self-gravitating fluid with active nuclear reactions (burning) taking place in the ejecta, and neutrino transport to carry energy from the inner core to the heating. The hydrodynamics is solved with a dimensionally split piecewise parabolic finite volume scheme. Gravity is solved by global multiple expansion of the Poisson equation. Nuclear burning is solved by the backward Euler method for a finite sized reaction network (14 or 16 species). Neutrino transport is solved by a finite difference implementation of energy resolved flux limited diffusion that couples all neutrino species to a comprehensive set of neutrino-matter interactions.

RESULTS & IMPACT

One simulation computes a high-resolution (1/4°) model with limited geometry (90°×90° wedge along equator) during the pre-explosive convection phase, which is compared with lower-resolution models (1/2° to 2°) computed elsewhere to understand the impact of the lower resolution on other simulations [3]. We were able to measure the development of the cascade of turbulent energy to small scales and show that it is similar to previous work [4] done using, parameterized neutrino transport. We also see (Fig. 1) that the large plumes with sinking streams in the low-resolution models (1/2°) are cut-out (permitted the formation of neutron-rich calcium-48. We are also computing 1° resolution models of several other stellar progenitors to examine the variation in initial stellar mass. All of these models are being compared to less costly 2D models of the same and similar progenitors to better leverage the 3D results to broader applicability. For all our models, we will compute post-processed isotopic production with very large (6,000 species) networks and compute neutrino and gravitational wave signals for comparison to future observations.

WHY BLUE WATERS

CCSN simulations are large, lengthy, and expensive, requiring 1000s+ coupled nodes. Even a single 3D simulation can overwhelm the available allocations for a single project at other large sites, but with Blue Waters we can perform about three simulations per year.
EXECUTIVE SUMMARY

Our main objective is to characterize the backflow contamination environment due to plasma created by electric-propulsion plumes, their interaction with the spacecraft environment, and neutralizer sources, using state-of-the-art high-performance petascale computations. In terms of modeling and simulation, we build on our earlier work where we have developed an object-oriented C++ Direct Simulation Monte Carlo (DSMC) code that uses adaptive mesh refinement (AMR) octree grids to capture the vast length scales inherent in supersonic expansions to vacuum for neutral–neutral and neutral–ion collisions. A key aspect of our computational work is to take advantage of our unique, recent advances in GPU (graphics processing unit) multi-thread parallelization applied to tree-based computational strategies. Blue Waters is especially suited to this modeling since we expect to use 128 to 256 GPUs per run for the plasma plume simulations on the 5K nodes. Initial results demonstrate that our approach has been successfully extended to couple DSMC and particle-in-cell (PIC) simulations on an AMR/octree grid using GPUs.

RESEARCH CHALLENGE

With the space environment becoming a home to constellations of small satellites and cube-satellites, improved predictability of key surfaces of solar cell arrays and spacecraft charging in the backflow environment of chemical and electric-propulsion (EP) thrusters is crucial. Fig. 1 shows a schematic of the multiple sources and processes that must be considered. Because there are both neutral and charged species in the backflow region of an onboard EP thruster, the modeling of these highly reactive ions with thruster and ambient neutral species involves multiple time and length scales. Indirect environmental exposure of spacecraft material such as the micron-sized coatings of solar array cover glass and aluminized Mylar can cause appreciable sputtering and erosion. This is often hard to quantify and predict over the lifetime of the mission because backflow ion fluxes are about five orders of magnitude less than those due to main ion beam impingement. The well-known charge-exchange (CEX) process between xenon thruster neutrals and beam ions also occurs among plume/beam ions and ambient species such as Xe+ and O, although these processes are not presently modeled in spacecraft environment effect models. In addition, external hollow cathodes are a source of electrons and also emit xenon ions and neutrals. Inclusion of their plumes is essential as the electrons play an important role in the charge density distribution, which influences the generation of the slow CEX ions, and, in turn, affects the erosion of solar cell panels.

METHODS & CODES

Our new plasma modeling is an outgrowth of our DSMC code, CHAOS (CUDA-based Hybrid Approach for Octree Simulations) that was developed to study neutral flows through porous media [1]. We have adapted this approach now to include both neutral and ion species in an external electric field. Because the local mean free path for collisions is about three orders of magnitude larger than the local Debye length (~10^-6 m), two linearized Morton-ordered forests of octrees (FOTs) have been implemented so that these grids can be adapted to meet these two diverse numerical criteria. To accurately model the electric field, using a Boltzmann distribution at a fixed electron temperature, or to solve Poisson’s equation, a leaf node should be only one level larger than its smallest face neighbor, i.e., a “2:1 criteria.” When this criterion is adhered to, any numerical discontinuities in the E-octrees are smoothed out and we are able to maintain first-order accuracy in the gradient calculations. In our AMR/octree approach, the DSMC cells of variable size satisfy the mean-free-path and Debye length criteria but do not automatically satisfy the 2:1 ratio. Implementation of this on an AMR/octree grid is nontrivial and has been accomplished through the use of local (on a single processor) and local–global (across processors) stages.

RESULTS & IMPACT

We have recently run simulations of mesothermal, collisionless plasmas for shifted electron and ion sources [2]. We discovered very interesting unsteady plume dynamics by modeling a shifted electron source (in contrast to co-located electron and ion sources) as summarized in Fig. 2. The electron source, with a radius of R_e = 0.01 m, was placed above the ion source, with an electron thermal to ion drift velocity of 1.25. The Xe+ electron number densities were 5 x 10^9 and 1.0 x 10^10 (cm^-3), respectively. Temperatures and bulk velocities were assumed to be 0.04 and 2 eV and 472,000 and 0 m/s, for ions and electrons, respectively. For this simulation, both ions and neutrals had the same time step of 2.8 x 10^-10 s. Fig. 2 shows a summary of the interesting time-varying behavior, which is very different for the ions versus the electrons. The electron dynamics were found to be complex, and as soon as electrons are released, they accelerate toward the beam and travel beyond the beam front. The positively charged beam front then deaccelerates the electrons that have escaped and reflects them back in the reverse direction, resulting in negative electron velocities. The continual exchange between electron kinetic and potential energy results in a meandering/bouncing movement, as can be seen in the top row of the figure, as time (normalized by plasma frequency) increases. The ion beam, however, more or less propagates away from its emitting source as time increases. Nevertheless, some diffusion of ions can be seen, particularly in the beam front at later times. In comparison to the electrons, the large ion mass keeps them moving primarily away from the source. In addition, as time progresses, the electron bouncing region is elongated with the multiple bouncing motions increasing the electron temperature. The increase in temperature is contrary to what is expected in an expanding plume and, not surprisingly, the actual electron velocity distribution obtained in the beam front region is very non-Maxwellian.

WHY BLUE WATERS

Blue Waters has allowed us to test on a large number of GPUs the potential computational savings to enable three-dimensional, fully kinetic plasma simulations. Compared to the present state-of-the-art plasma simulations, a uniform grid in 3D would require about a factor of seven more cells than our use of AMR/octrees. The use of a single K20x GPU decreases the runtime by a factor of five compared to a single Interlagos processor. Very conservatively, we estimate that the octrees in combination with GPUs, decrease the total runtime by at least a factor of 10, compared to uniform grid solvers on multi-core CPUs.
THREE-DIMENSIONAL NATURE OF COLLISIONLESS MAGNETIC RECONNECTION AT EARTH’S MAGNETOPAUSE

EXECUTIVE SUMMARY

Earth’s magnetosphere shields the planet from constant bombardment by supersonic solar winds. However, this magnetic shield, called the magnetopause, can be eroded by various plasma mechanisms. Among them, magnetic reconnection is arguably the most important process. Reconnection not only allows the transport of solar wind plasmas into Earth’s magnetosphere but also releases the magnetic energy and changes the magnetic topology. At Earth’s magnetopause, magnetic reconnection proceeds between the shocked solar wind plasmas and the magnetosphere plasmas. The magnetosheath plasma has a typical magnetic field strength of ~20 nanoteslas and density of ~5 per cc. The magnetosphere plasma has a magnetic field strength of ~60 nanoteslas and density of ~0.5 per cc. The magnetic fields on these two sides can shear at any angle. Many three-dimensional properties of magnetic reconnection in such asymmetric geometry remain unclear. We use first-principle simulations to explore the nature of magnetic reconnection in such asymmetric geometry.

RESEARCH CHALLENGE

Massive solar eruptions drive magnetic storms that impact Earth’s magnetosphere and produce space weather. The consequential electromagnetic waves, electric currents, and energetic particles can do harm to satellites, astronauts, GPS systems, radio communication, and power grids on the ground. Magnetic reconnection is the key player in such solar wind-magnetospheric coupling, and space weather in general. One fundamental question in reconnection study is: Is there a simple principle that determines the orientation of the reconnection x-line in such an asymmetric current sheet? The solution of this problem remains unclear with our current understanding of magnetic reconnection, and we aim to resolve this issue. Ultimately, we hope to develop an adequate understanding of the 3D nature of asymmetric magnetic reconnection itself, which is a crucial step in the quest for predicting the location and rate of flux transfer at Earth’s magnetopause. A better understanding of the nature of magnetic reconnection will advance the modeling of space weather.

METHODS & CODES

This project employs the particle-in-cell code VPIC [1]. VPIC solves the relativistic Vlasov–Maxwell system of equations using an explicit charge-conserving approach. Charged particles are advanced using Leapfrog with 6th-order Boris rotation, then the current and charge density are accumulated on grid points to update electromagnetic fields. Mander divergence cleaning is frequently employed to ensure the divergent free of the magnetic field. The level of error is bounded by the numerical round-off effect. These 3D kinetic simulations are now generating large amounts of data (~O(100) TB) for each run. We have met this challenge with help from visualization experts at LANL and LBNL to develop parallel readers that interface between these large VPIC data sets and the open-source package ParaView.

RESULTS & IMPACT

During the first award year, we have conducted several petascale simulations on Blue Waters to study the three-dimensional nature of the reconnection x-line. We identify features in these simulations that can potentially bring the most scientific merit. Our approach of inducing a solitary x-line is working well in a large-scale 3D simulation. We initialize the x-line at the center of the box and then let the x-line expand freely and form a well-defined orientation. Here, the orientation is measured as ~13°, close to the angle that bisects the total magnetic shear angle made by magnetic fields on two sides of the current sheet. In comparison with other existing work, where the interaction between multiple x-lines often complicates the orientation measurement, this approach provides an ideal setup for the x-line to grow in a certain “preferential” orientation. Also, we measured the expanding/spreading velocity of the reconnection x-line, decided whether there is an intrinsic reconnection x-line extent in a fully three-dimensional system, and studied the turbulent nature of reconnection.

The same principle that determines the x-line orientation could interplay with global geometrical effects to determine the location and orientation of magnetic reconnection at Earth’s magnetopause. Knowledge of the extent of x-line will enable space scientists to more accurately estimate the efficiency of flux transfer from solar wind to the Earth’s magnetosphere. The work proposed here is relevant to the study of dayside reconnection during the first phase of NASA’s Magnetospheric Multiscale Mission (MMS). The close tetrahedron deployment of the MMS spacecraft cluster and the equipped capability of high time-resolution provide an invaluable chance to reconstruct the three-dimensional geometry of the reconnection diffusion region.

WHY BLUE WATERS

Because the x-line has a dimension down to electron scale, a fully kinetic description is necessary. Given the available computational capability, it has become possible to use a first-principle kinetic simulation to investigate the dynamics of the x-line in a reasonably large 3D system, which spans from electron kinetic scale to magnetohydrodynamics scale. A representative 3D run in this project traces the motion of 2 trillion charged particles under the interaction of self-generated electromagnetic fields, which are evaluated on 6 billion grids. The output data easily have a size of hundreds of TBs for each run. Blue Waters not only provides the computational resource for the calculation but also the online storage for the output and restart files.
TRANSFORMATIVE PETASCALE PARTICLE-IN-CELL SIMULATIONS

EXECUTIVE SUMMARY
The NSF Leadership-class system at NCSA, Blue Waters, was used to study highly linear and kinetic processes that occurred in high-energy density plasmas, including plasma-based acceleration and laser fusion. These simulations have resulted in many high-impact publications in Physical Review Letters. In summary, Blue Waters resources allow the UCLA simulation group to perform high-impact research and train a new generation of plasma physicists capable of performing simulations on current and upcoming world-class supercomputers.

RESEARCH CHALLENGE
The research focused on three key areas with clearly developed science questions identified by the community [1–3]. This research, funded by both the National Science Foundation and the Department of Energy, addressed the following key questions:
• Can plasma-based acceleration be the basis of new compact accelerators for use at the energy frontier, in medicine, in probing materials, and in novel light sources?
• Can laser-plasma instabilities be controlled or even harnessed in inertial confinement plasmas?
• What are the collective processes responsible for the formation of speckles in collinear plasma lasers? Are collisionless shocks in plasmas responsible for the most energetic particles in the universe?

METHODS & CODES
Based on the highly nonlinear and kinetic processes occurring in high-intensity laser and beam-plasma interactions and in plasma based acceleration, we use PIC (particle-in-cell) codes [4,5], where Maxwell’s equations are solved on a grid using currents and charge densities calculated by weighting particles onto the grid. For this project, we employ the PIC codes OSIRIS, QuickPIC, and UPIC. These codes are all developed locally by the UCLA simulation group (and in collaboration with Instituto Superior Técnico), share many of the same algorithms and data structures, and have been optimized for heterogeneous leadership class supercomputers such as Blue Waters. These codes are freely available and QuickPIC and UPIC are open source (on GitHub).

RESULTS & IMPACT
Large-scale particle-in-cell simulations have been performed that will impact the design of future experiments in plasma-based accelerators and inertial confinement fusion. Blue Waters was used to study the generation of high-quality electron beams (with high brightness and low energy spread) using the density down ramp injection scheme. By carefully choosing the parameters of the plasma and the driver, Blue Waters simulations show that electrons with 1GeV (gigaelectronvolt) energy, 0.2% relative projected energy spread, and >10kA (kilampere) current can be generated in nonlinear plasma waves. This is suitable for X-ray FEL (free-electron laser) applications, which can provide a compact radiation source for nuclear science. Simulations were also run to study (6) the effects of ion motions on very-high-brightness beams such as those required for future collider design. In future colliders, which confine electrons within a radius of a few nanometers, the space charge forces around the beam can pull the plasma ions inward. This generates nonlinear force inside the wake, which can potentially lead to beam-emittance growth. The self-consistent simulation for this problem remains a big challenge because the simulation box has a transverse size of hundreds of microns. However, the cell size needs to be a few nanometers in order to resolve the electron beam, which results in a simulation box with 1010 cells. Blue Waters (because of its large memory size) is ideally suited to study this particularly demanding problem, and simulations performed on it showed that ion motion does not necessarily lead to catastrophic emittance growth.

In collaboration with colleagues at the Naval Research Laboratory, Blue Waters has been used to study the effects of temporal bandwidth (laser smoothing) on laser plasma interactions in laser fusion. In current inertial confinement fusion experiments, lasers can lose a large fraction of their energy to laser-plasma interactions where the laser decays into a backscattered light and a plasma wave. Large-scale 2D OSIRIS simulations showed that, given enough bandwidth (in the order of several terahertz), laser plasma instabilities can be suppressed. Blue Waters resources allow our group to simulate, for the first time, the interaction of many speckles in 2D over several picoseconds with beam smoothing. The additional realism provided by these simulations will lead to a better understanding of laser-plasma interactions in current and future experiments in inertial confinement fusion.

WHY BLUE WATERS
Blue Waters provides the largest, time-tested, and stable supercomputing platform in the world. It has a CPU side that is time tested and very stable, and has allowed us to perform large-scale simulations in a timely manner, and a GPU side that provides a testbed for code development. This combination suits almost all research supercomputing needs, and has provided a productive computational environment and a satisfactory experience since the very beginning of Blue Waters.
MINI-DISK DYNAMICS ABOUT SUPERMASSIVE BLACK HOLES BINARYs

Allocation: NSF PRAC/1,500 Knh
PI: Scott C. Noble
Co-Pi: Manuela Campanelli1, Julian H. Krolik1
Collaborators: Dennis Bowen2, Stephane d’Ascoli2, Vassilios Mewes2, Hotaka Shiokawa4

1The University of Tulsa
2Rochester Institute of Technology
3Johns Hopkins University
4Harvard-Smithsonian Center for Astrophysics

EXECUTIVE SUMMARY

The overall goal of our project is to provide the community with realistic predictions of electromagnetic (EM) signatures from supermassive black hole binaries (SMBHBs). These events are of much theoretical interest now since little is known about them, but they are ripe for discovery with current and future high-cadence large-sky observational campaigns (e.g., PAN-STARRS, Catalina Real-time Transient Survey, LSST), as well as future multi-messenger endeavors with space-based gravitational wave detectors. We passed the way toward this goal by performing a set of hydrodynamic and magnetohydrodynamic (MHD) simulations of these systems using the HARM3d code on the Blue Waters system. The simulations are frontier calculations and require incredible resources and significant effort to cover the large range in time and length scales inherent to their physical situation.

In previous years, we have performed 3D general relativistic MHD (GRMHD) simulations of the circular disk that resides outside the orbit of the binary [1, 2]. These long timescale runs were affordable because we excised the binary from the evolution, which allowed us to take larger time steps yet maintain numerical stability. Unfortunately, the excision technique prevented us from learning about the gas dynamics that occurs near each black hole (BH) and the processes that develop between the “mini-disks” of gas that form there. These mini-disk simulations were the focus of our work on Blue Waters this past year, and provided us with first-of-a-kind comprehensive GRMHD simulations of accreting SMBHBs. Including the mini-disks is critically important to understanding the EM signatures because the mini-disks make up a large fraction of the total luminosity from the system, and they give rise to the most variable emission, which is a key means by which astronomers hope to characterize and identify SMBHBs.

In order to start from initial conditions closer to a steady-state and to save resources, Rochester Institute of Technology graduate student Dennis Bowen wrote a method to generate mini-disks around each BH. The initial data method was used first to explore how general relativistic terms affect the dynamics of 2D nonmagnetized disks [3].

METHODS & CODES

All the results reported here used the flux-conservative, high-resolution, shock-capturing GRMHD code called HARM3d [4]. Besides using modern computational fluid dynamics techniques and a solenoidal constraint transport scheme for the magnetic field’s induction equation, GRMHD is written in a general covariant way so that any metric or coordinate system may be adopted. This last property has accommodated our implementation of a novel, time-dependent, nonuniform gridding scheme for resolving the small spatial scale features near each BH, and the larger spatial scale flow evolving circularly that extends much further out beyond the SMBHBs’ orbit [5]. This coordinate system enabled us to perform the first 3D GRMHD simulations of SMBHB mini-disks already mentioned.

RESEARCH CHALLENGE

In the 2D hydrodynamic simulations, we found that significant mass exchange, or “sloshing,” of material occurred between the two mini-disks. The fluctuations of the mass in the sloshing region resulted in a high-quality and characteristic signal, which we would expect to produce a similarly characteristic periodic EM signal because of the high rate of dissipation there. We further found that the amount of material in the sloshing region increases as the binary shrinks, suggesting that the sloshing emission will become brighter as the BHs grow closer and proceed to the relativistic limit. Our findings suggest that binaries near merger will be bright and periodic at a time scale associated with the SMBHB’s orbital period—a key to extracting information about an observed binary’s orbit.

Continuing our use of the new mini-disk initial data, we have performed the first full 3D MHD evolution of mini-disks about black hole binaries in the relativistic regime, including the accretion from the circumbinary disk. This full run will begin to address the importance of spiral shocks in the overall angular momentum budget relative to internal MHD stresses. Additionally, we will be able to explore the effects of accretion streams from the circumbinary on the overall structure of the mini-disks. Our results from this simulation will be submitted soon.

WHY BLUE WATERS

The 3D GRMHD mini-disk simulation ran for two orbital periods and used 12.9 million floating-point-core-hours, or 0.806 M node-hours, on Blue Waters. The simulation used 600x160x640, or approximately 60 million cells, on about 2 million time steps using 600 nodes or 19,200 Blue Waters cores. The simulation is challenging because of the large dynamic range of time scales between the fast behavior near the black holes and the relatively slow orbital velocity of the binary, resulting in a month-long run time. Members of the NCSA Blue Waters team, e.g., David Xing and Jing Li, were very helpful in arranging a reservation for our run during a busy time on the cluster. The reservation allowed us to finish the simulation before the end of the allocation period.

PUBLICATIONS AND DATA SETS


REALISTIC SIMULATIONS OF THE INTERGALACTIC MEDIUM: THE SEARCH FOR MISSING PHYSICS

EXECUTIVE SUMMARY
We have performed state-of-the-art simulations of the intergalactic medium (IGM) during the epoch of helium reionization. UV radiation from quasars ionizes the intergalactic helium over a period of several billion years beginning about 1 billion years after the Big Bang. We have carried out a suite of the first fully coupled radiation hydrodynamic cosmological simulations that treat the quasars as a time-varying population of point sources. We have performed multigroup radiative transfer self-consistently coupled to the cosmological hydrodynamics of the IGM at sufficient resolution and domain size to examine the photoionization and photoheating processes in detail. We have discovered that helium reionization completes significantly later compared to models that treat the quasar radiation as a homogeneous background. This modiﬁes the heating history of the IGM substantially, with a maximum mean temperature of 14,000 K being achieved at a redshift of 3, consistent with observations.

RESEARCH CHALLENGE
In the past decade, more precise observations of the intergalactic medium (IGM)—the hydrogen and helium gas between the galaxies produced in the Big Bang—have revealed a discrepancy with the well-established predictions of our computational models. In particular, precision observations of the IGM using the Keck telescopes in Hawaii show that the temperature and ionization state of the IGM is not what our standard cosmological simulations predict: The IGM is either somewhat hotter than ultraviolet radiation from stars in galaxies can make it, or the IGM is distributed differently in space than the models predict, or both. There could be missing sources of heat in our models, such as energy injection by decaying dark matter particles. The discrepancy is perplexing since the standard model predicts the galaxy distribution exceedingly well. The discrepancy suggests that the standard model lacks some essential ingredient that we refer to simply as “missing physics.” The signiﬁcance of this project to the nation is that it promotes the progress of science in the fundamental ﬁeld of cosmology where the U.S. is a world leader. The project is addressing the issue of whether we are overlooking a key component of the mass-energy content of the universe. Precise answers require powerful tools, and the Blue Waters supercomputer is the tool for the job.

METHODS & CODES
We have used an enhanced version of the hydrodynamic cosmology code ENZO to examine the possibility that inhomogeneous photoheating of the IGM by quasars is the missing physics. In the standard model of the Lyman alpha forest, quasar ionization is modeled as a homogeneous but time-varying radiation background. Photoheating is treated in the optically thin limit, which underestimates the heating behind optically thick ionization fronts. In reality, quasars are radiating point sources that ionize the helium in the IGM in their vicinity; it is the growth and eventual

Figure 1: Redshift evolution of the thermal and ionization state of the intergalactic medium on a slice through the 80 Mpc simulation volume. Left to right: redshift snapshots at z=4, 3, 3.5, and 2. Top to bottom: H II ionization, He II ionization rate, He II photoheating rate.

Figure 2: The time evolution of the IGM temperature for several simulations, overlaid on observational data that show that IGM temperature peaks around redshift 3. The standard model (blue line) peaks at z=3.5—earlier than observed. However, the MGFLD simulations, shown in other colored lines, peak around redshift 3, which is in agreement with observations. The reason for the difference has to do with the finite time it takes for an ionization front to propagate across the vast distances of intergalactic space. This discovery helps resolve one, but not all, of the above-mentioned discrepancies among observations and our earlier models. The impact of these results is that all future models of the IGM must be revised to include this finite time propagation effect.

WHY BLUE WATERS
Blue Waters is required because the simulation is extremely computationally intensive. Large spatial volumes need to be simulated at high resolution, to simultaneously resolve the Lyman alpha forest absorbers and at the same time encompass a representative sample of quasars. The ratio of the outer and inner scales is about 2,000, meaning that simulations require 2,000 grids. Additionally, a multifrequency treatment of radiative transfer is essential to accurately model the transport of the hard UV spectrum of quasars including the phenomena of ionization front pre-ionization and spectral hardening of the radiation field. The ability to do a suite of exploratory runs of this new type of simulation with excellent throughput, each of which is quite computationally intensive, is only possible on a Blue Waters-scale system. In this way we are able to home in on the model that best agrees with the high-precision observations.

PUBLICATIONS AND DATA SETS

Allocation: NSF PRAC/2,900 Kish
PI: Michael L. Norman1,2,3
Co-PI: James Bond34
Collaborators: David Tytler5, Pengfei Chen1, Dan Reynolds5

1University of California, San Diego
2Center for Astrophysics and Space Sciences
3San Diego Supercomputer Center
4Co-PI:
5PI:
6Michael L. Norman1,2,3
7Crafting Collaborators:
81University of California, San Diego
92Center for Astrophysics and Space Sciences
10Collaborators:
11Michael L. Norman1,2,3
12Allocation:
13THE SEARCH FOR MISSING PHYSICS
14REALISTIC SIMULATIONS OF THE INTERGALACTIC MEDIUM: THE SEARCH FOR MISSING PHYSICS
15EXECUTIVE SUMMARY
16WHY BLUE WATERS
17PUBLICATIONS AND DATA SETS
18

Figure 1: Redshift evolution of the thermal and ionization state of the intergalactic medium on a slice through the 80 Mpc simulation volume. Left to right: redshift snapshots at z=4, 3, 3.5, and 2. Top to bottom: H II ionization, He II ionization rate, He II photoheating rate.

Figure 2: The time evolution of the IGM temperature for several simulations, overlaid on observational data that show that IGM temperature peaks around redshift 3. The standard model (blue line) peaks at z=3.5—earlier than observed. However, the MGFLD simulations, shown in other colored lines, peak around redshift 3, which is in agreement with observations. The reason for the difference has to do with the finite time it takes for an ionization front to propagate across the vast distances of intergalactic space. This discovery helps resolve one, but not all, of the above-mentioned discrepancies among observations and our earlier models. The impact of these results is that all future models of the IGM must be revised to include this finite time propagation effect.

WHY BLUE WATERS
Blue Waters is required because the simulation is extremely computationally intensive. Large spatial volumes need to be simulated at high resolution, to simultaneously resolve the Lyman alpha forest absorbers and at the same time encompass a representative sample of quasars. The ratio of the outer and inner scales is about 2,000, meaning that simulations require 2,000 grids. Additionally, a multifrequency treatment of radiative transfer is essential to accurately model the transport of the hard UV spectrum of quasars including the phenomena of ionization front pre-ionization and spectral hardening of the radiation field. The ability to do a suite of exploratory runs of this new type of simulation with excellent throughput, each of which is quite computationally intensive, is only possible on a Blue Waters-scale system. In this way we are able to home in on the model that best agrees with the high-precision observations.

PUBLICATIONS AND DATA SETS

Allocation: NSF PRAC/2,900 Kish
PI: Michael L. Norman1,2,3
Co-PI: James Bond34
Collaborators: David Tytler5, Pengfei Chen1, Dan Reynolds5

1University of California, San Diego
2Center for Astrophysics and Space Sciences
3San Diego Supercomputer Center
4Co-PI:
5PI:
6Michael L. Norman1,2,3
7Crafting Collaborators:
81University of California, San Diego
92Center for Astrophysics and Space Sciences
10Collaborators:
11Michael L. Norman1,2,3
12Allocation:
13THE SEARCH FOR MISSING PHYSICS
14REALISTIC SIMULATIONS OF THE INTERGALACTIC MEDIUM: THE SEARCH FOR MISSING PHYSICS
15EXECUTIVE SUMMARY
16WHY BLUE WATERS
17PUBLICATIONS AND DATA SETS
18

Figure 1: Redshift evolution of the thermal and ionization state of the intergalactic medium on a slice through the 80 Mpc simulation volume. Left to right: redshift snapshots at z=4, 3, 3.5, and 2. Top to bottom: H II ionization, He II ionization rate, He II photoheating rate.

Figure 2: The time evolution of the IGM temperature for several simulations, overlaid on observational data that show that IGM temperature peaks around redshift 3. The standard model (blue line) peaks at z=3.5—earlier than observed. However, the MGFLD simulations, shown in other colored lines, peak around redshift 3, which is in agreement with observations. The reason for the difference has to do with the finite time it takes for an ionization front to propagate across the vast distances of intergalactic space. This discovery helps resolve one, but not all, of the above-mentioned discrepancies among observations and our earlier models. The impact of these results is that all future models of the IGM must be revised to include this finite time propagation effect.
SIMULATING GALAXY FORMATION ACROSS COSMOMIC TIME

EXECUTIVE SUMMARY

This simulation campaign addresses several pressing questions about galaxy formation and evolution by using a suite of physics-rich, high dynamic range adaptive mesh refinement simulations of cosmological structure formation. The two main thrusts of these simulations are: (1) to understand the connection of the first generation of galaxy formation with the Milky Way and its satellites, and (2) to understand the cycling of metal-enriched, magnetized plasma into and out of galaxies like the Milky Way, and the regulation of star formation in these galaxies. All of these problems require simulations with extremely high dynamic range in space and time, complex physics (including radiation transport and nonequilibrium gas chemistry), and large simulation volumes. We use the Enzo code (enzo-project.org), which has been modified to scale to large core counts on Blue Waters—the only machine available where the heavy data and communication needs can be satisfied.

RESEARCH CHALLENGE

Our goals are to understand two critical issues in galaxy formation: the formation of the earliest generations of galaxies and their connections to the Milky Way through hierarchical structure formation, and the “baryon cycle” in galaxies like the Milky Way—in other words, how gas gets into and out of galaxies, and what it does while it is there. Both of these questions are important to understanding observations of galaxies over the age of the universe using telescopes such as the 10-meter Keck telescope on Mauna Kea and the Hubble Space Telescope, which are both used to observe light from very distant galaxies as well as the absorption of light by the intergalactic and circumgalactic medium. The simulations used to properly model galaxies in both the early universe and the present day require extremely high spatial and temporal dynamic range, and also require complex physics—most importantly, radiation transport, magnetohydrodynamics, and nonequilibrium gas chemistry. Furthermore, large simulation volumes (and thus many resolution elements) are needed to model the many early galaxies that will merge together to create a Milky Way-like galaxy at the present day. Additionally, in our present-day galaxy simulations, huge numbers of cells are required to accurately resolve the circumgalactic gas. Taken together, this requires the use of a supercomputer with large memory and disk space (to accommodate the tremendous dataset sizes), large computational resources, and an extremely high bandwidth, low-latency communication network to enable significant scaling of the radiation transport code. Blue Waters is the only machine available to the academic community that fits all of these requirements.

WHY BLUE WATERS

The simulations used to properly model galaxies in both the early universe and the present day require extremely high spatial and temporal dynamic range, and also require complex physics—most importantly, radiation transport, magnetohydrodynamics, and nonequilibrium gas chemistry. Furthermore, large simulation volumes (and thus many resolution elements) are needed to model the many early galaxies that will merge together to create a Milky Way-like galaxy at the present day. Additionally, in our present-day galaxy simulations, huge numbers of cells are required to accurately resolve the circumgalactic gas. Taken together, this requires the use of a supercomputer with large memory and disk space (to accommodate the tremendous dataset sizes), large

collaboration: Devin W. Silvia, Cameron J. Hummels, Britton D. Smith

4 California Institute of Technology
5 University of California at San Diego

PI: Brian W. O’Shea
Co-PIs: David C. Collins, John H. Wise
Collaborators: Devin W. Silvia, Cameron J. Hummels, Britton D. Smith

1 Michigan State University
2 Florida State University
3 Georgia Institute of Technology

METHODS & CODES

Our simulation tool of choice is the Enzo code [1; also see http://enzo-project.org ], an open-source and community-developed software platform for studying cosmological structure formation. Enzo allows us to include all the critical physical components needed to study galaxy formation—gravity, dark matter dynamics, fluid dynamics, the microphysics of plasmas, and prescriptions for star formation and feedback—and to do so using a tool that can scale to large numbers of CPUs. All analysis was done with the yt code [2, also see http://yt-project.org ].

RESULTS & IMPACT

The analysis of the simulation performed as part of this campaign has only recently begun. However, our most important results thus far involve the cycling of gas into and out of galaxies. Both cosmological simulations of galaxy formation as well as more idealized calculations demonstrate that massively increased physical resolution in the “circumgalactic medium”—the gas outside the stellar disk of a galaxy but which is bound to the galaxy by gravity, and composes almost half of the mass of the baryons in the galaxy—is incredibly important. In particular, we find that increasing the resolution by more than an order of magnitude beyond previous state-of-the-art calculations results in the appearance of both spatial and chemical features that are seen in observations but not in previous models. Similarly, we find in our idealized simulations that galaxies can attain a dynamic equilibrium between cold gas condensing and falling into the galaxy (and thus fueling star formation), and the ejection of hot, metal-enriched gas into the circumgalactic medium. This work is changing our understanding of the interface between the stellar component of galaxies and the diffuse plasma component.

PUBLICATIONS AND DATA SETS

No publications, reports, or datasets have resulted from this simulation campaign yet. Several improvements to the open-source Enzo code have been made, and can be found at http://enzo-project.org.
3D GENERAL-RELATIVISTIC RADIATION-HYDRODYNAMIC SIMULATIONS OF CORE-Collapse SUPERNOVAE

EXECUTIVE SUMMARY

Core-collapse supernovae (CCSNe) are the magnificent explosions of massive stars. They are the birth sites of black holes and neutron stars, and they enrich the interstellar medium with the chemical elements produced by thermonuclear fusion. From these elements, planets form and life develops.

Using Blue Waters, we carried out the very first ab initio full 3D general-relativistic radiation-hydrodynamics CCSN simulations, focusing on the phase between initial collapse and the onset of explosion. We investigated the CCSN evolution of a 27-solar-mass progenitor star and followed the supernova engine for approximately 400 milliseconds in full 3D. We found the onset of an explosion driven by a combination of neutrino energy deposition and turbulent convection. The explosion develops in a large-scale asymmetric way, which is consistent with recent astronomical observations of supernova remnants.

METHODS & CODES

We employ our 3D general-relativistic radiation-hydrodynamics code, Zelmani. This is an open-source code based on the Cactus framework that uses components of the open-source Einstein Toolkit. Zelmani is the only U.S.-based code that is fully general-relativistic and based on numerical relativity. The key new development enabling our simulations is the Zelmani-M1 component, which implements full 3D multiphysics, multiequation radiation-hydrodynamics in the two-moment “M1” approximation. In M1, the Boltzmann equation is expanded in angular moments. The equations for radiation energy (a scalar) and flux (a 3-vector) are evolved and the system is closed with a local algebraic closure relation. Extensive tests show that this approach yields accurate results for CCSN neutrino radiation fields.

We employ a single 27-solar-mass progenitor star and carry out the initial collapse phase in 1D using our open-source code GR1D. We map to 3D at 20 milliseconds after core bounce (the time of proto-neutron star formation). In order to study the effects of numerical resolution on shock revival, we carry out simulations with spatial resolution of 1.5 km and 3 km in the region behind the stalled shock. We study the impact of imposed symmetries by simulating in full unconstrained 3D and in an octant of the 3D cube.

RESULTS & IMPACT

In this first set of full 3D CCSN simulations, we confirm that the well-studied “neutrino mechanism,” in combination with turbulent convection, is indeed able to drive 3D CCSNe. The neutrino mechanism relies on the deposition below the shock of a small fraction (5 to 10%) of the neutrinos emitted from the edge of the proto-neutron star. The neutrino energy is deposited in the gas, driving up temperature, thermal pressure, and establishing an entropy gradient that is unstable to buoyant convection. This convection is highly turbulent (theoretical Reynolds number of order $10^5$) since viscosity is extremely small. This turbulent neutrino-driven convection exerts effective turbulent pressure that, when combined with the thermal pressure, helps push out the shock and explode the star. Figure 1 shows a volume rendering of the specific entropy distribution within the CCSN shock as the explosion is developing. There are large high-entropy bubbles that develop over time (“volume filling convection”) and push out the shock. The sphere-like object in the center is the low-entropy region inside the proto-neutron star. The shock is strongly deformed and shows many small- to large-scale protrusions. These are created by convective bubbles that impinge on the shock from below. The explosion is globally asymmetric and elongated along one axis. Similar geometries are found in deep astronomical observations of CCSN remnants.

Furthermore, we find that low spatial resolution artificially favors explosion, because it traps turbulent kinetic energy at large scales where it can effectively help shock expansion. This result emphasizes the need for high-resolution simulations in addition to including the full set of physics ingredients. Comparing our octant 3D simulations with full 3D, we find that in the latter, explosions develop more easily. This is because the expanding large high-entropy bubbles that form in full 3D tend to have complex geometry that cannot be captured by a simulation that is constrained to an octant.

WHY BLUE WATERS

Blue Waters has been absolutely essential to this project. While our simulations typically use “only” 400 nodes, and other HPC systems could accommodate a single simulation, Blue Waters is the only U.S. resource that allows us to carry out multiples of such simulations with high throughput. Without Blue Waters, this project would have taken many years to complete.

REFERENCES


**EXECUTIVE SUMMARY**

Blue Waters time allocation through the NSF PRAC award was used to investigate physical phenomena occurring when the solar wind (SW) interacts with the local interstellar medium (LISM): (1) the origin of the SW on the solar surface and its further acceleration to supersonic velocities; (2) the effect of transient phenomena on space weather on Earth; (3) the SW propagation throughout the heliosphere and perturbations it creates in the LISM; (4) the effect of the heliopause boundary layer on the plasma oscillation events observed by Voyager 1 in the LISM; (5) MHD (magnetohydrodynamics) instabilities and magnetic reconnection; (6) the influence of the heliopause on the observed anisotropy of TeV galactic cosmic rays; and (7) using observations from multiple spacecraft to reconstruct otherwise missing properties of the SW and LISM. Our simulations are important for the explanation of IBEX (Interstellar Boundary Explorer), New Horizons, Voyagers, and Voyager measurements, as well as air shower observations.

**RESEARCH CHALLENGE**

The grand challenge of this research is to investigate physical phenomena that start on the solar surface and result in the SW acceleration and propagation through interstellar space toward the boundary of the heliosphere, where the SW interacts with the LISM. Our simulations are data-driven and help interpret observations from such space missions as IBEX, New Horizons, Voyagers, and Voyager measurements, as well as air shower observations.

**METHODS & CODES**

We solve the ideal magnetohydrodynamics (MHD) equations coupled with the kinetic Boltzmann equation describing the transport of neutral atoms. In a less strict approach, the flow of atoms is modeled with a few systems of the Euler gas dynamic equations describing different atom populations dependent on the domains of their origin. We have developed both fluid dynamics and kinetic models for PU and turbulence generated by kinetic instabilities of their distribution function. All these are components of a Multi-Scale Fluid-Kinetic Simulation Suite (MS-FLUKSS)—an adaptive mesh refinement code we have built on the Chombo framework from LANL.

**RESULTS & IMPACT**

As a result of the work supported by our Blue Waters allocation through the NSF PRAC award OCI-1615206 we have: (1) performed data-driven simulations of coronal mass ejections starting from the solar surface; (2) simulated the SW propagation along the Voyagers, Voyager, and New Horizons trajectories, and also at Uranus (Fig. 1); (3) modeled shocks propagating through the LISM and demonstrated good agreement with observational data; (4) explained the increase in the frequency of plasma waves observed by V1 in the LISM by the presence of a heliospheric boundary layer of depressed plasma density at the surface of the heliopause (Fig. 2); (5) performed high-resolution simulations of the heliopause instability, identified the areas of possible magnetic reconnection near the heliopause, and predicted that V2 is likely to observe more reconnection than V1; (6) showed that the heliospheric magnetic field may be dissipating in the heliosheath producing turbulence in the SW flow; (7) analyzed the heliostall flow and quantity distribution in the heliospheric bow wave for different LISM conditions, and showed that the observed multi-TeV cosmic ray anisotropy may be explained by the LISM magnetic field distortion by the heliosphere; and (8) investigated the heliosphere structure from IBEX observations and the implications for the 11-year solar cycle in the heliosheath. The results are published in 10 papers and reported at over 20 scientific meetings.

Our collaboration with the Blue Waters team further promotes the application of adaptive technologies to contemporary plasma physics problems through the development of publicly available packages suitable for multiple applications.

**WHY BLUE WATERS**

First, neutral atoms are modeled kinetically, and we need on the order of 10^6 particles in our Monte Carlo simulations. These simulations require particle splitting, multiple grids, and coupling with the MHD module. They also produce multiple data sets sometimes exceeding 1 terabyte, which require hybrid parallelization. Computational region sizes are very large, as in the case of long tail simulations to about 10,000 AU (astronomical units). (3) Very deep adaptive mesh refinement is necessary near magnetic reconnection sites.

**PUBLICATIONS AND DATA SETS**


**Figure 1:** Simulated interplanetary magnetic field direction, which exhibits the Sun’s rotation effect, solar wind radial velocity, and proton number density (scaled to 1 AU value) at the time of the New Horizons spacecraft’s closest approach to Pluto. Two cross-sections are shown by the solar equatorial plane and the plane containing the solar rotation axis and Pluto.

**Figure 2:** The distributions of plasma density (left panel) along the V1 trajectory and its comparison with the plasma wave events detected by the spacecraft beyond the heliopause (right panel).
EXECUTIVE SUMMARY

Clusters of galaxies are both a useful probe of cosmology and a laboratory for understanding galactic feedback processes. However, modeling galactic-scale feedback processes in the context of a cluster presents a computational challenge because of the large dynamic range involved. Through the use of a highly scalable N-body/Smooth Particle Hydrodynamics code running on Blue Waters, our project is beginning to tackle this challenging problem. Preliminary results show that models that have successfully reproduced the morphology and number densities of field galaxies can also produce realistic models of cluster galaxies. Large computational resources with high-performance networks are necessary for these calculations.

RESEARCH CHALLENGE

Groups and clusters of galaxies are the largest bound objects in the Universe, containing more than a third of the warm-hot diffuse gas and a significant fraction of the galaxies in the universe. Consequently, understanding the physical processes that occur in group and cluster environments, including the interactions among the dark matter, hot diffuse gas, stars, and active galactic nuclei (AGN), is key to gaining insights into the evolution of baryons and galaxies across the age of the universe. Furthermore, galaxy clusters are one of the few places where the majority of baryons are visible via X-ray and microwave. In contrast to field galaxies, where feedback from supernovae and AGN puts gas into a mostly invisible circumgalactic medium, feedback from cluster galaxies will impact the state of the intracluster medium (ICM). Hence, clusters will provide very tight constraints on our understanding of galactic feedback processes. Clusters of galaxies are also key probes of cosmology and large-scale structure. Their size makes them visible via X-ray and microwave. In contrast to field galaxies, where feedback from supernovae and AGN puts gas into a mostly invisible circumgalactic medium, feedback from cluster galaxies will impact the state of the intracluster medium (ICM). Hence, clusters will provide very tight constraints on our understanding of galactic feedback processes. Consequently, understanding the physical processes that occur in group and cluster environments, including the interactions among the dark matter, hot diffuse gas, stars, and active galactic nuclei (AGN), is key to gaining insights into the evolution of baryons and galaxies across the age of the universe. Furthermore, galaxy clusters are one of the few places where the majority of baryons are visible via X-ray and microwave. In contrast to field galaxies, where feedback from supernovae and AGN puts gas into a mostly invisible circumgalactic medium, feedback from cluster galaxies will impact the state of the intracluster medium (ICM). Hence, clusters will provide very tight constraints on our understanding of galactic feedback processes.

METHODS & CODES

This project uses the highly scalable N-body/hydrodynamics code ChaNGA to model the formation and evolution of a population of galaxies in a Coma-sized galaxy cluster, including their contribution to and interaction with the ICM. This code is built on the Charm++ [1] parallel programming infrastructure. It leverages the object-based virtualization and data-driven style of computation inherent in Charm++ to adaptively overlap communication and computation and achieve high levels of resource utilization on large systems. The code has been shown to scale well to 0.5 million cores on Blue Waters [2]. The ChaNGA code includes a well-constrained model for star formation and feedback, and improved implementation of supermassive black hole formation, growth, mergers, and feedback [3,4]. In our previous Blue Waters work we have demonstrated that these models can reproduce populations of field galaxies at intermediate-to-high redshift [5], and can reproduce the observed stellar mass-halo mass relationship of galaxies from dwarfs up to galaxy groups [6]. Our simulations will be compared to observations of cluster galaxies to understand the physical and temporal origin of their morphologies. The model ICM will be compared to X-ray and microwave data (via the Sunyaev–Zeldovich effect) to understand the relation between these observations and the underlying gas properties. Finally, the overall mass distribution will be used to better understand how these clusters act as gravitational lenses for background galaxies.

RESULTS & IMPACT

We have completed a prototype simulation of a cluster one-tenth the mass of our target cluster, but at the same mass resolution. Even this simulation is pushing the state of the art in modeling galaxy clusters. Preliminary analysis of this cluster indicates that we continue to produce the observed stellar mass to halo mass even at 10^10 solar masses. The model central galaxy also displays a morphology typical of observed brightest cluster galaxies: an old, extended stellar distribution and a complete lack of cold gas.

WHY BLUE WATERS

Our scientific goals require modeling over a large dynamic range in mass and space. We have demonstrated that we need mass resolutions on the order of 10^8 solar masses to accurately follow star formation and galaxy morphology. In addition, we need to model a galaxy cluster on the order of 10^9 solar masses that is comparable to those observed over a range of redshifts. Hence, 10 billion particles are needed. Such a simulation can only be run on the largest computers available. Furthermore, the long-range nature of gravity requires a high-performance, low-latency network to perform the calculation.
KINETIC SIMULATIONS OF LARGE-SCALE PLASMA TURBULENCE

Allocation: NSF PRAC/3,850 Knh
PI: Vladimirovitch K
Co-Pi: Yuri Omelchenko
Collaborators: John Podesta, William Matthews
1Space Science Institute
2University of New Hampshire
3University of Delaware

EXECUTIVE SUMMARY

This project seeks to gain a better understanding of turbulence in hot, rarified plasmas typical of space and astrophysical environments. Specifically, we use large-scale kinetic simulations tracking upwards of a trillion particles at a time to analyze how the energy of turbulent motions is dissipated in such plasmas. The insights generated by the project may explain such long-standing puzzles as the properties and origins of the solar wind or the anomalously high temperature of the solar corona.

RESEARCH CHALLENGE

Plasma turbulence plays a significant role in the dynamics of many systems in the universe, from laboratory fusion experiments, to the Sun, and to astrophysical objects such as accretion disks. While parameters, geometry, and some aspects of the physics may differ among these systems, there is also a large degree of universality, which makes understanding plasma turbulence a grand challenge problem relevant to many fields of study. This project seeks to conduct simulations of plasma turbulence using codes that are capable of faithfully describing microscopic physical effects. This is important since plasma turbulence is a truly multiscale phenomenon, where the very nature of physical processes governing dynamics changes with scales. For example, in plasmas that are sufficiently hot and not too dense (as is typical of space and astrophysical systems), processes operating at small scales determine the ultimate fate of the turbulent energy. Depending on which of the many possible processes dominates, the energy could be transferred to different species (e.g., electrons, protons, or heavier ions), or to distinct populations of the same species. Understanding these issues may help advance our knowledge of the systems where turbulence operates. For example, both the solar wind and the solar corona are significantly hotter than can be explained with simple models, with local heating by turbulence often proposed as one of the likeliest explanations.

METHODS & CODES

The most complete description of the plasmas of interest is provided by the Vlasov–Maxwell equations, a six-dimensional system of partial differential equations. In order to solve these equations, we use the particle-in-cell technique, where plasma is represented as a collection of particles while equations describing electromagnetic fields are solved on a grid. A typical large-scale simulation can simultaneously track upwards of a trillion particles in order to obtain reliable statistics. This requires petascale computational resources, such as Blue Waters. Some of the simulations performed in this project were guided by and directly compared against observations conducted by the WIND spacecraft in the solar wind.

RESULTS & IMPACT

During the first year of the project, we have mainly targeted two problems: study of current sheet formation in large-scale simulations with conditions modeling solar wind and the so-called selective decay problem. Solar wind is a stream of plasma, emitted by the Sun, which fills interplanetary space. Interaction of the solar wind with the Earth’s magnetic field gives rise to the Earth’s magnetosphere and its complex dynamical behavior, often referred to as “space weather.” Space weather has a significant impact on us: For example, extreme solar events may affect or damage the power grid and communication satellites. Understanding of the solar wind in general and its turbulence in particular may help develop better models that improve our ability to understand and forecast space weather. Moreover, thanks to the measurements performed by multiple spacecraft operating in the solar wind, it is the best-studied example of large-scale plasma turbulence, making it also very interesting from the perspective of basic research.

While considerable progress has been made, many puzzles remain. One of them concerns identifying specific mechanisms responsible for conversion of turbulent energy into thermal energy (heat) and energetic particles. In typical space plasmas, collisions between particles are relatively rare and cannot provide sufficient randomization, as would be the case in fluid or regular gases. The relevant theoretical ideas focus on collective processes that invoke interaction of charged particles with electromagnetic fields. Those processes are thought to operate most efficiently in the regions where the magnetic field is highly sheared. Geometrically, these regions typically resemble shears or strings embedded in the plasma and are referred to as the current sheets. Due to their small transverse scale, which often approaches internal plasma scales, the dynamics of current sheets inherently depend on microscopic (kinetic) physics.

On the other hand, the large-scale motions in plasma form current sheets. Thus, in order to satisfactorily describe the dynamics of current sheets, simulations need to resolve a large range of scales and include the relevant microscopic physics. We have conducted simulations using a hybrid approximation that treats heavier species in plasma (e.g., protons) using a first-principle kinetic formulation, while representing lighter electrons using an approximate fluid model. The simulations were initialized with conditions modeling solar wind plasma and were directly compared against spacecraft observations. The results were in remarkable agreement with observations [1], indicating that the hybrid model can properly describe dynamics of current sheets at scales comparable to proton kinetic scales. A detailed characterization of the current sheets was performed as the first step necessary to identify the physical processes responsible for their formation and evolution. Ultimately, this will help advance understanding of energy dissipation mechanisms in the solar wind.

The second problem considered in the project concerns the problem of selective decay in kinetic plasma. Broadly speaking, the concept of selective decay is the notion that a turbulent system decays toward special states that, somewhat counterintuitively, are characterized by a high degree of correlation among various quantities. A large body of existing work investigated these issues within the framework of a macroscopic approximation known as the magnetohydrodynamics (MHD) model. In the more complete kinetic model used in the present study, some of the inherent symmetries of MHD are broken, which may lead to significant differences between predictions of MHD and the kinetic models for decay toward certain states. Fig. 1 demonstrates a turbulent field in one of the simulations conducted for the study. While a full analysis of the results is ongoing, preliminary results point to substantial differences with existing predictions, possibly signaling previously unappreciated significance of microscopic physics for global evolution.

WHY BLUE WATERS

The simulations needed to address the scientific questions of this project solve a global problem that cannot be split into a series of smaller simulations. As such, they require large memory, fast on-node computation, and fast internode communications. For this reason, they require an HPC resource like Blue Waters and cannot be conducted on cloud resources.

PUBLICATIONS AND DATA SETS

**METHODS & CODES**

The GAMER code [3] has the following important features.

- Hybrid OpenMP/MPI/GPUs. GAMER uses GPUs as PDE solvers and CPUs to manipulate the AMR structure. It uses OpenMP for intra-node parallelization in CPUs, MPI for inter-node communication, and CUDA as the GPU programming interface.

- Overlapping computation. CPU computation, GPU computation, and CPU–GPU communication are allowed to overlap, greatly improving the overall throughput when CPUs and GPUs take a similar time to complete their own tasks.

- Hilbert space-filling curve for load balance.

- Efficient usage of memory. GAMER stores all the data in the CPU memory and only temporarily transfers data to GPU, which allows exploitation of the large CPU memory. Communication overhead between CPU and GPU is usually negligible since it can be overlapped by both CPU and GPU computations efficiently using CUDA streams.

- Bitwise reproducibility. The order of all floating-point operations in GAMER is carefully designed to be deterministic. It thus supports bitwise reproducibility when (1) running simulations with different numbers of MPI ranks and OpenMP threads, and (2) restarting simulations from checkpoint files. This feature is essential for scientific reproducibility.

- Inline analysis with yt. GAMER supports the ability to pass in-memory data structures to yt [4], a powerful Python-based package for analyzing and visualizing volumetric data. Moreover, using yt allows one to share the data analysis scripts to the community straightforwardly, which greatly improves the scientific reproducibility.

**RESULTS & IMPACT**

We have compared both the performance and accuracy between GAMER and FLASH based on the merging galaxy clusters simulation setup of [1]. These simulations include hydrodynamics, self-gravity, dark matter particles, and AMR, and thus provide a comprehensive test for the simulation codes. Fig. 1 shows the slices of gas temperature evolution visualized with yt, demonstrating that the results of the two codes are remarkably consistent. Fig. 2 shows the strong scaling. Blue and green lines indicate the comparison with a relatively lower resolution of 7 kiloparsecs.
EXECUTIVE SUMMARY

We perform magnetohydrodynamic (MHD) simulations in full general relativity (GR) of the collapse of radially unstable, uniformly rotating, massive stars to black holes (BHs). The stars spin at the mass-shedding limit, account for magnetic fields and obey a polytropic $\Gamma = 4/3$ equation of state (EOS). The calculations lift the restriction of axisymmetry imposed in previous simulations. Our simulations model the direct collapse of supermassive stars (SMSs) to supermassive BHs (SMBHs), with masses larger than $10^4 M_\odot$, at high cosmological redshifts $z$, which may explain the GRB sources. Thus, at high cosmological redshifts $z$, which may explain the electromagnetic and gravitational wave signatures of these events and discuss the detectability of such multimessenger sources.

RESULTS & IMPACT

Using our latest adaptive-mesh refinement GRMHD code we performed simulations of massive stars that undergo collapse [5]. We consider a purely hydrodynamic case, and two magnetized cases, one seeded with a poloidal magnetic field only in the stellar interior and the other extended from the stellar interior to the exterior (Fig. 1, upper left). In both cases, the adopted magnetic fields are dynamically unimportant initially. To ensure reliable evolution of the exterior field, we initially impose a low-density atmosphere such that the stellar exterior is described by a constant plasma parameter, defined as the ratio of the gas pressure to the magnetic pressure. To accelerate the collapse, the pressure is initially depleted by 1% in all our cases. We then evolve the stars through the collapse and subsequent BH formation immersed in a magnetized accretion disk (Fig. 1, top right and bottom panels). We find that, in terms of the BH mass, its spin and torus mass, the results from our hydrodynamic simulations are consistent with previous semi-analytic estimates and axisymmetric simulations in GR reported in [7-8]. We also find that the magnetic field do not affect these global quantities. In the magnetized cases, following BH formation we observe the formation of magnetically dominated regions above the BH poles where the magnetic field lines have been wound into a collimated helical funnel, within which the plasma flows outwards. This collimated outflow is mildly relativistic, and constitutes an incipient jet. Our analysis suggest that the Blandford-Znajek effect is likely operating in our simulations and could be the process powering these jets. The magnetization in the funnel reaches values $\gtrsim 200$, and since for steady-state, axisymmetric jets the magnetization approximately equals the jet terminal Lorentz factor, the jets found in our simulations may reach Lorentz factors of $\gtrsim 200$, and hence, explain gamma-ray burst (GRB) phenomena. The accretion torus lifetime is $\Delta t \sim 10^7 (1 + z)(M/10^5 M_\odot)$. Thus, collapsing SMSs with masses $\gtrsim 10^4 M_\odot$ at $z > 10$ are candidates for ultra-long GRBs, while collapsing Population III stars at $z \sim 5-8$ are candidates for long GRBs. We estimated that for observation times of $\sim 10^5$ s, FERMI and SWIFT could detect such ultra-long GRB events from these stars [9].

PUBLICATIONS AND DATA SETS


RESEARCH CHALLENGE
Numerical simulations of TDEs of Sun-like stars by supermassive BHs are extremely challenging because it is difficult to numerically resolve the debris stream because it is very thin relative to the black hole. Tilted disk simulations require high resolution to properly resolve nonaxisymmetric turbulence in the tilted disks. Both are huge numerical challenges.

METHODS & CODES
Using our new code H-AMR (pronounced “hammer”), which includes adaptive mesh refinement and efficiently runs on GPUs, we were able to overcome these challenges.

RESULTS & IMPACT
Until now, many different simplifications have been adopted when simulating TDEs. For instance, to save computational time, (1) stars were sent in on closed, elliptic orbits (instead of parabolic ones) or (2) the BH-to-stellar ratio order of magnitude was assumed to be lower than in reality (e.g., 1,000 instead of 10⁶) [1,2]. The adaptive mesh capabilities of H-AMR and the high efficiency with which it runs on GPUs allowed us to properly resolve the thin debris stream for a typical encounter featuring both a parabolic orbit and a mass ratio of 10⁶.

The first simulations of tilted accretion disks were carried out a decade ago [3]. They demonstrated that tilted disks undergo Lense–Thirring precession. However, whether such disks are capable of producing relativistic outflows, or jets, and whether those jets point along the disk axis or the BH spin axis has remained a mystery. Our simulations for the first time established that such disks are indeed capable of producing jets and that the jets undergo precession together with the accretion disk. This is the first demonstration that jets can be used as probes of disk precession.

Blue Waters was instrumental in enabling our simulations, which are the highest resolution GRMHD simulations to date and set the new state of the art in the field. On a CPU (central processing unit) cluster, the effective simulation cost would have been 400-million CPU core-hours, accessible only on Blue Waters.

WHY BLUE WATERS
Our simulations require a high degree of parallelism as they run on hundreds to thousands of GPUs in parallel.
ADVANCED SPACE WEATHER MODELING

EXECUTIVE SUMMARY

Our effort aims to achieve a breakthrough advance in the understanding of space weather. The most destructive forms of space weather are due to major solar eruptions: fast coronal mass ejections (CMEs) and eruptive X-class flares. These destructive events originate with magnetic fields emerging from the solar interior, forming the active regions from where CMEs erupt into the heliosphere. Upon impacting the Earth, interplanetary CMEs impact the magnetosphere and produce geomagnetic storms. This process is controlled by the microphysics of magnetic reconnection. Our goal is to answer the most salient questions of space weather: How the buildup of magnetic energy results in solar eruptions and how magnetic reconnection results in geomagnetic storms.

RESEARCH CHALLENGE

Major space weather events are caused by large-scale expulsions of magnetized plasma from the Sun, which are known as coronal mass ejections (CMEs) and that typically travel to Earth in one to three days. These eruptions occur frequently, as often as several times per day during solar maximum, and cause geomagnetic storms by triggering sudden reconfigurations of the magnetosphere by magnetic reconnection. Extreme space weather events are caused by the most energetic CMEs, which drive sudden and extensive changes in the Earth’s magnetic field producing among other effects large-scale electric impulses that can melt transformers and cause cascading blackouts. Repair times for replacing the high-voltage transformers is estimated to be several months.

METHODS & CODES

We combine the efficiency of global fluid-type models with the physical capabilities of expensive but physically accurate local kinetic models. The resulting magnetohydrodynamic with embedded particle-in-cell (MHD-EPIC) model is 100 to 10,000 times more efficient than a global kinetic model. In addition, we found that the kinetic scales can be artificially increased, which can dramatically—by many orders of magnitude—reduce the computational cost of the embedded PIC model. Fig. 1 shows an MHD-EPIC simulation of the magnetosphere of Earth.

RESULTS & IMPACT

We have used this unique opportunity to simulate space weather events with the MHD-EPIC model, where the reconnection is handled by a kinetic PIC code. With this approach, we focused on modeling of the fundamental process of reconnection and its impact on the global dynamics. Currently, the MHD-EPIC model is the first three-dimensional global study of the complex process reconnection process using a high-fidelity kinetic model for the magnetic reconnection. We also made breakthrough advances in simulating flux emergence at active-region scale in spherical geometry. This work addresses the most salient questions of space weather.

WHY BLUE WATERS

Our project uses the Blue Waters petascale computing facility to perform unprecedented space weather simulations. This capability allows us to simulate magnetic flux emergence from the convection zone into the corona to form active regions that may result in coronal mass ejections. Using Blue Waters allows us to model a whole active region with sufficient grid resolution to capture magnetic energy build up. Blue Waters also allows us to model the reconnection process in the magnetosphere with the MHD with embedded PIC model and gain a better understanding of the intricate interaction between the small kinetic scales and the global scales that result in magnetospheric storms.

PUBLICATIONS AND DATA SETS


EXECUTIVE SUMMARY

Our team is simulating brief events in the interiors of evolved stars that result in ingestion of unprocessed new fuel into convection zones above nuclear burning shells. The new fuel can burn very violently under the much hotter conditions in the convection zone after reaching a sufficient depth within it. This burning sets off a series of reactions that dramatically affects the nucleosynthesis of heavier elements and, hence, the ultimate expulsion of heavier elements into the surrounding interstellar gas. Expulsion can be either a relatively slow expulsion of the outer envelope of the star or by an explosion of the star itself, if it is sufficiently massive. This work involves large and very detailed 3D simulations of the entirety of the stellar interior, for which Blue Waters is ideally suited.

RESEARCH CHALLENGE

Our simulations involve brief but important events in the lives of stars that can greatly affect the heavier nuclei that the stars produce. We have been concentrating so far on hydrogen ingestion flashes in which unburned hydrogen-rich fuel is brought into the convection zone above the helium-burning shell. The ingested hydrogen reacts with carbon in the convection zone to set off a sequence of nucleosynthesis reactions that is called the $i$-process, since the neutron fluxes that result are intermediate between the slow and rapid $s$- and $r$-process nucleosynthesis. In the beginning of our project on Blue Waters, we simulated hydrogen ingestion events in evolved stars, and during the last year we have been gearing up to attack the more challenging problem of the potential merger of multiple nuclear burning shells in massive stars. This has involved detailed studies of the ingestion process, particularly in massive star contexts, as well as the aggressive development of a new simulation code. Our results on $i$-process nucleosynthesis are important as inputs for the study of the chemical evolution of galaxies. Our new work with massive stars, just begun, could have a large impact on the conditions just before those stars explode, and also on the injection of heavier elements from these explosions into the surrounding interstellar medium.

METHODS & CODES

Our work to date simulating hydrogen ingestion flashes exploits the piecewise parabolic method (PPM) coupled with the piecewise parabolic Boltzmann (PPB) moment-conserving advection scheme for the multifluid volume fraction. PPM delivers more than double the resolving power of the PPB scheme for the single, very important variable representing the volume fraction of entrained fluid. Together with the already high resolving power of PPM, we are able to obtain very accurate results on a uniform grid. We must simulate a great many large-eddy overturning times in the convection zone above a nuclear burning shell in order to accurately approach a nonlinear, global oscillation of the burning of ingested hydrogen that increases the hydrogen ingestion rate by as much as two orders of magnitude. We are able to cover this long approach to the violent ingestion event because our PPM code scales to nearly 14,000 nodes on Blue Waters, at which scale it advances the simulation by roughly 20 time steps per second. Thus, the millions of time steps we need to simulate an ingestion event accurately are practical on Blue Waters with our code.

We have been turning our attention this year to massive stars, where the ingestion of material from above a burning shell of, for example, oxygen can allow the convection zone above that burning shell to eat its way outward in radius until it reaches the carbon-burning shell above it. A merger of these two burning shells can then result: Simulating this process is very challenging. We have been studying the relevant ingestion process in considerable detail over this last year with the goal of alternating between 1D and 3D simulation in order to span the time necessary in leading up to a shell merger. Our 3D simulations would keep models used in the ID intervals that stitch one 3D run to the next validated as a half to the development of a completely new code. This code adds a Level 3 adaptive mesh refinement (AMR) grid that will enable us to contain multiple nuclear burning shells and their respective convection zones in a single simulation. It is designed to scale to 14,000 nodes while running roughly twice as fast as our older code per node by exploiting 32-bit precision and GPU (graphics processing unit) acceleration [1–3].

RESULTS & IMPACT

We are producing a database of detailed simulations that investigates the phenomenon of convective boundary mixing at unprecedented accuracy for convection zones that extend over ranges in radius of more than a factor of two (see www.lcse.umn.edu). Global convection modes play an important role in these situations, making simulation difficult and costly [4–6]. Convective boundary mixing plays an important role in stellar evolution. In particular, in ingestion events that we study, it can have a dramatic impact on nucleosynthesis, which in turn affects galactic chemical evolution [7].

WHY BLUE WATERS

We have carried out our simulations on Blue Waters because of its special ability to enable our simulation code to be run at a sufficiently large scale that our large computations can each be completed in less than one week [8]. This allows our team to pose questions and get answers on a timescale that is conducive to productive thought and dynamic adjustment of our research direction.

PUBLICATIONS AND DATA SETS

For publications, see reference list in back of book. For shared data sets, see www.lcse.umn.edu.
High-Resolution Digital Surface Models of the 2016 Mw 7.8 Kaikoura Earthquake, New Zealand

Image Processing to Build a Multi-Temporal Vegetation Elevation Model (MTVEM) of the Great Lakes Basin (GLB)

The Terra Data Fusion Project

Large-Eddy Simulation of Sediment Transport and Hydrodynamics at River Bifurcations, Using a Highly Scalable Spectral Element-Based CFD Solver

The Reference Elevation Model of Antarctica

Untangling Entrainment and Precipitation in Convective Clouds

Understanding the 4-D Evolution of the Solid Earth Using Geodynamic Models with Data Assimilation

Physics-Based Modeling of High-Frequency Ground Motions and Probabilistic Seismic Hazard Analysis

Enhanced Digital Elevation Model for the Arctic

Simulating the most devastating tornadoes embedded within supercell thunderstorms

High-Resolution Earth System Modeling Using Blue Waters' Capabilities

Location-Specific Space Weather Hazards to Electric Power Grids Calculated on a Global Scale

Impact of Ocean Coupling on Simulated Tropical Cyclone Activity in the High-Resolution Community Earth System Model

Petascale Modeling of Convective Storms Under Climate Change and Variability

Impacts of Orientation and Morphology of Small Atmospheric Ice Crystals on in-situ Aircraft Measurements: Scattering Calculations

Pure-Scale Simulation of Multiphase Flow in Porous Media with Applications to Geological Sequestration of Carbon Dioxide

3D Particle-Resolved Aerosol Model to Quantify and Reduce Uncertainties in Aerosol–Atmosphere Interactions

Particulate Matter Prediction and Source Attribution for U.S. Air Quality Management in a Changing World
HIGH-RESOLUTION DIGITAL SURFACE MODELS OF THE 2016 Mw7.8 KAIKOURA EARTHQUAKE, NEW ZEALAND

EXECUTIVE SUMMARY

The 4D evolution of topography is critical for understanding faulting and landslide processes and the linkages between the two. For major events that deform and erode landscapes, such as earthquakes, we lack data at an event scale for understanding such processes because of the prohibitive cost of repeat land and airborne surveys, and because pre-event data is usually lacking. However, recent advances in high-performance computing coupled with stereo-satellite imagery collection allow the opportunity to provide regional-scale, high-resolution topographic time-series. Such evaluations could more accurately constrain fracture zones and landslide parameters for future models and could more accurately assess the impacts of landslides for the affected population. The November 14, 2016, Mw7.8 Kaikoura, New Zealand, event represents the first opportunity to execute such efforts for a regional scale, large-magnitude earthquake because a data for a large area of New Zealand were collected with stereo satellite imagery following the 2011 Christchurch earthquakes.

RESULTS & IMPACT

Regional assessment of landsliding and infrastructure damage in the aftermath of large earthquakes is a broad societal problem involving loss of life and property. Because hazard assessment is limited by access and resources in the immediate aftermath of an earthquake, satellite-based photogrammetry methods using high-resolution imagery and high-performance computing potentially provide an avenue to rapidly assess land surface changes and infrastructure damage in a cost-effective and time-sensitive manner. Whereby commonly used radar-based methods rely on coherent ground deformation for constraints on finite fault slip models, these techniques break down where there is surface disruption, such as surface faulting, landsliding, and infrastructure damage. New methods using photogrammetry techniques take advantage of in-line stereo imagery capture of high-resolution (0.5 m Ground Sampling Distance) Worldview satellite data. Digital surface models derived from these data provide valuable information for quantitative assessment of surface change. Immediate collection preserved these data in the most pristine form, and regional assessment provides a synoptic view not provided by typical post-event land- or airborne-based surveys. The third and fourth dimension provided by satellite-derived surface models will significantly advance auto-recognition algorithms by combining image analysis with 3D surface data.

METHODS & CODES

We want to understand how the landscape evolves in time after an earthquake, examining fine-scale land-surface rupture and landslide processes. We do this by making repeat models of the terrain in northeast South Island New Zealand from a series of stereo satellite images collected both before and after the Kaikoura Earthquake. We use the Surface Extraction with TIN-based Search-space Minimization (SETSIM) algorithm, developed at The Ohio State University, in a high-performance computing (HPC) workflow that bulk corrects hundreds of raw stereo satellite imagery pairs for image distortions and then performs a photogrammetric analysis. We order the surface model output in time and then use co-registering routines to make sure that the models are then further oriented in the right place in space. With this time series of landscape change we can see the immediate effects of the earthquake, how faults move in three dimensions, and how the landscape adapts to a huge number of mass flow events.

WHY BLUE WATERS

Blue Waters-scale computing was necessary to complete the project in a timeframe useful to coordinate perishable field data collection and provide rapid feedback on the event history. Maps and volume estimates of current slope failures estimated from the surface model will immediately assist New Zealand partners in monitoring and assessing dam hazards and slope reactivation, which will continue to affect the area for months to come. Our code and workflow is optimized for Blue Waters and allowed us to instantly activate the processing chain, which would likely have taken weeks to months on cloud resources.
IMAGE PROCESSING TO BUILD A MULTI-TEMPORAL VEGETATION ELEVATION MODEL (MTVEM) OF THE GREAT LAKES BASIN (GLB)

EXECUTIVE SUMMARY

Deriving ecosystem characteristics from stereo imagery has been used for decades, yet frequent 3D digital mapping and monitoring has not been feasible until very recently. By more frequently monitoring surface changes in the vegetation throughout the growing season in the Great Lakes Basin with satellite remote sensing techniques, land managers will be supported with new, enhanced information to address emerging stand-to-landscape scale changes in ecosystem habitats. The amount of data in a study area of this size needed to be processed and analyzed is well beyond those available from academic, private, and government systems. This computational reality is precisely why we need a leading-edge petascale resource such as Blue Waters. The results from this project will enable scientists to understand their land with greater detail in three dimensions, thereby making management, conservation, and protection of important ecosystems as modern and effective as possible, which will enhance the nation’s natural resources for future generations.

RESEARCH CHALLENGE

Ecosystem management requires knowing the type, size, structure, and density of vegetation over time. These important features need to be repeatedly mapped. Stereo submter, optical satellite imagery, and derived surface vegetation models can be used to better characterize these features, and their changes over time, with the added dimension of height. High-resolution vegetation surface canopy mapping over large geographic regions, such as the Great Lakes Basin (GLB), has never been obtained from either aerial or satellite surveys. Additionally, the binational management (Canada and U.S.) of the GLB limits consistent, repeatable coverage by either country working independently. While a few scattered vegetation surface models exist from expensive airborne-active laser sensors (LiDAR) within the GLB, these datasets represent a single date and were not planned as continuous, basinwide acquisitions. Not having this information limits the kind of science that can be done to address the multitude of questions that surround the ecosystems of the GLB. The question remains: How are the ecosystems of the GLB changing and what can we, as a society, do about it? Continuous monitoring of surface elevation will detect both natural changes (such as flooding, forest blowdown, fire, insects, and disease outbreaks) and anthropogenic changes (such as harvest and land cover change). Further, MTVEM will improve habitat and biological modeling. Finally, MTVEM will be used binationally to better visualize canopy change in forest habitats and freshwater wetland resources within the Great Lakes Basin.

METHODS & CODES

We currently have over 8,000 satellite image stereo pairs available and we estimate there will be an additional 5,300 pairs to be collected over the next 18 months, totaling about 13,300 image stereo pairs. Each stereo pair task is run on a single node, submitted in batches of 2 to 8 tasks per job. Complete processing of one stereo pair to 2m takes on average about 36 node hours, totaling about 483,590 node hours. We are adding ~10% time in addition for unforeseen events. Thus we asked for an allocation of 540,000 node hours. We will prioritize images based on science value to fit into our allotment, where the primary concentration is on the Great Lakes Basin but will extend temporal and geographical footprint as efficiencies and capacities improve.

RESULTS & IMPACT

As the data are processed, the resulting surface canopy models will become openly available through partner’s online distribution systems, such as NOAA’s Digital Coast and GEOSS Portal (www.geoportal.org). The final product, a seamless and registered surface elevation model (MTVEM) of the GLB, will enable a large range of science activities at substantially higher resolution than currently available. These canopy maps and change-detection products will provide positional accuracies of less than a couple meters with added ground control points. We have processed our entire archive of satellite image stereo pairs once. We are assessing change in priority GLB areas where LiDAR digital surface models from 2011 (DSM) are available (Figs. 1 and 2). Both gains and losses in vegetative cover over the five-year difference in the acquisition were detected in these examples. In addition, we are able to begin looking at seasonal differences by processing surface models from satellite stereo image pairs from within a single growing season. These preliminary results show great promise for providing valuable data to a myriad of coastal and terrestrial ecosystem science research questions that need to be addressed across the entire GLB.

WHY BLUE WATERS

Stereo satellite imagery allows for the generation of highly accurate surface elevation models and we have already tasked stereo-mode acquisition through DigitalGlobe over the entire GLB. Processing of this dataset will require an allocation well beyond those available from most academic, private, and government HPC systems, including the standard XSEDE allocations, which is precisely why the leading-edge petascale resource Blue Waters is necessary to address this research.

PUBLICATIONS AND DATA SETS

THE TERRA DATA FUSION PROJECT

Allocation: Blue Waters Professor/140 Kb
PI: Larry Di Girolamo
Collaborators: Guang-Hao Zhan, John Towns, Shaoan Wang, Yan Liu, Kent Yang
1University of Illinois at Urbana-Champaign
2National Center for Supercomputing Applications
3The HDF Group

EXECUTIVE SUMMARY

The Terra Data Fusion Project that we initiated in 2015 continues to make significant progress through collaborative efforts among NASA, HDF Group, and NCSA. Large Terra data transfers (~1PB) involving the Blue Waters team and NASA are now complete. The software tool that fuses all the Terra radiance granules was developed and has produced one year of basic fusion data on Blue Waters. Besides the scientific usage demonstrated in the previous reports, the dataset has been further used to (1) characterize ice crystal roughness of cirrus clouds, resulting in a better understanding of ice cloud optical properties, and (2) gauge the effects of cloud heterogeneity on microphysical retrievals, concluding that cloud spatial heterogeneity is insufficient to explain the large biases in the MODIS Re standard product.

RESEARCH CHALLENGE

The Terra satellite was launched in 1999 and continues to collect Earth sciences data using five instruments: the Moderate-Resolution Imaging Spectroradiometer (MODIS), the Multi-angle Imaging SpectroRadiometer (MISR), the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), the Clouds and Earth’s Radiant Energy System (CERES), and the Measurements of Pollution in the Troposphere (MOPITT). Terra data is among the most popular of NASA’s datasets, serving not only the scientific community but also governmental, commercial, and educational communities.

The need for data fusion and for scientists to perform large-scale analytics with long records has never been greater [1]. The challenge is particularly acute for Terra, given its growing data volume (~1 petabyte), the storage of different instrument data at different NASA centers, the different file format and projection, and inadequate cyberinfrastructure [2]. We recently initiated the Terra Data Fusion Project to tackle two long-standing problems: (1) How do we assemble and deliver Terra data fusion products; and (2) How do we facilitate the use of Terra data fusion products by the community in generating new products and knowledge through national computing facilities, and disseminate these new products and knowledge through national data sharing services?

Solutions to these questions will: 1) facilitate greater ease in creating new geophysical retrieval algorithms that provide greater accuracy than the current single-instrument algorithms, 2) provide an easy mechanism for users to access and process the entire Terra record, (3) greatly reduce error and redundancy in the science community among these researchers using multiple instrument datasets, (4) provide greater insight into geophysical processes through synergistic use of fusion products, and (5) provide a framework for fusion that could extend to other NASA missions and constellations. The end result will facilitate discovery and accelerate progress in Earth Science research. Use cases are presented below.

METHODS & CODES

Key steps in the Terra Data Fusion Project include: (1) transferring the entire Terra record (Level 1B radiance; >1 petabyte) to Blue Waters from NASA centers, (2) building software optimized for whole-mission processing on Blue Waters to create basic fusion products, (3) optimizing data granularity and HDF API settings that best support parallel I/O on Blue Waters; and (4) archiving and distributing Terra fusion products through existing NASA services.

Thus far, key steps (1) and (2) are accomplished. We successfully transferred mission-scale radiance data from all the five Terra instruments to Blue Waters. We built a software tool to merge all the Terra radiance granules into one basic fusion granule, which contains not only radiance measurements but also their uncertainties, geolocation, sun-view geometry, and observational time. The tool has produced one year of basic fusion data on the Blue Waters. We are currently working on key step (3) to prepare for mission-scale processing with help from the Blue Waters team. The overview of this project and progress report was given at the MISR science team meeting and data users’ science symposium [3].

Initial science investigation using this Terra fusion dataset was carried out in two studies. Fusing MISR and MODIS products, we characterized the effective radius (Re) of the liquid cloud drop size distribution over oceans. Details were given in Blue Waters 2015 Annual Report and in [4]. Our results paint a radically different picture of the distributions of cloud drop sizes in our atmosphere compared to what was previously determined from the original MODIS data; e.g., 2 to 10 μm differences in the zonal means. We also processed the 15-year MISR data to examine the decadal trends in the Earth’s color and texture [5]. Globally, we show that the Earth has been appearing relatively more blue (up to 1.6% per decade from both nadir and oblique views) and smoother (up to 1.5% per decade, only from oblique views) over the past 15 years.

RESULTS & IMPACT

We continue to explore the scientific usage of this dataset in two new studies. The first was that we used the dataset to generate one full year (2013) of MISR-MODIS cloud element fusion data (~50 terabytes) and share the data with Prof. Ping Yang’s group at the Texas A&M University via GridFTP. The fusion data were used to build ice crystal roughness parameters that are important to our meteorological understanding of ice clouds. Our intermediate results were presented at recent meetings and conferences [6-7]. Further scientific analyses are being conducted.

The second was that we examined the effect of cloud spatial heterogeneity on satellite-retrieved liquid cloud Re using 13 years of MISR and MODIS fusion for the month of January. We did find some dependency of Re on cloud spatial heterogeneity. Further stratification of the biases by both cloud spatial heterogeneity and cloud optical depth for various solar zenith angles and geolocations reveals that more heterogeneous and optically thick clouds tend to have larger Re bias, indicating cloud spatial heterogeneity is insufficient to explain the large biases in the MODIS Re standard product. Our results were presented at the MISR data users’ science symposium [8].

WHY BLUE WATERS

Key advantages of using Blue Waters for access, usage, and distribution of Terra fusion products are that the Terra data and processing are local, with access and sharing that are global. It has been demonstrated that having the Terra data local, with processing tuned to a massively parallel system with excellent sharing services, in one of the largest storage and bandwidth computing facilities in the country, provides an optimum framework for large-scale processing, analytics, and mining of the entire Terra record. In addition, the project staff provides expertise critically needed to optimize workflows. Terra is just one earth science data set. Fusion with other instrument records and meteorological reanalysis data (all of which are growing exponentially) for advancing earth science requires—and will continue to require—a Track-1 system that is accessible by the community.

PUBLICATIONS AND DATA SETS

Liang, L., Di Girolamo, L., and Sun, B. MODIS cloud drop effective radius for oceanic water clouds as deduced from optical thickness variability across scattering angles. J. Geophys. Res. Atmos., 120:15 (2015), DOI: 10.1002/2015JD023256. [The main data set applied to this study is the fusion product generated by fusing the MISR Level 1B radiance product (Version F03_0024) and the MODIS level2 cloud product (Collection 6).]

Zhao, G., et al., Regional changes in Earth’s color and texture as observed from space over a 15-year period. IEEE Trans. Geosci. Remote Sens., 54:7 (2016), DOI: 10.1109/TGRS.2016.2538723. [The main data set applied to this study is the MISR level1B radiance product (Version F03_0024).]
LARGE-EDDY SIMULATION OF SEDIMENT TRANSPORT AND HYDRODYNAMICS AT RIVER BIFURCATIONS: USING A HIGHLY SCALABLE SPECTRAL ELEMENT-BASED CFD SOLVER

EXECUTIVE SUMMARY

Bifurcations are fundamental features of all river systems. This study focuses on a specific class of bifurcations called diversions. Experiments have shown that the distribution of near-bed sediment between the downstream channels at a diversion is not proportional to the water flow distribution, with a disproportionately higher amount of sediment going into the lateral channel. A better understanding of the aforementioned non-linear phenomena will help in efficient design of river diversions. Diversions are used for navigation and flood mitigation, and have also been put forth as a solution for reclaiming deltas going under the sea due to increasing sea level. The current study investigates the mechanisms behind this phenomenon through Large Eddy Simulations (LES) and Direct Numerical Simulations (DNS) of the flow in idealized diversions of different configurations, with the sediment being modeled as Lagrangian particles. The simulations have been conducted using a highly scalable spectral element based incompressible Navier-Stokes solver, Nek5000. The simulation conditions are comparable to the laboratory experiments, which make these simulations large enough to require the use of Blue Waters. The simulation results provide new insight into the mechanism behind the phenomenon, and have identified yet unexplored parameters that influence the distribution of sediment at a diversion.

RESEARCH CHALLENGE

Bifurcations are fundamental features found in river systems. This current study focuses on a specific class of bifurcations called diversions, where one source channel splits into two: a main channel and a lateral branch. Laboratory experiments conducted by Bulle in 1926 [1,2] and later by other investigators, have shown that the distribution of near-bed sediment between the downstream channels is not proportional to the flow distribution, with a disproportionately higher amount of sediment going into the lateral channel. This non-linear phenomenon is often referred to as the Bulle Effect. The current study investigates the mechanisms behind this phenomenon through high-resolution numerical simulations of the flow and sediment transport at the scale, and for the configurations similar to Bulle’s experiments. A better understanding of the aforementioned phenomenon will help in efficient design of river diversions, which among uses like navigation and flood-mitigation have also been put forth as a solution for reclaiming deltas sinking under the sea due to rise in sea level, which is a consequence of anthropogenic climate change [3]. A prime example is the Mississippi River delta, for which different potential diversion designs are being studied for diverting water and sediment from the Mississippi River [4]. Better understanding of the phenomenon will eventually help in more accurate prediction of the short and long-term geomorphological evolution of river bifurcations, thus furthering the state of art in the field of river mechanics. It will also provide insights that will help improve reduced-order models of flow and sediment transport at bifurcations. Better understanding of the fundamental mechanism behind Bulle-Effect will also help shed light on vorticity-driven transport at bifurcations in human body (e.g. the carotid bifurcation [6]) and manmade systems (e.g. grit chambers in water reclamation plants [5]).

RESULTS & IMPACT

High-resolution Large Eddy Simulations (LES) and Direct Numerical Simulations (DNS) of the flow at different configurations of the idealized diversions have been conducted, with sediment being modeled as Lagrangian particles. A simulation being LES or DNS depends on the Reynolds number of the flow, as simulations were conducted for a range of Reynolds numbers 10–25,000. For the cases with Reynolds number of 25,000 the conditions are comparable to the laboratory experiments. The simulations have been conducted using the open-source, spectral element based higher-order incompressible Navier-Stokes solver Nek5000 [7]. The Spectral Element Method (SEM) combines the accuracy of spectral methods and the flexibility of Finite Elements Method [8]. Sediment transport in the current study has been modeled using Lagrangian particle tracking. For simulating transport of poly-disperse sediment particles efficiently, a semi-implicit Lagrangian particle algorithm was developed for the current study [1]. Initial results had shown that for a 90-degree diversion, and the flow being equally divided between the channels, the simulation was able to capture the preferential movement of bedload sediment into the lateral-channel [2]. Hence, simulations were conducted for a range of Reynolds number and flow divisions, further confirming the primary driver of the phenomenon that is preferential movement of the flow near the bottom of the channel into the lateral-channel [3,4]. This characteristic of the flow was also observed for different diversion angles (e.g. for 30,150 degrees see fig. 1). Simulations were also conducted for sediment with different diameters, which made them travel at different depths in the water column. It was found that the phenomenon was not just valid for sediment traveling as bedload, but for any sediment traveling at the lower 30% of the water column. To illustrate that, thousands of neutrally buoyant fine particles were released and their paths monitored. These particles by design will follow the flow, thus providing a sense of how the flow divides between the two channels. Figure 2 illustrates one of the above mentioned cases, where most of the particles starting near the bottom of the channel can be observed to primarily move into the lateral-channel. This study not only provides new insights into the hydrodynamics and sediment transport at bifurcations, it also shows that high-resolution LES can be used to study complex river-mechanics problems.

WHY BLUE WATERS

The current study pushes the limit of the scale at which high-resolution LES have been used to study complex multi-phase river mechanics problems, warranting the use of a computational resource that can provide sustained computing power at an unprecedented scale, thus the need to use Blue Waters. For the current study simulations have been conducted for up to 243,648 million computational points, with the code scaling strongly up to 32,768 MPI ranks. Without access to supercomputers like Blue-Waters, completing the study within a realistic timeframe would be impossible. Visualization of a phenomenon is an effective way to understand (and explain) it’s mechanics, thus we are currently working with Blue-Waters project staff to create an animation of the phenomenon, using data from one of the simulations.

PUBLICATIONS AND DATA SETS


Figure 2: The figure illustrates the division of neutrally distributed (in the water column) neutrally buoyant fine particles between the two channels, for Reynolds number of 25,000 and 30-90 flow division. The particles in the panels (a), b) are colored according to their position in the water column. Particles near the bottom are blue and those near the top are yellow. (c) shows that most of the blue particles have moved into the lateral channel and most of the yellow particles have continued into the main channel.
THE REFERENCE ELEVATION MODEL OF ANTARCTICA

Allocation: Innovation and Exploration/1,200 Knh
PI: Ian Howat1
Collaborators: Paul Morin2, Claire Porter2

1The Ohio State University
2Polar Geospatial Center of the University of Minnesota

EXECUTIVE SUMMARY

The goal of this project is to create the Reference Elevation Model of Antarctica (REMA), a continuous, high-resolution (8 m), high-precision (accuracy better than 1 m) reference surface for a wide range of glaciological and geodetic applications. REMA will be constructed from stereo-photogrammetric Digital Elevation Models (DEM) extracted from pairs of sub-meter resolution Worldview satellite imagery and vertically registered using ground control from both GPS (Global Positioning System) surveyed points and coordinated airborne LiDAR (Light Detection and Ranging) surveys by Operation IceBridge of the U.S. National Aeronautics and Space Administration. The Worldview imagery is archived and provided at no cost from the Polar Geospatial Center (PGC), which will also openly distribute REMA through its web data portal. REMA processing will be done using fully automated DEM extraction and coregistration software on the Blue Waters supercomputing system.

RESEARCH CHALLENGE

Accurate surface elevation is an essential dataset for glaciology, required for mapping bed topography from ice thickness; measuring ice thickness changes; constraining ice flow and geodynamic models; mapping glacial geomorphology, terrain corrections, and filtering of remote sensing observations; and many other science tasks. It is also critical for mapping ice traverse routes, landing sites, and other field logistics planning. Continuous DEMs of the continent, however, have spatial resolutions of hundreds of meters or more, accuracies of tens of meters in coastal and mountainous areas, and no definitive time stamping for change detection, limiting their utility. The growing archive of sub-meter stereo imagery held by the Polar Geospatial Center is arguably among the most underutilized collections of satellite data in existence, with nearly complete coverage of the polar regions below 85° of latitude. The project will be the first continental-scale application of these data for terrain mapping, upscaling an established processing and distribution system that has been used to map large regions of the Arctic successfully, and will be the next step toward eventual global mapping.

METHODS & CODES

In collaboration with PGC, our team has spent three years developing an efficient algorithm for constructing photogrammetric DEMs from satellite imagery with the objective of creating a fully automated system capable of handling large amounts of data. Development of the Surface Extraction from TIN-based Search-space Minimization (SETSM) algorithm was begun to facilitate an automated processing pipeline for the PGC operations. SETSM DEMs have been extensively validated; SETSM is node parallelized using OpenMP and has been applied to processing large-area DEM mosaics in proof-of-concept studies. SETSM’s structure uniquely eliminates the need for an existing (i.e., “seed”) DEM for a priori constraint or any data-specific, user-defined search parameters, making it a truly automated algorithm. SETSM is called from a single command line with the only required inputs being the filenames of the two stereo images and the RPC (Rich Photoenseric) formatfile, typically provided in XML (eXtensible Markup Language) format. SETSM is written in stand-alone C-code with no external dependencies and requires no libraries, ensuring simple, multi-platform installation, support, and optimization. SETSM is available as open source on Github.

RESULTS & IMPACT

Antarctica’s ongoing rapid changes—and potential for near-future change—are of global concern due to increased rates of sea level rise. REMA will provide a benchmark for detecting these changes. It will be useful for a wide range of applications beyond glaciology, ranging from geodynamics to logistics planning.

As we near the end of the first year of the project, we are well ahead of schedule in our activities thanks to receiving a 1.2-million node-hour “innovation award” allocation for the Blue Waters supercomputer at the National Center for Supercomputer Applications (NCSA). This allocation has allowed us to complete a full, double-coverage processing of the entire Antarctic continent and subantarctic islands to 8-m resolution, totaling nearly 0.5 million individual DEMs, and resulting in 98% coverage prior to filtering. Using the remaining allocation, we are now selecting gaps in coverage due to both prior data unavailability or cloud cover, and we are processing additional coverage to fill gaps.

Processed data are undergoing manual quality control by our student assistants. In addition, we are implementing automated filtering and processing methods to improve the efficiency and consistency of the filtering. We have also finalized our algorithms for mosaicking data into seamless elevation model tiles for distribution. As a test, we produced a prototype set of tiles for the Thwaites Glacier area.

In collaboration with colleagues at the University of Washington, we have successfully implemented a method for registering the DEMs to Cryosat-2 radar altimetry, providing an elevation accuracy within the 1-m specification for REMA. Cryosat-2-derived registration has been tested against airborne LiDAR elevations. The spatial and temporal coverage of Cryosat-2 makes it an ideal source of DEM registration.

Finally, we have worked with the PGC to develop a web distribution system that will be implemented soon.

WHY BLUE WATERS

Our processing requires scheduling and rapid throughput of thousands of individual jobs, each with highly variable and unpredictable wall times. Only Blue Waters has the capacity to provide efficient throughput of this job volume. Further, our processing was able to take advantage of Blue Water’s backfill scheduling capabilities, further increasing efficiency and system utilization.
UNTANGLING ENTRAINMENT AND PRECIPITATION IN CONVECTIVE CLOUDS

EXECUTIVE SUMMARY

Internal cloud circulation patterns introduce dry air inward from outside the cloud, which is called entrainment. Its effects can limit storm development, longevity, and various interdependent microphysical processes that may ultimately produce precipitation. Our understanding of entrainment and precipitation links has been limited in the past by inadequate model resolution. We are using high-resolution 3D simulations of convective clouds and storms, along with our diagnostic entrainment algorithm, to untangle the intricate web of connections between entrainment and its effects upon the generation of precipitation. Our latest results show that closer spacing between storms along a storm line can initially delay and decrease precipitation by competing for air flowing into the cloud bases; decreases in entrainment may not appear until the spacing is so small that the storms’ edges are no longer distinct. If ultimately the closer-spaced storms can generate even weak precipitation outflows, they can combine due to their proximity to produce a new generation of stronger storms that precipitate much more.

RESEARCH CHALLENGE

Deep convective clouds produce the majority of the earth’s precipitation, and yet it is difficult to predict whether developing cumulus clouds will attain the depth and longevity required to produce heavy rainfall and/or become severe thunderstorms. Entrainment is the term for the process by which the circulation within clouds bring dry air from outside the cloud in stored. In time, entrainment not only reduces the cloud buoyancy, limiting its vertical development, but also depletes its liquid water by evaporation, limiting precipitation formation. A long-standing problem in meteorological models has been to understand why they tend to predict rain formation too early and in excessive amounts. While others are researching what details of precipitation processes might be incorrectly represented in models, our approach is to investigate if the under-prediction of entrainment could explain the over-eagerness of the models to produce convective rain. This problem affects a broad range of atmospheric science problems, ranging from short-term weather forecasts from numerical weather prediction models to climate forecasts from regional and global climate models.

METHODS & CODES

We are using the National Center for Atmospheric Research’s CM1 model [1] to simulate convective clouds and storms at high resolution by employing its MPI capabilities on the many nodes available on Blue Waters. We make use of the National Severe Storms Laboratory microphysical scheme [2] within CM1. The simulations are conducted in both idealized and realistic environments. We evaluate entrainment with our new code [3] that calculates mass fluxes into the user-defined core of the cloud as the clouds and storms evolve and relate this entrainment to the ability of the clouds to produce precipitation and hail. We relate the calculated entrainment to the storm longevity and the amount of precipitation it produces.

RESULTS & IMPACT

Many atmospheric scientists have worked to try to produce parameterizations (larger-scale approximations) for cumulus entrainment, when the models they use for daily weather prediction cannot resolve the smaller cloud motions that can still be important for entrainment but over the larger domains required for thunderstorms. Blue Waters, with its huge number of nodes, its high speed, and its large storage capability for high-resolution model output and analysis allows us to push the spatial scale limit much farther than in the past. The hardware needed to run these kinds of simulations quickly exceeds the limits of most computers. Blue Waters staff have helped us to learn new and practical ways to visualize the output for easier analysis.

Figure 1: Visualization of the 3D core of a developing thunderstorm. Warm colors (yellow, red) denote areas of greater amounts of precipitation mass; cool colors (green, blue) denote areas with little water mass remaining. We are studying how (likely limited) entrainment resulting from the storm internal motions allows the greater precipitation mass to exist higher in the cloud. Once it falls out (not shown), the upper portions of the cloud exhibit multiple thermal circulations (nodes along the top of the storm).

Figure 2: Vertical cross-section through a portion of a line of convective clouds that are precipitating. Rain mass indicated by color bar below; wind motions and strength denoted by direction and magnitude of plotted vectors. Rain precipitating from adjacent clouds form downdrafts that may collide near the ground to form new strong storm updrafts in between the original clouds.

WHY BLUE WATERS

Our Blue Waters allocation is essential for achieving the high resolution required within a given simulation to properly represent the smaller cloud motions that can still be important for entrainment but over the larger domains required for thunderstorms and groups of thunderstorms. Blue Waters, with its huge number of nodes, its high speed, and its large storage capability for high-resolution model output and analysis allows us to push the spatial scale limit much farther than in the past. The hardware needed to run these kinds of simulations quickly exceeds the limits of most computers. Blue Waters staff have helped us to learn new and practical ways to visualize the output for easier analysis.
**EXECUTIVE SUMMARY**

We use numerical models with data assimilation to reproduce the complex processes of Earth dynamic evolution, which requires consideration of large amounts of data distributed over many computer nodes. With Blue Waters, our research projects during the past year include: 1) investigating mantle dynamics behind the widespread intra-plate volcanism associated with the Yellowstone hotspot, 2) simulating subduction and mantle flow beneath South America since 100 mega anum (100 million years). Our third project concerns the mechanical strength and its lateral variation of continental lithosphere [6]. We demonstrate that the electrical conductivity of rocks is an excellent proxy for the effective viscosity of the lithosphere [Fig. 2]. Consequently, this finding provides a practical way to “measure” the strength of the Earth’s rigid outer shell that is otherwise difficult to estimate. This also opens a new direction of geodynamic research. Our second project is to reproduce the subduction and mantle flow history below South America since 100 mega anum (100 million years). We found that the seismically observed flat slabs beneath the continent are mostly due to the subduction of thick and buoyant seafloor anomalies (oceanic plateaus) [3]. We also found that the formation of these flat slabs tears the slab itself, leading to the absence of earthquakes within the flat portion and formation of special volcanisms above [4]. Our third project concerns the mechanical strength and its lateral variation of continental lithosphere [6]. We demonstrate that the electrical conductivity of rocks is an excellent proxy for the effective viscosity of the lithosphere [Fig. 2]. Consequently, this finding provides a practical way to “measure” the strength of the Earth’s rigid outer shell that is otherwise difficult to estimate. This also opens a new direction of geodynamic research.

**METHODS & CODES**

To quantitatively understand the current and past dynamic processes within the deep Earth, we adopt various data assimilation techniques in large-scale geodynamic modeling. For example, we employ seismic and magnetotelluric tomography images to represent the present-day mantle temperature and viscosity profiles. We assimilate the motion of tectonic plates that are geologically reconstructed as velocity boundary conditions of a geodynamic model. We combine these data into a single physics-based numerical model using either sequential (forward) or adjoint (inverse) data assimilation techniques. The software package we use is a well-benchmarked finite element code, CitcomS [1]. We have further expanded the scalability and I/O performance of this code on Blue Waters.

**RESULTS & IMPACT**

In our first project, we evaluated the traditional hypothesis that Yellowstone volcanism is caused by a deep-rooted mantle plume [2]. We found that this hypothesis is flawed, because a hot upwelling plume is always blocked from above by the sinking oceanic plate. Therefore, Yellowstone has to have been formed by a different mechanism. This work challenges the traditional view on this important topic of intra-plate volcanism. Our second project is to reproduce the subduction and mantle flow history below South America since 100 mega anum (100 million years). We found that the seismically observed flat slabs beneath the continent are mostly due to the subduction of thick and buoyant seafloor anomalies (oceanic plateaus) [3]. We also found that the formation of these flat slabs tears the slab itself, leading to the absence of earthquakes within the flat portion and formation of special volcanisms above [4]. The slab-induced mantle flow dominates the deformation of the surrounding mantle [Fig. 1], as confirmed by the model predicting the observed seismic anisotropy in the region [5].

Our third project concerns the mechanical strength and its lateral variation of continental lithosphere [6]. We demonstrate that the electrical conductivity of rocks is an excellent proxy for the effective viscosity of the lithosphere [Fig. 2]. Consequently, this finding provides a practical way to “measure” the strength of the Earth’s rigid outer shell that is otherwise difficult to estimate. This also opens a new direction of geodynamic research.

**WHY BLUE WATERS**

The enormous amount of data processing and computation makes Blue Waters the best platform. The code CitcomS is community-based software, and this code has been designed and tested mostly on traditional supercomputers.

**PUBLICATIONS AND DATA SETS**


PHYSICS-BASED MODELING OF HIGH-FREQUENCY GROUND MOTIONS AND PROBABILISTIC SEISMIC HAZARD ANALYSIS

EXECUTIVE SUMMARY
A research collaboration, led by the Southern California Earthquake Center (SCEC), which includes earth scientists, engineers, and computer scientists, used Blue Waters to run physics-based earthquake simulations that improve our understanding of earthquake processes and their effects on seismic hazard. SCEC’s earthquake system science research program develops detailed earth models and high-performance computing software needed to perform realistic, physics-based earthquake simulations and motion simulations. This past year, the SCEC team used NCSA Blue Waters to develop more accurate and scalable computational models of earthquakes and to calculate the first physics-based probabilistic ground motion forecasts for Central California.

RESEARCH CHALLENGE
Probabilistic Seismic Hazard Analysis (PSHA) [1] is the scientific framework for many seismic risk and disaster engineering and retrospective design activities. The U.S. Geological Survey (USGS) currently uses empirical PSHA to promote seismic safety engineering and disaster preparedness across the United States, including California. SCEC’s research goal is to develop physics-based seismic hazard models for California and elsewhere that are more accurate than the empirical USGS National Seismic Hazard Map Project [2] standard models. Our long-term goal is to extend physics-based PSHA across the full bandwidth needed for seismic building codes and other purposes.

METHODS & CODES
This year, SCEC researchers added improved physics into our wave propagation software and improved our software’s performance on GPUs (central processing units) and CPUs (graphics processing units). For high-frequency ground motion simulations, our codes must model frequency-dependent attenuation [3], surface-free topography [4], and nonlinear yielding effects [5]. With improved codes and support through the Blue Waters PAID program, we performed the first 4-Hz nonlinear magnitude 7.7 earthquake simulation using 4,200 GPUs on Blue Waters [6,7] using a highly optimized implementation of a nonlinear computational method developed by SCEC researchers. We continued to validate our software by simulating well-recorded California earthquakes and comparing our simulations against the recorded ground motions [7].

Also this year, we used Blue Waters to perform CyberShake Study 17.3. This study applied the CyberShake [8] PSHA computational method to Central California for the first time. Study 17.3 calculated two seismic hazard models for Central California using a traditional 1D seismic velocity model and the other using a more accurate 3D velocity model, with results shown in Fig. 1. Results using the 3D velocity model show ground motion levels in the California Central Valley that are markedly lower than the levels produced from the standard ground motion prediction equations (GMPEs) currently in widespread use by earthquake engineers. This is an important contrast to our results in Southern California, where CyberShake predicts stronger shaking than the GMPEs in the deep, low-velocity sedimentary basins. The differences are related to the lateral extents of the basins, which govern their resonance frequencies and amplitudes. These results provide new evidence that CyberShakes physics-based approach can substantially improve our estimates of strong ground shaking.

We are preparing selected research codes to run efficiently on next-generation supercomputers. We have improved the performance of our wave propagation and Strat Green Tensor codes on next-generation GPUs and Xeon Phi systems [9]. To scale up the I/O performance of our software along with our improved compute performance, we optimized I/O performance of our anelastic wave propagation (AWP) software by increasing our use of third-party HPC (high-performance computing) I/O libraries including AIOUS, HDF5, and PnetCDF.

RESULTS & IMPACT
CyberShake simulations for Southern California are under review as inputs to a new Los Angeles urban seismic hazard map under development by the USGS. The SCEC committee for Utilization of Ground Motion Simulations (UGMS) is working within the framework of the Building Seismic Safety Council activities to develop long-period, simulation-based, spectral response acceleration maps for the Los Angeles region. Our CyberShake hazard maps are under consideration for inclusion in the National Earthquake Hazards Reduction Program, the American Society of Civil Engineers 7–10 Seismic Provisions, and for the Los Angeles City building codes. The UGMS group is using CyberShake simulations to quantify the effects of sedimentary basins and other 3D crustal structures on seismic hazard. The SCEC’s Off-fault Deformation and Digital Earth software is the first tool in the National Institute of Building Sciences’ Project 17 recommendations for tall buildings. It has no overt statement to say that our sustained work on Blue Waters is transforming and modernizing earthquake science and engineering, and thus represents a major contribution to Strategic Goal 1 (Transform the Frontiers of Science and Engineering) of the NSF 2014–2018 Strategic Plan [10].

WHY BLUE WATERS
SCEC’s earthquake system science research program needs access to Blue Water’s super computing resources for several reasons. SCEC computational research requirements continue to expand in many ways including in terms of algorithmic sophistication, geographical range, and time resolution. New simulations require more computational, memory, and storage resources. Our computational demands continue to grow because our calculations do not yet span the full range of resolution and parameter space, not all important physics have yet been included, and because individual earthquake simulations do not “solve” a problem when run just once or twice. Great uncertainty remains in the ground motions expected in future earthquakes, and society would be remiss in delaying a better resolution of such a critical scientific and public safety challenge.

PUBLICATIONS AND DATA SETS


BLUE WATERS ANNUAL REPORT

78

79
ENHANCED DIGITAL ELEVATION MODEL FOR THE ARCTIC

EXECUTIVE SUMMARY

The Earth’s land surface topography is arguably the most fundamental single data set in the geosciences, geographical sciences, and civil engineering. It is essential to research ranging from the location of rivers and the extent of watersheds in hydrology, to permafrost collapse in built-up areas, to the change in the shape of volcanoes in volcanology. The Polar Geospatial Centre (PGC) and its partners at The Ohio State University and Cornell University will adapt PGC’s digital elevation model (DEM) production capabilities from small-area, on-demand production to systematically process and mosaic the entire Arctic sub-meter stereo commercial imagery archive. Such a DEM would not only catapult the Arctic from the worst- to among the best-mapped regions on Earth, it would also allow for precise detection of change over time, creating a benchmark for measuring this rapidly evolving landscape.

RESEARCH CHALLENGE

There is a lack of high-resolution, consistent, high-quality elevation data available for the Arctic. In 2000, the Shuttle Radar Topography Mission began by acquiring Synthetic Aperture data for the Earth that was processed into an elevation model with a 30-m posting. This mission had limited geographic coverage between the latitudes of 60⁰N and 56⁰S because of the shuttle’s limited orbital inclination. The National Geospatial Intelligence Agency (NGA), Digital Globe (DG), and PGC have built up a near-seamless archive of polar sub-meter stereo imagery that consists of almost 70,000 stereo pair strips from the Worldview 1, 2, and 3 satellites. Using photogrammetric algorithms, we are able to construct digital elevation models (DEM) from the stereo pairs, allowing for mapping of surface features at the same scale as Airborne LiDAR (Light Detection and Ranging) without the cost or logistics constraints of LiDAR. This imagery collection increases at a rate of more than one California-area equivalent per day while there is sufficient Arctic sun. These data will be used by the Arctic research community to support activities that include transportation, national defense, land management, sustainable development, and scientific studies. Further, repeat DEMs with frequencies of months or even days, can be used for change detection, with applications ranging from studies of land use, to transportation, national defense, land management, sustainable development, and scientific studies. Further, repeat DEMs with frequencies of months or even days, can be used for change detection, with applications ranging from studies of land use, to transportation, national defense, land management, sustainable development, and scientific studies.

METHODS & CODES

Our team has spent three years developing an efficient algorithm for constructing photogrammetric DEMs from satellite imagery with the objective of creating a fully automated system capable of handling large amounts of data. Development of the Surface Extraction from Terrain-based Search Space Minimization (SETSM) algorithm was begun to facilitate an automated processing pipeline for the PGC operations. SETSM DEMs have been extensively validated [1], are node parallelized using OpenMP, and have been applied to processing large-area DEM rasters. Once the source imagery is corrected, SETSM takes the two source images and the RPC (remote procedure call) file, typically provided in XML format.

The DEM extraction workflow starts with a preprocessing step that corrects the source imagery for sensor-specific detector alignment artifacts and outputs a GeoTIFF-formatted set of source rasters. Once the source imagery is corrected, SETSM takes the two source images and derives increasingly detailed elevation models using its pyramid-based approach.

RESULTS & IMPACT

Thus far, we have produced over 57,000 individual 2-m posting DEMs that total approximately 80,000,000 km² of the Arctic. This means that the Arctic is covered four times, on average, though some areas are more poorly covered and some have over 100 time steps. These data were also processed into continuous mosaics for over 92% of the 20,000,000 km² Arctic. All of these data have been released to the science community and the public through ArcticDEM.org, and Esri has developed viewer and Amazon Web services. These data are now being used by scientists, national geographic surveys, and regional and local governments for a broad range of scientific, civil engineering, and mapping applications.

WHY BLUE WATERS

No other academic computer had the capacity, at the time, for this project. Blue Waters was able to execute the ArcticDEM workload without significantly impacting throughput of other projects. The project had a timeline that precluded the use of cloud services because of time required for development and porting of the SETSM code.
EXECUTIVE SUMMARY

Supercell thunderstorms produce the strongest, longest-lived tornadoes, ranked EF4 or EF5 on the Enhanced Fujita Scale. While such tornadoes are uncommon, they cause the vast majority of fatalities and damage. Forecasting such tornadoes requires a deeper understanding of supercell thunderstorms. We have simulated several EF5 tornadoes embedded within supercell thunderstorms at resolutions up to 15 meters, revealing flow features that greatly enhance our current understanding of tornado formation and maintenance. Interactions between the storm’s cold pool and updraft create and reorganize vorticity (spin and shear) in such a way as to concentrate it into streams and sheets that help initiate and maintain the EF5 tornado.

RESEARCH CHALLENGE

Tornadoes embedded within supercell thunderstorms create the strongest winds found in nature at the earth’s surface. Accurate forecasting of tornado behavior is of great benefit to millions of people who live in tornado-prone regions, such the Great Plains of the United States. However, the process of tornado formation, maintenance, and decay remains elusive to scientists. Our work aims to explore the factors involved in the creation within supercell thunderstorms of the strongest, most long-lived tornadoes.

METHODS & CODES

We used the CM1 model, developed at the National Center for Atmospheric Research. We modified the output driver of the model in order to enable efficient, memory-buffered, high-frequency, lossily compressed floating point HDF5 output, and wrote middleware to read, convert, and visualize raw model output utilizing techniques that included volume rendering and trajectory clouds. We explored different environmental conditions, and the most promising environments were simulated at extremely high resolution, resulting in EF5 tornadoes.

MODEL DATA are not currently published online; however, several scientific talks in video format, including model visualizations, may be found at http://orf.media. The code written thus far is on a private github server; we intend to make the code available to the public toward the end of this allocation.

RESULTS & IMPACT

Every year Americans die from tornadoes, despite attempts to warn the public of severe weather before it occurs. Our ability to forecast tornado behavior is limited by our current knowledge of how tornadoes form, are maintained, and decay in supercell thunderstorms. The focus of our research is on the most devastating, long-lived tornadoes embedded within supercell thunderstorms. We have simulated dozens of supercells at ultra-high resolution, and some of these supercells form long-lived EF5-strength tornadoes. These simulations have revealed newly identified flow features such as the streamwise vorticity current (SVC), a primarily horizontally oriented, helically flowing “tube” of rain-cooled air that is tilted vertically into the supercell’s updraft. The SVC helps maintain the updraft’s vigor near the ground, which appears to force tornado formation and assist in its maintenance.

WHY BLUE WATERS

In order to capture flow features that are critical to tornado formation, maintenance, and decay, the atmosphere must be resolved at extremely high resolution. As the resolution of the model is doubled, it requires approximately 16 times more computing resources. Our highest-resolution simulations (at 15-meter grid spacing) utilized 360,000 Blue Waters cores, or roughly half of the machine. Further, extremely high memory, communication, and I/O bandwidth were essential in these simulations, something which is not currently sufficient with XSEDE resources, much less cloud resources.

PUBLICATIONS AND DATA SETS

HIGH-RESOLUTION EARTH SYSTEM MODELING USING BLUE WATERS’ CAPABILITIES

EXECUTIVE SUMMARY
This work contributes to a larger body of ongoing research aimed at using high-resolution climate and weather forecast models to better understand high-impact events in the present day and future warmer scenarios. Simulations completed this year include those using the Community Earth System Model (CESM) at the highest resolution currently feasible for long climate integrations (0.25° atm/land – 1° ocn/ice) as well as the highest resolution currently possible (0.25° atm/land – 0.1° ocn/ice). For both resolutions, the simulations contribute to a set of control climate sensitivity, twentieth-century transient, and future scenario simulations. In addition to ongoing studies of tropical cyclones and midlatitude storms, we also investigated atmospheric rivers and other extreme events in the northwest Atlantic Ocean.

METHODS & CODES
For long climate integrations we use the Community Earth System Model (CESM), a coupled climate model for simulating the earth’s climate system. Composed of six component models that simulate Earth’s atmosphere, ocean, land surface, sea ice, land ice, river transport, and one central coupled component, the CESM allows researchers to conduct fundamental research into the earth’s past, present, and future climate states.

For regional downscaling over the contiguous United States, we use output from three GCMs—Community Climate System Model (CCSM4); Geophysical Fluid Dynamics Laboratory Earth System Model (GFDL-ESM2G); and Hadley Centre Global Environment Model, version 2-Earth System (HadGEM2-ES)—to force the Weather and Research Forecasting (WRF) model. A substantial number of science papers have been published and a number of papers are currently in preparation.

RESULTS & IMPACT
Building on research from our previous work, we expanded storm tracking into the midlatitudes to investigate potential changes in storm number and strength over the current century. As found with tropical cyclones, the total number of midlatitude storms increases in the future in all ocean basins (Fig. 1). The atmosphere tends to be more stable in the future in the midlatitudes, contributing to fewer storms. Atmospheric rivers (ARs), a subcategory of midlatitude storms that are able to transfer moisture from the tropics to the midlatitudes, had not previously been analyzed in the 0.25° atm/land-1° ocn/ice version of the CESM. Higher model resolution allows for better representation of counts and frequency of ARs. Results indicate that future changes in ARs depend on changes in the atmospheric jets and therefore respond differently in the different ocean basins.

Certain regional processes within the ocean may be much better represented by the high-resolution ocean (0.1°) simulations. An observed hotspot of localized warming in the Gulf of Maine is much better represented in this high-resolution simulation compared to the nominal 1° simulations. Such extremes in ocean temperatures can have a significant economic impact on fisheries.

Investigations of extreme temperature and precipitation change over the contiguous United States reveal startling realities for future climate. Two future climate scenarios were simulated: a moderate warming scenario (RCP4.5) and a strong warming scenario (RCP8.5). Fig. 2 shows the change in the number of days exceeding 95°F for the late twenty-first century under each forcing scenario. Some regions, such as the Northeast, will not experience much change due to the relative rarity of events in that region. While other areas, such as the middle part of the country, are projected to experience large changes in both scenarios. For much of country, there is an additional one to two months each year that will exceed 95°F with the RCP8.5 scenario as compared to RCP4.5. This is significant because in historical simulations, the vast majority of 95°F days take place only during June–July–August (JJA) for these regions. An additional 30–50 days per year means the “summer” months will span much more than just JJA, which would have significant energy and agricultural impacts.

WHY BLUE WATERS
Multiple century-long simulations are needed in order to quantify CESM model characteristics and sensitivity and to produce a sufficiently long, stable preindustrial control simulation, followed by historical and numerous future scenarios. Furthermore, ensemble members are needed to quantify and reduce uncertainty. The climate modeling community has refined horizontal resolution to 0.25° for the atmosphere and 0.1° for the ocean, allowing for a full eddy-resolving ocean simulation within the modeling system. These simulations and analyses, at high resolution, at a minimum require petascale computing resources and cannot be completed without a computational platform like Blue Waters. Because these simulations use a modest number of nodes for long periods of time, our project requires the help of the Blue Waters staff to achieve good throughput.

PUBLICATIONS AND DATA SETS
LOCATION-SPECIFIC SPACE WEATHER HAZARDS TO ELECTRIC POWER GRIDS CALCULATED ON A GLOBAL SCALE

Allocation: NSF PRAC/5,000 Knh
PI: Jameson J. Simpson¹
¹University of Utah

EXECTIVE SUMMARY

The largest documented geomagnetic storm due to a coronal mass ejection on the sun occurred in 1859. This storm caused telegraph operators communicating over 100 km-long wire lines to experience electric shocks, some nearly fatal. The historical record suggests that extreme space weather is likely to impact the Earth again in the future. However, modern electrotechnologies will be affected by space weather to a much larger degree than in the past. We are utilizing Maxwell’s equations models of the Earth–ionosphere waveguide to calculate location-specific space weather hazards to electric power grids in order to prevent blackouts. Blue Waters permitted us, for the first time, to study electric field behavior near ocean–continent boundaries using realistic coastal geometries. Additionally, we developed stochastic models of electromagnetic wave propagation through uncertain and variable ionosphere regions. These algorithms are allowing us to determine the confidence level that a communication or similar system will work as expected during disturbed conditions.

RESEARCH CHALLENGE

The largest documented geomagnetic storm on Earth resulting from a coronal mass ejection (CME) on the sun occurred in 1859. That storm caused telegraph operators communicating over 100 km-long wire lines to experience electric shocks, some nearly fatal. Further, business transactions requiring telegraphic exchanges were completely shut down in the world’s major capitals. A 2008 National Academies report indicates that extreme space weather events, “though rare, are likely to occur again sometime in the future.” However, a recurrence of an 1859-magnitude space weather storm could disrupt today’s society to a much greater degree than in 1859 due to the proliferation of vital but vulnerable electrotechnologies. Interruptions to radio communications, commercial airline flight plans, satellite operations, transportation, banking, financial systems, home and industrial computer electronics, and electric power grids are just some examples.

The focus of our research was to greatly enhance our understanding of the near-Earth electrodynamics associated with historically intense CMEs. The methodology of the proposed work was to advance and apply detailed, high-resolution Maxwell’s equations models of the Earth-ionosphere waveguide developed by the PI over the past ~15 years [1]. These models are based on the robust finite-difference time-domain (FDTD) method [2]. They currently extend from ~400 km below the Earth’s surface to an altitude of ~100 km. They uniquely account for the Earth’s complete topography, oceans, lithosphere composition, geomagnetic field, and magnetized ionospheric plasma according to altitude, position, and time of day, all while solving the full-vector Maxwell’s equations. As a result, those models created by the PIs group are the most advanced electrodynamics models of the Earth-ionosphere waveguide. They can model the entire world’s response to a space weather event and provide location-specific information on possible hazards to societal infrastructure. Previous studies were limited to only the sinusoidal steady-state and involved simplifying geometries and physics (e.g., infinite line currents, layered lithosphere models, nonphysical constant magnetic fields assumed within each layer, only solutions to Poisson’s equation, etc.).

METHODS & CODES

The first goal of our research was to help power grid stations better understand their individual risks to different space weather impact scenarios depending on their orientation and location. Of particular focus this past year was the electromagnetic field behavior in coastal regions in order to determine whether space weather poses unique hazards to power grids along coastlines. For this work, we used the FDTD method to calculate electromagnetic fields at the surface of the Earth near ocean-continent boundaries. As a time-domain method, FDTD permits modeling of arbitrary source time waveforms, variable current source orientations as shown in Fig. 1, and even the finite propagation velocity of the ionospheric currents. Further, as a grid-based method, FDTD permits modeling of complex geometries, such as sloping coastlines combined with finite depth oceans (rather than a coastline having a constant, infinitely-long slope as in the previous analytical studies).

The second goal of our research was to develop the first efficient, grid-based stochastic electrodynamics models of the Earth-ionosphere waveguide. Nearly all electrodynamics solvers assume average (mean) electrical properties of materials and solve for average (mean) electric and magnetic fields. However, assuming numerically only an average state of the ionosphere yields calculated output electromagnetic field waveforms that are not as rich and complex as measured electromagnetic fields. Further, there is great uncertainty in the content of the ionosphere at any given moment. The FDTD models we developed this past year solve for both the mean and variance of electromagnetic wave propagation through a varying/uncertain ionosphere.

RESULTS & IMPACT

In 2015, Prof. Simpson was among 20 researchers who participated in the NASA Living with a Star Working Group Institute on Geomagnetically Induced Currents. The issue of ocean-continent boundaries and whether they pose significant risks to power grids was among the list of topics that were considered important and unresolved. Blue Waters allowed us to take a closer look at this issue this past year using an established numerical technique (FDTD) that offers more flexibility than previously possible using analytical approaches, and provides more rigorous (full-vector Maxwell’s equations) solutions. In order for space weather to pose a unique risk to electric power grids at ocean-continent boundaries, high electric fields must extend at least 100 km inland from the coast. Counter to the conclusions of previously published work, we were able to determine that space weather does not induce intense electric fields over a sufficiently large area to pose a risk to power grids [3].

WHY BLUE WATERS

We were able to use Blue Waters to develop higher grid resolutions than previously possible by over an order of magnitude (~1 x 1 x 1 km vs. ~40 x 40 x 5 km). This has opened up a wide variety of new applications because we can model higher frequencies of electromagnetic waves and also model smaller geometries. For example, we are now working with the Defense Advanced Research Projects Agency to examine the use of our models for developing a new electromagnetic system for geolocation. We are also now working with the Office of Naval Research to detect objects submerged in the ocean. Other possibilities due to the higher grid resolutions achieved include communications during space weather events, new remote sensing applications, and studying very low-frequency signals detected by spacecraft around the times of earthquakes.

PUBLICATIONS AND DATA SETS

EXECUTIVE SUMMARY

This research seeks to advance our understanding of the relationship between tropical cyclones (TCs) and Earth’s climate system using a high-resolution state-of-the-art Earth system model, Community Earth System Model, or CESM. In this report, we highlight results from a set of 30-year simulations in which the high-resolution (25 km) atmosphere component is configured with three different levels of ocean coupling: prescribed sea surface temperature (SST), mixed-layer ocean, and full three-dimensional dynamic ocean with nominal 1-degree horizontal resolution. We find that the inclusion of ocean coupling significantly affects TC characteristics, including storm frequency, geographic distribution, maximum wind, and storm intensification. Key differences in storm numbers and distributions can be attributed to variations in the modeled large-scale climate mean state and variability that arise from the combined effect of intrinsic model biases and air–sea interactions. This work addresses the importance of storm-induced ocean–atmosphere feedbacks in Earth’s coupled climate system, which can help improve our understanding of how TC activity may change in the future.

RESULTS & IMPACT

We find that TC number, geographical distributions, and intensity are sensitive to ocean coupling (Figs. 1 and 2). Differences in TC characteristics are mainly attributed to model differences in local air–sea flux exchanges and large-scale climate conditions. The fully coupled model with dynamic ocean simulates the most realistic annual TC number in the northwestern Pacific, though an expanded Indo-Pacific warm pool is likely contributing to more TCs forming in the central Pacific and Indian Ocean. The coupled model underestimates TC activity in the north Atlantic and northeastern Pacific, which is mainly due to biases in sea surface temperature and vertical wind shear, as well as the southwestward shift of the Inter-Tropical Convergence Zone. The partially coupled simulation with a mixed-layer ocean exhibits key features of TC distributions similar to the fully coupled simulation, including the cold surface ocean temperature bias and the underestimation of TCs in the north Atlantic. By design, ocean heat transport in the mixed-layer ocean model is specified based on the modeled SST, and this bias in ocean heat transport in the coupled model is likely propagated to the partially coupled simulation. The uncoupled atmosphere-only simulation appears to perform better in the north Atlantic than the coupled model. However, the model exhibits an asymmetric bias in TC activity across the Pacific Ocean, including an overestimation of TC number in the eastern Pacific and underestimation of TC number in the northwestern Pacific. Our results show that this asymmetric pattern can be related to a weakened tropical zonal atmospheric overturning circulation (Walker Circulation) and associated biases in large-scale vertical motion.

METHODS & CODES

In this project, we assess the impact of ocean coupling on simulated TC activity using a high-resolution configuration of the CESM with a 25-km resolution atmosphere [2]. We performed three, 30-year simulations in which the atmosphere model is configured with three different levels of ocean coupling: (1) prescribed monthly-varying sea surface temperature based on the observed climatology; (2) a mixed-layer ocean model that allows thermodynamic exchanges between the atmosphere and the ocean mixed layer but does not account for ocean dynamics; and (3) a 1-degree ocean general circulation model with full dynamics and ocean–atmosphere fluxes. The models are configured with increased air–sea coupling frequency, modified surface wind drag law, and high frequency of history file output, in order to focus on ocean–atmosphere interactions associated with TCs. Each simulation is run for 30 years under the preindustrial climate conditions and with an active carbon–nitrogen cycle.

RESEARCH CHALLENGE

TCs are among the world’s deadliest natural hazards. How TC activity will vary with the changing climate is a topic of great interest. Recent research has shown that Atmosphere General Circulation Models (AGCMs) can be capable of generating TC-like vortices, and the representation generally improves with increasing horizontal resolution [1]. The current generation of high-resolution (less than 50 km) AGCMs have been shown to capture realistic global TC activity and can resolve the most intense Category 5 TCs [2-4]. Correct representation of air–sea interactions under TCs is important for simulating realistic storm intensities and track durations. Local feedbacks due to ocean mixing and surface fluxes can inhibit storm development and intensification as well as influence larger-scale ocean and atmospheric circulations. These feedbacks can influence TC characteristics and statistics in a global coupled Earth System model, and they can also contribute to the tropical mean state and variability on interseasonal to interdecadal timescales. AGCMs with prescribed surface ocean conditions do not capture these feedbacks and are thus missing important physical processes that can influence climate variability.

PUBLICATIONS AND DATA SETS


WHY BLUE WATERS

Given the substantial computational expense of high-resolution Earth system models, it is difficult to apply these models to study tropical cyclones because of the necessary grid resolution (¼ degree), integration length (multiple decades), and high frequency output (sub-daily). Blue Waters provides the unique capabilities to handle the computational demand associated with running the model at ultra-high resolutions, including scalability to over 15,000 cores, high-frequency input and output, and post-processing and visualization of model results.
PETASCALE MODELING OF CONVECTIVE STORMS UNDER CLIMATE CHANGE AND VARIABILITY

Allocation: Blue Waters Professor/240 Ksh
PL: Robert J. Trapp

EXECUTIVE SUMMARY

This research seeks to answer the basic question of how present-day extreme storm events might be altered by human-induced climate change. The pseudo-global warming (PGW) methodology has been adapted for this purpose. Modified atmospheric states drawn from global climate model (GCM) output were used to constrain an ensemble of Weather Research and Forecasting (WRF) model simulations of select extreme events. Highly resolved grids allow for process-based analyses of the simulated events; ensembles of such simulations facilitate quantification of uncertainty.

Our most recent work has focused on hail storms and tornadoes spawned by landfalling hurricanes. The severe hail events considered thus far are more intense but generate less-extensive and fewer hail swaths under PGW. Simulations of Hurricane Ivan (2004) are showing that the conditions under PGW promote a relatively more intense hurricane at landfall, and lead to a much higher incident of mesoscale vortices with tornadic potential.

RESULTS & IMPACT

Exemplifying the hailstorm results are the simulations of the May 19, 2013, outbreak of tornadoes, damaging wind, and hail (Fig. 1). Relative to the CTRL experiments, the PGW experiments across the three microphysical schemes (Morrison: MO, Milbrandt-Yau: MY, NSSL: NS) and three GCM drivers (GFDL, MIROC, and NCAR), the individual hailstorms under PGW tended to be more intense but generated less extensive and fewer hail swaths (Fig. 2). Our analyses thus far are showing that this response is due to the PGW-induced: (1) changes in the thermodynamic environment (e.g., higher convective available potential energy, higher convective inhibition, and higher freezing level), (2) increases in precipitation loading of individual updrafts, and (3) reductions in environmental vertical wind shear. One hypothesis that we are currently exploring is that a reduction in wind shear leads to a reduction in updraft area and thus in the area that accommodates hail growth.

The hurricane simulations are of Hurricane Ivan (2004), which spawned a record 118 tornadoes and caused significant damage to inland communities. We are finding that the conditions under PGW promote a relatively more intense hurricane at landfall and lead to a much higher incidence of mesoscale vortices with tornadic potential. The relatively higher convective available potential energy under PGW is one hypothesized physical explanation for this response.

In addition to continued analyses of the hailstorm and hurricane simulations described above, we are currently developing the capability to employ the Model for Prediction Across Scales (MPAS) for studies of convective storms under climate change and variability. MPAS is one of the emerging global atmospheric models with variable-resolution grids and provides us with the ability to isolate effects of low- and high-latitude processes (e.g., from Arctic sea ice and tropical oceans) on deep convective storms that are well resolved in middle latitudes. With the assistance of the National Center for Supercomputing Applications personnel, the MPAS model codes have recently been compiled on Blue Waters, and MPAS experimentation is currently under design.

WHY BLUE WATERS

The relatively small size of thunderstorms and the ranges of relevant scales within tropical cyclones, coupled with their episodic occurrence, necessitates a research approach that can account for temporal scales from decades to minutes and spatial scales of thousands of kilometers to hundreds of meters. In other words, we require very large geospatial domains that have fine grid point spacings and long-time integrations with high rates of model output. Moreover, quantifications of uncertainty require that such realizations be repeated over multiple experiments. The Blue Waters allocation is providing us with the resources needed to achieve this unprecedented level of climate simulation.
IMPACTS OF ORIENTATION AND MORPHOLOGY OF SMALL ATMOSPHERIC ICE CRYSTALS ON IN-SITU AIRCRAFT MEASUREMENTS: SCATTERING CALCULATIONS

EXECUTIVE SUMMARY

The single-scattering properties of ice crystals with maximum dimensions smaller than 50 μm were calculated at a wavelength of 0.55 μm using a numerically exact method (i.e., the discrete dipole approximation) and Mie theory. For these calculations, hexagonal columns and spheres were used to represent the shapes of the small ice crystals in natural clouds. Furthermore, because the morphological features of nonspherical ice crystals are closely related to their single-scattering properties, varying aspect ratios were used to characterize the hexagonal column shapes in the scattering calculations. The results show the impacts of orientation, shape, and aspect ratio on the directional intensity of scattered light.

METHODS & CODES

The discrete dipole approximation (DDA), a flexible technique that can calculate the scattering properties of irregularly shaped particles. In DDA, a particle is discretized into a number N of elementary polarizable units called dipoles. Specifying the location and polarizability of these dipoles allows calculations of the scattering and absorption of light. The number of dipoles N into which a particle is divided can be assigned to N CPUs (or cores) with the single-scattering properties then calculated within a parallel environment (e.g., message passing interface). In our study, a numerical code, the Amsterdam DDA (ADDA [1]) was used to calculate the scattering and absorption of electromagnetic waves by ice crystals.

RESULTS & IMPACT

We calculated the single-scattering properties (i.e., phase matrix, asymmetry parameter, and extinction efficiency) of hexagonal crystals with Dmax < 50 μm at a nonabsorbing wavelength (i.e., λ=0.55 μm) using Blue Waters. Because molecules in ice crystals form a hexagonal lattice structure, the most common crystal habits are hexagonal prisms. To represent natural variations of hexagonal ice crystals, six different aspect ratios (AR=0.5, 1.0, 2.0, and 4.0) were used. The single-scattering properties of hexagonal columns with a width of up to 36 μm and a length of up to 48 μm were determined.

The single-scattering properties of hexagonal ice crystals with AR=0.5, 1.0, 2.0, and 4.0 were calculated using DDA, and the results show that the impacts of this nonsphericity on derived particle size parameters are greater for those with Dmax<10 μm (due to interference structures) and with Dmax>30 μm (due to nonsphericity) than for those with Dmax>10 μm. The differences were up to 112% (170%) in the forward (backward) direction depending on the degree of nonsphericity. Calculations for larger particles are required and should be based on accurate calculations for larger size parameters. Although it is well known that the scattering properties of nonspherical ice crystals differ from those of spherical shapes, the impacts of this nonsphericity on derived in situ particle size distributions (PSDs), upon which satellite retrievals and parameterizations for large-scale numerical models depend, are unknown. To improve in situ airborne measurements of small crystals and PSDs, precise relationships among the intensity of light scattered in multiple directions by a particle and its size and shape are required and should be based on accurate calculations of single-scattering properties. Such calculations demand large computing time and memory that rapidly increase with particle size.

Figure 1: Non-zero scattering phase matrix elements of hexagonal ice crystals with aspect ratio of 1.0 at wavelength of 0.55 μm as functions of scattering angle and volume equivalent size parameter.

Figure 2: Calculated differential scattering cross sections of hexagonal ice crystals and spheres in forward (left) and backward (right) directions.

RESEARCH CHALLENGE

Current in situ airborne probes (e.g., forward scattering spectrometer probes) that measure the sizes of ice crystals with maximum dimensions Dmax less than 50 μm are based on the concept that the measured intensity of light scattered by a particle in the forward and sometimes backward direction can be converted to particle size. The retrieval of ice crystal size from satellites also relies on relationships between light scattering and particle size, as do parameterization schemes for numerical models. The relationship between particle size and scattered light used to process data from current forward scattering probes is based on Mie theory, which assumes the refractive index of a particle is constant.

The accuracy of radiative transfer models and satellite retrieval algorithms depends heavily on accurate calculations of single-scattering properties of ice crystals. Blue Waters is an important resource in completing these calculations by allowing more exact calculations for larger particles.

PUBLICATIONS AND DATA SETS


Um, J., and G. M. McFarquhar, Quantifying uncertainty in forward scattering probes due to non-sphericity of atmospheric ice crystals. 17th International Conference on Clouds and Precipitations, Manchester, UK, July 25–29, 2016.

PORE-SCALE SIMULATION OF MULTIPHASE FLOW IN POROUS MEDIA WITH APPLICATIONS TO GEOLOGICAL SEQUESTRATION OF CARBON DIOXIDE

EXECUTIVE SUMMARY

Advances in noninvasive imaging of rock along with continued developments in computing power open up the exciting possibility of direct numerical simulation of pore-scale physics that are relevant to a variety of subsurface energy processes. In particular, we can investigate the fundamental pore-scale flow processes that control migration and trapping of supercritical carbon dioxide during carbon capture and storage, which is a key technology to mitigate emission of greenhouse gases into the atmosphere. Direct numerical simulation of pore-scale multiphase flow physics is a grand computational challenge since a fine spatial grid is required to capture the complex pore geometry, while a large spatial grid must be included for a statistically representative sample. Although the lattice Boltzmann method (LBM) is generally known to be an effective numerical scheme, we did considerable work to develop a flexible code that is optimized for manycore processors. We used our code in conjunction with unique microfluidics experiments to investigate the role of inertial effects during rapid pore-scale displacements.

RESEARCH CHALLENGE

Understanding the migration of multiple fluids within pore spaces in subsurface geological formations is critical for addressing important problems such as enhanced oil recovery, groundwater pollution from leaking tanks or pipelines, geothermal energy production, and geological sequestration of carbon dioxide (CO2). The last application, capture and geological storage of CO2 provides a means of reducing CO2 emissions into the atmosphere, by capturing CO2 from major stationary sources and injecting it into suitable deep rock formations. Direct numerical simulation of CO2 invasion patterns in porous media under reservoir conditions is critical for addressing large-scale phenomena such as the relative permeability relationship and residual trapping of fluids. Therefore, it is of great importance to study the detailed pore-scale displacement patterns of CO2 injection in porous media under reservoir conditions.

RESULTS & IMPACT

Our further studies of the effects ignored in most pore-scale studies due to the complex geometry and complex interfacial dynamics, require further development of high-performance codes to utilize GPU computing power as well as to improve parallel efficiency.

EXECUTIVE SUMMARY

ADVANCES IN NONINVASIVE IMAGING OF ROCK ALONG WITH CONTINUED DEVELOPMENTS IN COMPUTING POWER OPEN UP THE EXCITING POSSIBILITY OF DIRECT NUMERICAL SIMULATION OF PORE-SCALE PHYSICS THAT ARE RELEVANT TO A VARIETY OF SUBSURFACE ENERGY PROCESSES. IN PARTICULAR, WE CAN INVESTIGATE THE FUNDAMENTAL PORE-SCALE FLOW PROCESSES THAT CONTROL MIGRATION AND TRAPPING OF SUPERCRITICAL CARBON DIOXIDE DURING CARBON CAPTURE AND STORAGE, WHICH IS A KEY TECHNOLOGY TO MITIGATE EMISSION OF GREENHOUSE GASES INTO THE ATMOSPHERE. DIRECT NUMERICAL SIMULATION OF PORE-SCALE MULTIPHASE FLOW PHYSICS IS A GRAND COMPUTATIONAL CHALLENGE SINCE A FINE SPATIAL GRID IS REQUIRED TO CAPTURE THE COMPLEX PORE GEOMETRY, WHILE A LARGE SPATIAL GRID MUST BE INCLUDED FOR A STATISTICALLY REPRESENTATIVE SAMPLE. ALTHOUGH THE LATTICE BOLTZMANN METHOD (LBM) IS GENERALLY KNOWN TO BE AN EFFECTIVE NUMERICAL SCHEME, WE DID CONSIDERABLE WORK TO DEVELOP A FLEXIBLE CODE THAT IS OPTIMIZED FOR MANYCORE PROCESSORS. WE USED OUR CODE IN CONJUNCTION WITH UNIQUE MICROFLUIDICS EXPERIMENTS TO INVESTIGATE THE ROLE OF INERTIAL EFFECTS DURING RAPID PORE-SCALE DISPLACEMENTS.

RESEARCH CHALLENGE

UNDERSTANDING THE MIGRATION OF MULTIPLE FLUIDS WITHIN PORE SPACES IN SUBSURFACE GEOLOGICAL FORMATIONS IS CRITICAL FOR ADDRESSING IMPORTANT PROBLEMS SUCH AS ENHANCED OIL RECOVERY, GROUNDWATER POLLUTION FROM LEAKING TANKS OR PIPELINES, GEOTHERMAL ENERGY PRODUCTION, AND GEOLOGICAL SEQUESTRATION OF CARBON DIOXIDE (CO2). THE LAST APPLICATION, CAPTURE AND GEOLOGICAL STORAGE OF CO2 PROVIDES A MEANS OF REDUCING CO2 EMISSIONS INTO THE ATMOSPHERE, BY CAPTURING CO2 FROM MAJOR STATIONARY SOURCES AND INJECTING IT INTO SUITABLE DEEP ROCK FORMATIONS. DIRECT NUMERICAL SIMULATION OF CO2 INVASION PATTERNS IN POROUS MEDIA UNDER RESERVOIR CONDITIONS IS CRITICAL FOR ADDRESSING LARGE-SCALE PHENOMENA SUCH AS THE RELATIVE PERMEABILITY RELATIONSHIP AND RESIDUAL TRAPPING OF FLUIDS. THEREFORE, IT IS OF GREAT IMPORTANCE TO STUDY THE DETAILED PORE-SCALE DISPLACEMENT PATTERNS OF CO2 INJECTION IN POROUS MEDIA UNDER RESERVE CONDITIONS.

RESULTS & IMPACT

OUR FURTHER STUDIES OF THE EFFECTS IGNORED IN MOST PORE-SCALE STUDIES DUE TO THE COMPLEX GEOMETRY AND COMPLEX INTERFACIAL DYNAMICS, REQUIRE FURTHER DEVELOPMENT OF HIGH-PERFORMANCE CODES TO UTILIZE GPU COMPUTING POWER AS WELL AS TO IMPROVE PARALLEL EFFICIENCY.
EXECUTIVE SUMMARY

This research aims at reducing key uncertainties in quantifying the impact of atmospheric aerosol particles on the Earth’s climate. Aerosol particles can be brought into the atmosphere by a wide range of human activities or by natural sources. They profoundly impact the large-scale dynamics of the atmosphere because they interact with solar radiation, both directly by scattering and absorbing light and indirectly by forming cloud droplets. These impacts depend on the particles’ sizes and their compositions, which continuously change in the atmosphere. In uncertainties in quantifying these impacts originate from scale interactions and the high computational cost required for modeling these. To tackle this problem, we developed the particle-resolved 3D model WRF-PartMC-MOSAIC, which has the unique ability to track size and composition information at a per-particle level. Particle-resolved simulations at the regional scale not only require efficient numerical algorithms but also a computational resource with the capabilities of Blue Waters. Together, these methods and the petascale resources allow for ultra-high-detail simulations that are needed to quantify the impact of aerosol particles on weather and climate at the regional scale.

RESULTS & IMPACT

We present results from the first-ever particle-resolved aerosol simulation for a realistic, spatially resolved three-dimensional domain (North Carolina). Aerosol and trace gas emissions were taken from the 2005 National Emission Inventory [5], and the meteorology corresponded to January 24, 2000. On the order of 100 billion computational particles were tracked in this simulation, including their composition changes due to gas-to-particle conversion, and coagulation events. Fig. 1 shows the modeling domain and the spatial distribution of black carbon-containing particle number concentrations near the surface after eight hours of simulation. Black carbon aerosol is of interest because of its adverse health impacts and because of its warming impact on climate. While this is a fundamental bulk quantity, common to any chemical transport model, the particle-resolved aerosol representation provides unprecedented detail of particle composition and source tracking. Fig. 2 (top) shows the originating sources of all particles within a given grid cell. This allows source attribution for any location within our domain. Fig. 2 (bottom) shows an example of the complex continuum of aerosol composition that exists within a single grid cell; particles of similar diameters can have very different chemical composition—information that is usually lost when using traditional aerosol models. The variations in particle composition are determined by their emission source characteristics, here with highway vehicles containing the largest black carbon mass fractions. During the simulation, aerosol composition evolves due to coagulation and condensation of secondary gas species, resulting in a complex continuum.

As the model tracks composition and source information of thousands of computational particles per grid cell, individual particles may also be explored. For example, a single particle with a particular size and black carbon mass fraction may be examined, marked with a red dot in Fig. 2. By tracking mass of constituent species (not shown), we can determine that the selected particle has grown considerably due to the condensation of nitrate. Additionally, by tracking source history (not shown), the composition of the aerosol emissions sources can be determined where this selected particle has undergone multiple coagulation events with particles from different emission sources such as agriculture and fossil fuel combustion. These capabilities will be useful in future studies for quantifying how much individual source categories are contributing to the pollution at a certain location. Aerosol modeling is challenging because of the multiscale nature of the problem—the macroscopic aerosol impact on climate is determined by microscale processes on the particle scale. The WRF-PartMC-MOSAIC model provides a tool that represents many of these microscale processes explicitly, which allows for an improved process-level simulation of the key interactions among aerosols, clouds, and radiation. This model framework therefore serves as the first benchmark for more approximate models, and provides a basis for rigorous coarse-graining to develop physically robust parameterizations for use in larger scale models.

WHY BLUE WATERS

Access to Blue Waters allows for a cutting edge model formulation that pushes both science and computing by combining the large-scale features of state-of-the-art 3D models with the process level physical representation of box models. Modeling 3D domains with on the order of 100 billion tracked particles creates many computational challenges due to computationally intensive equations per particle and memory requirements to track high-dimensional particle composition. To enable simulations of aerosols at both a high spatial and compositional resolution, there is a need for tens of thousands of cores, fast interconnections among those cores, and sufficient memory per process.

PUBLICATIONS AND DATA SETS

PARTICULATE MATTER PREDICTION AND SOURCE ATTRIBUTION FOR U.S. AIR QUALITY MANAGEMENT IN A CHANGING WORLD

EXECUTIVE SUMMARY

The objectives of this study are to better understand how global changes in climate and emissions will affect pollution in the United States, focusing on particulate matter and ozone, to project their future characteristics and quantify key source attributions. We are using a state-of-the-science dynamic prediction system that couples a global climate–chemical transport model with regional climate–air quality models over North America to determine individual and combined impacts of global climate and emissions changes on air quality, with uncertainty estimate, from the present to 2050 under multiplexible scenarios. We are doing the long-term global climate chemistry runs using Community Earth System Model CESM1.2 with fully coupled chemistry using CAM5-chem at 0.9° x 1.25° horizontal resolution, then comparing the results with observational data to evaluate the model simulation.

RESEARCH CHALLENGE

Our goal is to determine the individual and combined impacts of global climate and emissions changes on U.S. air quality from the present to 2050 under multiple scenarios, quantifying pollution sources and assigning their attribution—natural vs. anthropogenic emissions, national vs. international agents, natural variations vs. climate changes—with associated probability and uncertainty. We are developing a timeline for the global change factors to become significant such that effective actions can be taken. The level of significance is defined following the cross-state air pollution rule as one percent of nonattainment areas with the goal of bringing all areas into attainment for the National Ambient Air Quality Standards. Our hypothesis is that the integration of the most advanced modeling system, most updated emissions treatment, multiple processes representation, and a multi-climate–emission scenarios assessment will improve the predictive capability and result in more reliable projection of future changes in particulate matter, ozone, and related pollutants as well as their global and regional sources.

This research presents a state-of-the-science approach for advancing quantitative knowledge of the impacts of global changes in climate and emissions on U.S. air quality. The Global Climate Chemistry Transport model (GCCCT) integrates global climate change with long-range pollutant transport that links worldwide natural and anthropogenic source emissions, while providing lateral boundary conditions that drive the Regional Climate–Air Quality model (RCAQ) for regional climate and air quality prediction. RCAQ incorporates more comprehensive physical representation (surface, precipitation, convection, cloud, aerosol, and radiation); comprehensive chemical mechanisms (e.g., secondary organic aerosols and SOAs); and detailed emissions treatment. Hence, it more realistically simulates interactions between surface and atmospheric processes at regional–local scales that in turn affect local air quality. This nested GCCCT/RCAQ dynamic prediction system is being evaluated against observations, and subject to process-level understanding and source attribution of U.S. air quality episodes under present and future conditions.

METHODS & CODES

We are conducting three primary experiments using the dynamic prediction system: (1) historical simulations for period 1994–2013 to establish the credibility of the system and refine process-level understanding of U.S. regional air quality; (2) projections for the period 2041–2060 to quantify individual and combined impacts of global climate and emissions changes under multiple scenarios; and, (3) sensitivity analyses to determine future changes in pollution sources and their relative contributions, from anthropogenic and natural emissions, long-range pollutant transport, and climate change effects.

We are also conducting a series of 20-year runs using CESM1.2 (CAM-chem). We have completed short-term global CAM-chem simulations driven by NASA Modern-Era Retrospective analysis for Research and Applications (MERRA) reanalysis data at 0.9° x 1.25° horizontal. We used the FSTARMAM7 component set, which uses seven modes to model aerosols, prescribed ocean and ice, and CAMS physics with carbon and nitrogen in the Community Land Model (CLM). CAMS-chem has 160 species with 427 reactions, and has strict enforcement of the conservation of total (organic and inorganic) chlorine and total bromine under advection. The heterogeneous chemistry module has also been upgraded to reflect the underestimated of supercooled ternary solution and surface area density, in addition to an improved aerosol treatment and inclusion of aerosol–cloud interactions, with extensive tropospheric and stratospheric chemistry. It has 30 vertical levels with model top at about 40 km. The dust emission is calibrated in the model so that global dust aerosol optical depth is between 0.025 and 0.030, and the system couples tropospheric aerosol to chemistry through heterogeneous chemistry. In this run, we used CESM1.2 default emissions, which represent surface emissions of approximately 30 species of aerosols. The surface emission of each species is composed of all possible sources of emissions, including those from biomass burning, domestic sources, transportation, waste treatment, ships, industry, fossil fuels, and biofuels, and were composed from POET, REAS, GFEDv3 and FINN emissions databases [1].

RESULTS & IMPACT

We did short-term global CAM-chem simulations driven by MERRA reanalysis data at 0.9° x 1.25° horizontal resolution and FSTARMAM7 component set. Fig. 1 shows the global O3 concentration for May 2005. We will be doing a variety of additional modeling simulations and comparing the results. Previous studies have shown that trans-Pacific dust and aerosols contribute significantly to North American aerosol inflow [2], while export-related Chinese pollutants contributed 3–10% of annual mean sulfate concentration, 0.5–1.5% of ozone, and one or more day of noncompliance of ozone standards over many U.S. regions in 2006 [3]. We are in the process of investigating the long-range transport of pollutants from China to the United States.

WHY BLUE WATERS

The computational demand of the high-resolution climate model used in this project is extensive. We are using the fully coupled model of the Earth’s climate system with chemistry, which is computationally expensive. Blue Waters, with its petascale computational facility, large number of nodes, storage capability for the output from the high-resolution model simulation, is essential for our project. Blue Waters staff have been critical in figuring out the various issues arising with the long-term fully coupled climate chemistry runs with CESM. The staff have helped figure out and resolve various issues with the CESM1.2.2 models. Blue Waters has given us the computational resource, data management, and support staff to perform our research.
EXPLORING CONFINEMENT VS. ORIENTATION EFFECTS IN RIGID AND SEMI-FLEXIBLE POLYMERS USING A MASSIVELY PARALLEL FRAMEWORK

PI: David Ackerman
Collaborator: Balasubramanian
Collaborator: Iowa State University

EXECUTIVE SUMMARY

The overarching goal of this work is to develop and utilize a highly parallel computational framework to model equilibrium structures of semi-flexible polymers efficiently in complex, confined, non-periodic geometries. This will allow simulation of a wide range of technologically and biologically relevant systems such as DNA, RNA, and viral capsids; polymer chains with stiff backbones used in organic electronics (field effect transistors, photovoltaics, thermoelectrics); liquid crystals, where orientation and alignment create ordered phases; and rod-like polymers used in membranes and sensors.

To handle the range of applications, the simulation system is designed to be capable of scaling to large numbers of processors for large, complex systems of semi-flexible polymer systems. It is also designed to use high-throughput simulations for parameter sweeps of smaller or less complex systems. This work has developed tools designed to use high-throughput simulations for parameter sweeps of smaller or less complex systems.

RESEARCH CHALLENGE

An understanding of the morphology of polymers under different conditions aids in materials design and discovery. Self-consistent field theory (SCFT) [1] simulations are a tool for efficiently modeling polymer microstructures. While the method has proven very successful, most SCFT simulations are not designed for scaling to large systems with arbitrary geometries and boundary conditions. They are also not designed to handle polymer systems with semi-flexible chains. This puts limits on the applicability of those simulations. A generic, highly scalable SCFT simulation tool capable of modeling semi-flexible polymers greatly expands the types of polymer systems that SCFT studies can address. These tools will be of great interest to polymer scientists and materials researchers. It will also enable the ultimate goal of studying confinement effects on semi-flexible chains.

METHODS & CODES.

To address the desired goals of highly scalable simulations of systems having arbitrary geometries and boundaries with complex chain models, we utilized a finite element approach with several advanced techniques. Our implementation uses a custom finite element library built upon the Portable, Extensible Toolkit for Scientific Computation (PETSc) [2]. The SCFT method utilizes an evolving chain contour, analogous to a time-dependent problem. To provide enhanced scaling, we use a coupled space-time model in the contour solver. In contrast to typical time-dependent finite element models, we utilized a finite element volume mesh. At each nodal point within the volume mesh, a spherical surface mesh handles the orientation. The coupling of the two simulation spaces enables orientationally dependent systems to be simulated.

RESULTS & IMPACT

This work has had three major results: First, the development, testing, and deployment of a scalable finite element SCFT code. This has been utilized in a high-throughput study of polymer confinement and will be used in several future studies. Development of this code included design of an efficient process for the construction of polymer structure phase diagrams. Second, implementation and testing of a simultaneous space-time finite element method. Access to the Blue Waters supercomputer enabled testing of this system by providing access to numbers of CPU cores unavailable elsewhere. This ensured a full view of the capabilities and limitations of this method. This is invaluable for a method aimed at enhancing scaling. Third, development of a framework for modeling semi-flexible polymers using SCFT. This enables morphology studies of polymers with complex, orientation-dependent properties. The computational demands of polymer models have limited past work on these types of polymers. The resources of Blue Waters are well suited to the computational demands of this work.

WHY BLUE WATERS

The target systems for this work are polymer systems on a physically relevant scale. The finite element framework has been designed to model systems with large numbers of nodal points. At each finite element nodal point there are 180 or more degrees of freedom for which to solve. For even the smallest problem, this results in billions of unknowns evolving under complex physical processes. Solving for these values yields the polymer structure. The solving process requires significant, sustained computational resources. Solving for a structure in a reasonable time takes even more resources. Utilizing a highly scalable framework and thousands of nodes makes these structure determinations feasible. The processing resources are not available outside the Blue Waters system.
**EXECUTIVE SUMMARY**

Geopolymers are a class of inorganic polymeric, X-ray amorphous materials consisting of alumina, silica, and alkali metal oxides [1]. Geopolymer composites exhibit appealing properties such as high thermal stability [2], small carbon dioxide footprint [3], and high strength, which makes them suitable for many applications including alternative composites materials [4], passive cooling systems [5], low-level nuclear waste encapsulation [6], and potential novel biomaterials. The research goal is to understand the impact of nanoporosity on stiffness and strength via molecular dynamics and finite element modeling. Nanoporosity is governed by monovalent cation and water content. In contrast, microporosity results from the mixing procedure and from poor bonding between the geopolymer matrix and the filler phase. To increase the performance of geopolymer composites several strategies were found: reduce the microporosity, select strong bonding between the geopolymer matrix and the filler phase. The constitutive behavior of the individual constituents was prescribed based on independent experiments. Tensile tests with prescribed periodic boundary conditions were simulated to yield the effective response.

**RESULTS & IMPACT**

The strength and stiffness of amorphous geopolymers were correlated to their chemistry and density. The nanoporosity was found to be a function of the alkali and Si:Al ratio. At the mesoscale, the mechanical resistance of geopolymer composites to permanent deformation was elucidated as a function of the type, nature, size, and shape of the inclusions. Our theoretical and computational framework was validated on 31 different geopolymer-based systems, based on experiments carried out by various researchers over seven years. To our knowledge, it is the first time that a theoretical model has been proposed to upscale the constitutive behavior of geopolymers. In future efforts, we will model the evolution of the fracture toughness of pure geopolymer for various densities via molecular dynamics.

**WHY BLUE WATERS**

The Blue Waters platform was essential to carry out our molecular dynamics simulations as well as our finite element simulations. Standard desktop workstations do not have the required memory and computational power to resolve the complex systems involved. For molecular dynamics, it was crucial to be able to simulate large systems in a timely fashion. In the case of finite element simulations, the nonlinearity of the equations due to friction, contact, and plastic flow made it impossible to obtain results using a desktop workstation. In both cases, the Blue Waters supercomputer has enabled us to obtain novel results that bring new insights into the origins of strength and toughness in geopolymer composites.

**PUBLICATIONS AND DATA SETS**


HARNESSING PETASCALE COMPUTING TO EXPLAIN FUNDAMENTAL MECHANISMS DRIVING NANOPATTERNING OF MULTICOMPONENT SURFACES BY DIRECTED IRRADIATION SYNTHESIS

Allocation: Illinois/250 Knx
PI: Jean Paul Allain
Collaborators: Michael A. Lively, Brandon J. Holybee, Michal Toriyama, Scott A. Norris

1University of Illinois at Urbana-Champaign
2Southern Methodist University

EXECUTIVE SUMMARY

Ion beam-driven synthesis of quantum dots on III–V semiconductor surfaces is a promising approach toward a device-scale fabrication process for solar cells, quantum dot lasers, and other applications. However, the physical mechanisms that determine the pattern formation and properties such as quantum dot spacing remain poorly understood. To close this knowledge gap, massive-scale molecular dynamics simulations have been performed studying ion beam bombardment of gallium arsenide (GaAs) and gallium antimonide (GaSb) surfaces with 500 eV Ne+, Ar+, and Kr+ ions. Incident energetic ions are able to induce the formation of single-element clusters within the GaSb surface, with Sb clusters being far more prevalent than Ga clusters. These clusters may provide the basis for long-timescale mechanisms beyond the scope of molecular dynamics, such as radiation-enhanced diffusion, to create the compositional depth profile observed in experimental studies.

INTRODUCTION

The ability to synthesize unique nanostructures and nanopatterns at surfaces by ion beam irradiation holds great potential as a scalable technique to create novel nanomaterials in a single process step. In particular, the ability to fabricate quantum dots on III–V semiconductor surfaces, such as Gallium Arsenide (GaAs) or Gallium Antimonide (GaSb), using ion beam irradiation could be used to efficiently functionalize these materials for use in high-efficiency solar cells or quantum dot lasers. However, since the first discovery of these nanodots [1], theoretical models have been unable to provide a complete, predictive description of the nanopattern formation on these surfaces. Consequently, the dependence of nanopattern formation on III–V semiconductor surfaces on experimental parameters such as the incident ion mass or energy remains poorly understood.

METHODS & RESULTS

Molecular dynamics (MD) simulations of ion bombardment are carried out with the LAMMPS package [5] for 500 eV Ne+, Ar+, and Kr+ ions incident on an initially pristine GaSb surface. The simulation cells were 25 × 25 nm2, providing sufficient space for lateral compositional variations to evolve on the same scale observed in previous simulations. For each ion species, the GaSb surface was irradiated to a fluence of 7.5 × 1015 ions/cm2, an experimentally relevant fluence (the sum of the energies of the particles per unit area contained in the particles with which a material is irradiated) on the same order of magnitude as the threshold fluences for pattern formation that have been observed in experiments.

Snapshots of the resulting surfaces after ion irradiation are shown in Fig. 1 for the 500 eV Kr+ case. Three major observations can be made from these snapshots. First, despite irradiation to a fairly large fluence, no compositional depth profile was observed. This indicates that additional long-temporal scale mechanisms are necessary to accurately model the compositional evolution of the surface, such as radiation-enhanced diffusion or stress-driven flow. Second, while a depth profile does not emerge, the formation of smaller clusters of Ga or Sb can still be observed due to the prompt ion effects. This indicates that even at short timescales the energetic ions still induce compositional changes in the surface, which may provide “seed” structures to grow into larger-scale lateral compositional gradients. Finally, Sb atoms show a much stronger clustering tendency than Ga atoms, which may be indicative of the higher mobility of Ga compared to Sb seen in experimental studies [6].

Interestingly, different ion species were shown to have differing effects on the surface compositional evolution, as shown in Fig. 2. To address this knowledge gap, atom-level computational modeling is necessary to observe and characterize the ion-induced changes in the surfaces as they occur, providing detailed insight into the ion-surface interactions at length and time scales far smaller than what can be experimentally observed. At the same time, any simulation efforts must also be able to connect the small-scale atomistic mechanisms to the evolution of the larger surface at experimentally observable length and time scales. Determining both the large and small scales of ion-driven nanopatterning can only be done with high-performance computational power on a massive scale, well beyond that provided by conventional computing platforms.

In previous work conducted on Blue Waters [2], massive-scale molecular dynamics simulations were conducted to connect experimental observations of the compositional depth profile [3,4] to pattern-forming surface instabilities. The results showed, for the first time, that an ion-induced compositional depth profile led to lateral phase separation, providing a pathway for the pattern-forming instability to emerge. The present simulations seek to elucidate the ion-induced mechanisms causing the formation of the compositional depth profile that drives these changes.

PUBLICATIONS AND DATA SETS


Figure 1: Snapshots of the GaSb surface after irradiation by 500 eV Kr+ ions to a fluence of 7.5 × 1015 ions/cm2. (a) snapshot of the entire surface after ion bombardment, showing the lack of any notable compositional depth profile. (b) snapshot showing only Ga atoms in clusters. (c) snapshot showing only Sb atoms in clusters. Sb atoms show a significantly greater tendency to form clusters within the irradiated GaSb surface.

Figure 2: Fraction of Sb atoms which are members of clusters as a function of increasing fluence for each ion species. While all three ion species initially induce similar clustering behavior, for fluences greater than 2.4 × 1015 ions/cm2 Kr+ irradiation causes a decrease in clustering compared to the lighter two ion species.
EXECUTIVE SUMMARY

Distinguishing the different subclasses of immunoglobulin G (IgG) antibodies in blood serum can enable breakthrough advances in mapping the immune system and the health status of the human body. In this study, using petascale-based molecular simulations (containing up to ~1,000,000 atoms) and a total aggregate simulation time of 2.7 microseconds (µs), we demonstrate that an atomically thin graphene nanopore is capable of sensing and discriminating among different subclasses of IgG antibodies despite their minor and subtle variations in atomic structure. Using machine learning, we featurized and clustered the ionic current and the dwell times data obtained from the device during multiple antibody translocation events. In addition, the histogram of ionic current for each segment of IgG can provide high-resolution spatial detection of antibody segments. Parallel nanofluidic studies during IgG translocation reveal distinct water flux rates for IgG subclasses facilitating an additional recognition mechanism.

DESCRIPTION OF RESULTS

DNA sequencing using nanopore technology has evolved significantly during the last few years. Oxford Nanopore Technologies Ltd. currently is fabricating a USB-stick-sized device that can sequence the DNA in a few hours. In recent years, both biological and synthetic nanopores were used for “label free,” high-resolution DNA sequencing. In addition to DNA sequencing, detection of antibody proteins can lead to advances in improving human health. The challenges posed to biological molecule detection using nanopore technology are the low signal-to-noise ratio, pore degradation due to multiple uses, the identification of single bases in real time, and the high speed of translocation [1,2]. Engineering the translocation of molecules through biological/synthetic nanopores has been defined as one of the challenging problems of biotechnology. By using extensive molecular dynamics (MD) simulations, this study shows that an atomically thin graphene nanopore is capable of sensing and discriminating between different subclasses of IgG antibodies.

Protein detection via a graphene nanopore is accomplished using ionic current, dwell time, and water flux calculations. A total aggregate of 2.7 µs of simulation time has been carried out for systems containing up to ~1 million atoms. All parts of the antibody (Fig. 1) are distinguishable by ionic current measures. More specifically, the Fab, Fc, and hinge regions exhibit a unique current level when translocating through the pore. Using k-means clustering, the ionic current–dwell time and water flux–ionic current feature plots lead to clusters with distinguishable centroids. We also compared the performance of the single-layer graphene nanopore with that of a solid-state nanopore (Si3N4). IgG subclasses are not distinguishable when using thick nanopores of Si3N4 because some of the atomic details cannot be captured. In conclusion, we have shown that ionic current, dwell time, and water flux can detect different antibodies with high precision.

WHY BLUE WATERS

We performed extensive MD simulations that involved up to 1,000,000 atoms. These expensive computations would have not been possible to perform without a petascale supercomputer. Also, the MD package (NAMD) we used scales almost linearly with the number of cores up to 1,000 in our test on Blue Waters.

PUBLICATIONS AND DATA SETS

This project focuses on high-performance calculations for materials and devices of high current interest and on development of petascale methods for such simulations. This past year, we concentrated on three challenges. First, we investigated two paradigmatic sensor configurations for detection of biologically important molecules through ab initio calculations: a noncovalently functionalized nanotube for glucose detection, and a covalently functionalized nanotube for ethylene detection. Second, we analyzed high-performance dielectric materials that store and release energy electrostatically through depolarization and depolarization. We showed that blending two dipolar strongly polar molecules, e.g., a poly(ethylene oxide) (PEEPO) and an aromatic polypeptide (APPTU), K = 4.4), the resulting mixture exhibits a very high dielectric constant, K = 7.5, while maintaining low dielectric loss (c 1%). Third, in the pursuit of atomically precise and bottom-up fabrication of graphene-based electronics, we determined and understood the atomic mechanism responsible for controlled conversion of a polymer to a nanoribbon, stimulated by hole injection through an scanning tunneling microscope (STM) tip.

**RESEARCH CHALLENGE**

Biomolecular detection is a rapidly growing field in biochemical and biomedical sciences. It is widely recognized as one of the key technologies for environmental monitoring and disease diagnosis. Functionalized carbon nanotube (CNT)-based nanosensors have shown great potential for detection, due to their sensitivity and selectivity. We computationally investigated the sensing mechanisms and capabilities of two paradigmatic sensor configurations based on single-walled carbon nanotubes (SWCNT): the covalently functionalized sensor for ethylene detection and the noncovalently functionalized sensor for glucose detection.

Dielectric materials store energy electrostatically through various polarization mechanisms and release it by depolarization. Dielectric capacitors are unparalleled in flexibility, adaptability, and efficiency for electrical energy storage, filtering, and power conditioning. We have discovered a practical and widely applicable mechanism for enhancing the dielectric constant through nanostructure engineering of dipolar polymers.

In the pursuit of atomically precise and bottom-up fabrication of graphene-based electronics, impressive advances have been made with the synthesis of graphene nanoribbons (GNR) by polymerizing monomers on different catalytic metal surfaces. We investigated the formation of GNRs from quasi-free-standing polymers assisted by hole injections from an STM tip. While a catalytic cyclo-hydrogenation occurs in a cooperative domino-like conversion process during the thermal annealing, the hole-injection-assisted reaction happens at controlled molecular sites selected by the STM tip. We have established how the bottom-up synthesis of a graphene nanoribbon can be controlled by charge injections from an STM tip. From our experiments and first-principles calculations, it was found that the hole injections from an STM tip can trigger a cooperative domino-like cyclo-hydrogenation even when the polymers are quasi-free-standing with suppressed substrate effect. The hole injections greatly reduce the energy barrier in the key step of the carbon-to-carbon bond formation. The hydrogen atoms migrate to the edge and dissociate into the vacuum as H₂ molecules. As the STM tip treatment can be performed at selective molecular sites without involving a traditional catalytic effect from the metal substrate, the results point to a new way for bottom-up and controllable synthesis of free-standing GNRs and heterojunctions, which is critical for practical GNR-based nanodevices.

**RESULTS & IMPACT**

The sensing mechanism of glucose detection that we established can be generalized to detect other carbohydrate molecules, which are also known as saccharides. Carbohydrates are involved in a wide range of biological and pathological processes, such as cancer metastasis, cell signaling, protein function regulation, and cellular communication. Detection of specific carbohydrates is thus essential for investigating their roles in numerous natural processes, as well as for obtaining insights into the mechanisms involved and diseases provoked. The glucose sensor configuration and mechanism we described could be very useful in the design of other carbohydrate sensors. Moreover, the glucose sensing mechanism and the computational framework we developed can be applied for detection of the prostate cancer biomarker, which is known as osteopontin (OPN). Experimentalists have demonstrated that nanotube devices show detectable sensitivity to OPN when covalently functionalized with a receptor called scFv. One possible explanation is that the charged sites on the scFv's surface are neutralized by opposite charges associated with the bound OPN, leading to a change of electrostatic potential at the nanotube surface. However, the precise mechanism for the observed sensing response remains to be determined. The computational approach established here could provide a quantitative understanding of the sensing mechanism.

It is a great challenge in dielectric polymers to achieve a high dielectric constant while maintaining low dielectric loss and high operating temperatures. We show that by blending two dipolar polymers in glassy state, i.e., poly(ethylene oxide) (PEEPO, K = 4.7) and an aromatic polypeptide (APPTU, K = 4.4) to form a nanosystem; the resulting blend exhibits a very high dielectric constant, K = 7.5, while maintaining low dielectric loss (c 1%). The experimental and computer simulation results demonstrate that blending these dissimilar dipolar polymers causes a slight increase in the interchain spacing of the blend in its glass state, thus reducing the barriers for reorientation of dipoles in the polymer chains along the applied electric field and generating a much higher dielectric response than the neat polymers.

We have established how the bottom-up synthesis of a graphene nanoribbon can be controlled by charge injections from an STM tip. From our experiments and first-principles calculations, it was found that the hole injections from an STM tip can trigger a cooperative domino-like cyclo-hydrogenation even when the polymers are quasi-free-standing with suppressed substrate effect. The hole injections greatly reduce the energy barrier in the key step of the carbon-to-carbon bond formation. The hydrogen atoms migrate to the edge and dissociate into the vacuum as H₂ molecules. As the STM tip treatment can be performed at selective molecular sites without involving a traditional catalytic effect from the metal substrate, the results point to a new way for bottom-up and controllable synthesis of free-standing GNRs and heterojunctions, which is critical for practical GNR-based nanodevices.

**WHY BLUE WATERS**

The applications described above require a very large parallel supercomputer with a high-speed interconnect among the nodes (due to the frequent exchange of substantial amounts of data among nodes). Each project required many runs to explore the various scientific issues, with a substantial amount of analysis...
REDUCING JET AIRCRAFT NOISE

Allocation: Blue Waters Professor/100 Kgh
PL: Daniel Bodony
Collaborator: Mohammad Mehrabadi

1University of Illinois at Urbana-Champaign

EXECUTIVE SUMMARY

Developing quieter, more fuel-efficient jet aircraft has been a goal since the 1950s. The movement of engine designs from the original turbojet to today’s high-to-ultra-high-bypass-ratio turbosfans has sustained noise and fuel-use reduction goals. However, future growth in aircraft utilization and more stringent requirements on acceptable noise suggest further demand on noise reduction. This requires detailed understanding of the turbulent flow in the fan stage of the engine that gives rise to the dominant broadband component of the fan noise, which cannot be provided by current reduced order models. Therefore, it is essential to utilize high-fidelity numerical tools to predict the turbulence and its radiated noise.

RESEARCH CHALLENGE

The dominant noise source of jet engines has changed as engines have evolved. Large bypass-ratio fans (Fig. 1, upper-left inset image) produce noise that predominantly originates from the fan itself, rather than from the exhaust streams from the core and fan. Fan noise is mostly due to the interaction of small unsteady flow perturbations with the fan (rotor) and stationary guide vane (stator) blade rows. Depending on the nature of these sources, they are categorized as tonal noise and broadband noise.

Broadband noise contains a wide range of frequencies and is associated with the interaction of a turbulent flow with a solid boundary (Fig. 1, main image). The prediction of the broadband component of fan noise (Fig. 1, lower-right inset image) is more complicated than tonal noise predictions since it arises from an inhomogeneous and anisotropic turbulence interacting with fan blades. Although the methods developed to predict rotor-stator interaction tonal noise can, in practice, also be applied to the broadband component prediction, representing a collection of random incident disturbances in the time domain or a wide range of frequencies in the frequency domain—the ensemble of which represents a turbulent flow—requires a highly resolved computational domain with a significant computational cost.

METHODS & CODES

The turbulence and its radiated sound are governed by the compressible Navier-Stokes equations that are solved using a high-order, finite difference code written by the PI, capable of describing complex geometries using multiple, overset meshes. From fan airfoil geometry, including small roughness elements near the leading edge (Fig. 1, main image), a computational grid is designed which sufficiently resolves the flow around the blade as well as the radiated field far away from it (Fig. 1, lower-right inset image). The resulting mesh contains 2.5 billion degrees of freedom, which are saved in HDF5 format on the file system using a new, scalable, collective MPI-I/O-based infrastructure (Fig. 2).

RESULTS & IMPACT

The direct numerical simulation of sound generated by an idealized blade shown in Fig. 1 requires 2.5 billion degrees of freedom and generates a database on the order of 1 terabyte. While useful for generating reduced-order models for broadband sound generation from isolated blades, the simulation domain shown in the main image of Fig. 1 is only a fraction of that needed to predict the noise from the entire fan shown in the upper-left inset image of Fig. 1. Thus, the main results of the simulation are being used to develop a wall-modeled large-eddy simulation (WMLES) approach to broadband noise prediction wherein the inner boundary layer dynamics (below the log layer) are modeled using a less expensive model while the outer boundary layer dynamics and external flow field are modeled using traditional LES techniques. Determining the accuracy of the WMLES simulation for on-blade and acoustic field predictions is the primary impact of the simulations. If successful, using WMLES for broadband noise prediction enables the aeroacoustic community to predict and, ultimately, reduce the environmental impact from most aircraft noise sources.

WHY BLUE WATERS

Blue Waters is critical to the research because of the system’s scale, the performance of its I/O system, and the capability its software teams have in improving code performance, including I/O performance. These runs would not be feasible on any XSEDE computer.
PROPERTIES OF DENSE HYDROGEN

The phase diagram of high-pressure hydrogen is of great interest for fundamental research, planetary physics, and energy applications. A first-order phase transition in the fluid phase between a molecular insulating fluid and a monoatomic metallic fluid has been predicted. The existence and precise location of the transition line is relevant for planetary models. Recent experiments reported contrasting results about the location of the transition. Theoretical results based on density functional theory are also very scattered. We performed highly accurate coupled electron-ion Monte Carlo calculations of this transition, finding results that lie between the two experimental results. We have also explored the transition between molecular and atomic hydrogen at lower temperatures, in the solid phase.

METHODS & CODES

Recent experiments reported contrasting results about the location of the transition, different by a factor of two in pressure. Theoretical results based on density functional theory are also very scattered and hence not predictive. These findings motivated us to repeat our earlier calculations on Blue Waters, to better control the convergence and utilize recent improvements in methodology. In addition, recent experiments have probed the solid phase leading up to the atomic transition and the melting temperature.

RESULTS & IMPACT

The properties of hydrogen under extreme conditions of pressure and temperature are of great interest for fundamental research, planetary physics, and energy applications. Hydrogen accounts for much of the visible mass in the universe. The properties of hydrogen and helium are needed to understand the giant planets, Jupiter and Saturn, but experiments under the relevant conditions are challenging. Even though hydrogen is the first element in the periodic table, calculating its properties is not simple since both the electronic and protonic motions are quantum and correlated. It has long been an open question how hydrogen makes a transition from a molecular insulating state to an atomic metallic state as pressure and temperature are increased. A first-order phase transition in the fluid phase between a molecular insulating fluid and a monoatomic metallic fluid has been predicted.

EXECUTIVE SUMMARY

The phase diagram of high-pressure hydrogen is of great interest for fundamental research, planetary physics, and energy applications. A first-order phase transition in the fluid phase between a molecular insulating fluid and a monoatomic metallic fluid has been predicted. The existence and precise location of the transition line is relevant for planetary models. Recent experiments reported contrasting results about the location of the transition. Theoretical results based on density functional theory are also very scattered. We performed highly accurate coupled electron-ion Monte Carlo calculations of this transition, finding results that lie between the two experimental results. We have also explored the transition between molecular and atomic hydrogen at lower temperatures, in the solid phase.

RESEARCH CHALLENGE

The properties of hydrogen under extreme conditions of pressure and temperature are of great interest for fundamental research, planetary physics, and energy applications. Hydrogen accounts for much of the visible mass in the universe. The properties of hydrogen and helium are needed to understand the giant planets, Jupiter and Saturn, but experiments under the relevant conditions are challenging. Even though hydrogen is the first element in the periodic table, calculating its properties is not simple since both the electronic and protonic motions are quantum and correlated. It has long been an open question how hydrogen makes a transition from a molecular insulating state to an atomic metallic state as pressure and temperature are increased. A first-order phase transition in the fluid phase between a molecular insulating fluid and a monoatomic metallic fluid has been predicted.

METHODS & CODES

Over the past decade, we have developed new quantum Monte Carlo simulation methods to treat quantum systems at low temperature. The quantum Monte Carlo method we use (Coupled Electron Ion Monte Carlo) works with the full interaction between the electrons and protons and treats both particles fully quantum mechanically. In contrast to density functional calculations, all effects of electronic correlation are explicitly included. This is particularly important in hydrogen, because of possible self-interaction effects, difficulty in treating the hydrogen bond breaking, and the large van der Waals interactions. We model hydrogen with about 100 electrons and protons in a periodic cell. Special methods are used to extrapolate accurately to the thermodynamic limit. With our method, we simulated hydrogen for temperatures in the range of 200⁰K up to 5,000⁰K, and at relevant pressures, 100GPa to 500GPa.

RESULTS & IMPACT

For temperatures below 2,000⁰K we observe a first-order transition between an insulating molecular liquid and a denser metallic atomic liquid. Our predicted transition pressures are intermediate between the two experimental observations. In recent unpublished work, we performed a detailed study of solid molecular hydrogen at temperatures between 200⁰K and 414⁰K and pressures between 180GPa and 520GPa, to understand the detailed crystal structures, the melting temperature of the solid and the transition into the metallic/atomic phase. There has been a recent controversial report of the experimental observation of this transition. Shown in the figure are snapshots of a layer of one of the crystal structures (Cmca12) as it evolves from 180GPa to 420GPa.

Our calculations are needed to validate both our computational method and to resolve the different experimental measurements. It is essential for progress in the high-pressure community to resolve the difference between the experiments and computation. After validation, the method can be used with more confidence in modeling the wide variety of astrophysical objects being observed, composed largely of hydrogen and helium under extreme conditions. As a byproduct of our simulation, we are also judging the accuracy of various density functionals and force-fields for the simulations of these systems.
SCALABLE NANOPATTERNING OF GRAPHENE BY HYDROGEN-PLASMA ETCHING

Alidential: University of Illinois/280 Kith
PI: Huck Beng Chew

EXECUTIVE SUMMARY

Scalable and precise nanopatterning of graphene is an essential step for graphene-based device fabrication. Hydrogen-plasma reactions have been shown to narrow graphene from the edges, or to selectively produce circular and hexagonal holes in the basal plane of graphene, but the underlying plasma-graphene chemistry is unknown. The petascale Blue Waters supercomputing resources have enabled us to quantify the mechanisms of hydrogen-plasma etching of graphene supported on SiO₂ substrate across the range of plasma ion energies. Specifically, our molecular dynamics simulation results, based on a reactive force-field potential, have uncovered distinct etching mechanisms, operative within narrow ion energy windows, which explain the differing plasma-graphene reactions observed experimentally. These simulation results have provided rich insights into the complex plasma-graphene chemistry, opening up a means for controlled patterning of graphene.

RESEARCH CHALLENGE

The technique of plasma surface nanopatterning of materials has played an important role in the microfabrication over the past several decades of semiconductors of ever-reducing dimensions. Silicon, which has so far been the primary material for manufacturing of transistors, has reached its technological limitation with the recent 7-nm scale transistors. Use of novel materials such as graphene has been proposed in order to achieve the next milestone of the 5-nm ITRS node (International Technology Roadmap for Semiconductors). However, graphene has been patterned to achieve its full potential in the domain of electronic applications. Several methods such aslausination, boron nitride doping, irradiation by heavy ions, electron beam lithography and etching by hydrogen plasma have been used in the past for this purpose. Among these techniques, hydrogen plasma etching has shown the most promise in terms of scalability and cost effectiveness. While there exists ample experimental evidence for the patterning of graphene by hydrogen-plasma treatment, the reported etching reactions and the resulting graphene nanostructures have been vastly different. These reactions range from selective edge etching with no damage to its basal plane, to combined basal plane and edge etching of graphene resulting in isotropic and anisotropic hole growth in the basal plane of graphene. To date, controlling the patterned graphene nanostructures by hydrogen-plasma treatment has not been achieved due to a lack of fundamental understanding of the complex hydrogen plasma-graphene chemistry. Furthermore, the complete parameter space of substrate temperature, ion energy, and incident flux has not been systematically studied due to the cost limitations of plasma experiments.

METHODS AND CODES

We perform length-scale bridging by delineating the contributions of the edge and basal plane etching using ReaxFF-based molecular dynamics (MD) and linking these processes together via a mechanistic model. All our simulations were performed on the C3s-based open source LAMMPS code.

RESULTS & IMPACT

We quantify the extent of basal plane damage for each ion energy by defining a damage parameter Dₐ as the fraction of broken C–C bonds in the graphene sheet. The basal plane of graphene remains nearly undamaged at ion energies of 1 eV and 25 eV, but displays a nonmonotonous relationship with ion energies in between. Peak etching is observed at ion energy of 10 eV, as shown in Fig. 1a. The basal-plane etching initiates with the chemisorption of the H atoms on two neighboring C atoms and proceeds with the further hydrogenation of the dangling C–C bonds, as shown in Fig. 1b. To quantify the edge-etching of graphene, we define an edge-etching rate Dₑ and measure its variation with ion energy for both zigzag and armchair orientations, as shown in Fig. 1c. Our results demonstrate distinct ion energy regimes for isotropic versus anisotropic etching. Between 1 and 5 eV, the average armchair and zigzag etching rates are similar with overlapping error bars. At ion energies of between 7 and 30 eV, however, armchair etching consistently proceeds faster than zigzag etching by ~15%, suggesting that the patterned holes should indeed be hexagonal with zigzag-oriented edge structures, in agreement with experimental observations. Fig. 1d shows the mechanisms of etching at the zigzag and armchair edge. H-induced etching as well as direct impact damage plays a role in zigzag edge, while only the latter is active in the more stable armchair edge. These distinctive etching mechanisms, which are operative within narrow ion energy regimes, fully explain the differing plasma-graphene reactions observed experimentally.

WHY BLUE WATERS

The Blue Waters computational capacities were necessary for several reasons. First, the complex chemistry and plasma surface interactions involved in the hydrogen etching of graphene require the use of fully reactive MD potential, allowing for potential reactions between the Si, O, C, and H species at each MD time step. Second, the impact dynamics of impinging H atoms on graphene requires the use of a small time step (0.15 fs), which further increases computational cost. Third, studying the edges of the multilayer graphene presents a wide range of possible configurations, as the edges can be partially or completely covered by a graphene layer. Finally, because of the random process of H deposition, a large number of simulation runs are required to obtain statistically significant findings.

PUBLICATIONS AND DATA SETS


Figure 1: a) Steady state basal plane etching rate versus ion energy for monolayer graphene; b) Mechanism of etching showing the three possible configurations of damage nucleation and progression; c) Etching rates of the zigzag and armchair configurations versus ion energy; and d) Mechanism of etching for zigzag and armchair edges.
hPIC: A SCALABLE ELECTROSTATIC PARTICLE-IN-CELL FOR PLASMA–MATERIAL INTERACTIONS

Allocation: Exploratory/50 Knx
PI: Davide Curreli1
Co-PI: Rinat Khaziev1
Collaborators: Cameron Dart1, Alyssa Hayes1
1University of Illinois at Urbana-Champaign

EXECUTIVE SUMMARY

The hPIC code is a kinetic–kinetic electrostatic particle-in-cell (PIC) application targeted at large-scale simulations of plasma–material interactions (PMI). hPIC models the plasma sheath and presheath in strongly magnetized conditions, as normally encountered at the wall of magnetic fusion reactors. The Poisson solver is based on the PETSc conjugate gradient with BoomerAMG algebraic multigrid preconditioners. Scaling tests on the Blue Waters supercomputer have demonstrated excellent weak scaling and considerable strong scaling.

RESEARCH CHALLENGE

In the edge region of magnetically confined plasmas, plasma–material interactions can compromise the thermos–mechanical integrity of the wall, currently limiting the development of nuclear fusion reactors. When exposed to a fusion plasma environment, plasma-facing materials exhibit evidence of surface morphology modifications, nanostructuring, and neutron damage. The boundary layer between the edge plasma and the material surface is called the plasma sheath, a region where the plasma ions accelerate up to supersonic conditions in order to maintain the quasi-neutrality of the plasma. In the strong magnetic field of a fusion machine (of the order of several Tesla), the plasma sheath is a complex, three-dimensional multi-layer structure, where substantial electrostatic forces accelerate the ions toward the wall, causing damage. The hPIC code is a large-scale particle-in-cell under development at the University of Illinois and is able to perform detailed analyses of the plasma sheath in the strongly-magnetized conditions encountered in a fusion reactor. We report progress on the scaling tests (weak and strong scaling) and code optimization performed on Blue Waters. All scaling plots have been reported as a function of the number of Blue Waters’ XE6 nodes; one Blue Waters node corresponds to 32 cores.

METHODS & CODES

As an electrostatic particle-in-cell, hPIC requires the solution of a large-scale Poisson problem at each time step. A well-known problem of elliptic solvers is the intrinsic nonlocality of the differential operator, hindering an efficient and scalable implementation on a large-scale computer. hPIC uses a finite-difference discretization of the Laplace operator solved with conjugate gradient and an algebraic multi-grid (BoomerAMG) as a preconditioner. The PETSc library was used as an interface for the parallel implementation. The PETSc numerical library manages the field functions of the code. The interface between hPIC and PETSc also ingests the charge density inputs required by the Poisson solver and returns electric field and electric potential. hPIC and PETSc also ingests the charge density inputs required by the Poisson solver and returns electric field and electric potential.

RESULTS & IMPACT

The largest weak scaling test performed was a Poisson problem on 2,048 Blue Waters nodes (65,356 cores). A 500–500 grid ran on each core, resulting in 16.304 billion degrees of freedom. Fig. 1 shows the efficiency (w.r.t. single node) of the full Poisson code, including both particle and field functions, for different numbers of particles-per-cell (ppc). A weak scaling efficiency greater than 78% and between 83% and 88% for all conditions relevant for practical applications has been attained.

The tests have shown that as a rule of thumb for good performance, the number of particles per cell should be kept larger than 100, with preference to 500–600 particles-per-cell. With a small local grid of 50x50 on each MPI process, hPIC is capable of achieving up to 3,000 particles per cell, enough to achieve a satisfactory level of statistical noise. Currently, hPIC uses double precision for all physical quantities on the XE6 nodes of Blue Waters containing 64 gigabytes of random access memory per node. Thus, each node stores approximately 250 million particles. Similarly, strong scaling tests were performed on 4 to 8,192 nodes. The problem was a Poisson equation with a spatial grid of N×N points and a corresponding Laplacian matrix of size (N×N)×(N×N). At the largest scale of 8,192 nodes, the time to solution for the two allocatable cases of NN=16,000 and NN=32,000 was 1.6 seconds. The greatest problem of NN=100,000 was solved on 256, 512, and 2,048 Blue Water nodes. This solution required PETSc’s index to switch from a 4 byte signed integer to an 8 byte signed long long. The problem of size NN=100,000 has a grid size of NN×NN=1010 nodes, and a corresponding Laplacian matrix of NN2×NN2=1020 degrees of freedom with matrix indices running from 0 to 1020–1, which can be allocated only by using 64-bit unsigned long integers. Such a problem was solved in 13.4, 6.8, and 2.1 seconds respectively on 256, 512, and 2,048 Blue Waters nodes.

WHY BLUE WATERS

Access to Blue Waters has accelerated the development of hPIC, allowing for testing and optimization of a number of massively parallel features of the code. The code is now at the level of meeting the large computational demands of fully kinetic simulations of near-wall plasmas encountered at the edge of fusion reactors, simulations feasible only on a machine the size of Blue Waters. Resolving the plasma–material interface at scales from nano to millimeters and from picoseconds to the time scales of plasma–sheath equilibrium requires computing at the petascale and beyond. Blue Waters offers a unique opportunity to push this research forward.

Figure 1: Weak scaling efficiency of the electrostatic version of the full cycle of hPIC for a hydrogen plasma with ion and electron temperature equal to 1eV and electron density equal to 1017 m–3; each rank stores 100×100 grid points, with a number of particles per cell (ppc) as indicated in the figure.
EXECUTIVE SUMMARY

The Blue Waters system enables us to achieve large-scale modeling of light- and medium-mass nuclei, including short-lived nuclei not yet accessible to experiment but key to understanding astrophysical processes, which are the focus of current and next-generation rare isotope experimental facilities. The scale of computational challenges inherent in the modeling of such intricate quantum many-body systems makes the Blue Waters resources essential for addressing long-lasting questions of importance to nuclear theory and experiment, as well as to astrophysics. Breakthrough theoretical advances [1,2] coupled with the Blue Waters cutting-edge computational power have opened a new regime for the intermediate-mass nuclei from Fluorine to Calcium isotopes, for first investigations with the Blue Waters system, allow us to reach medium-mass nuclei that are inaccessible experimentally and to other computational methods [2,5].

SA-NCSM investigations of the lowest-lying eigenvectors that correspond to the experimental ground state of deformed and, in the future, heavy nuclei of interest to nuclear science. While the above-mentioned applications focus on the study of single nuclides, the Blue Waters system enables us to reach medium-mass nuclei that are inaccessible experimentally and to other ab initio methods [1,2].

RESULTS & IMPACT

The nucleon of interest represent a considerable challenge requiring computational power of nearly the entire Blue Waters (BW) machine and its system memory. Two graduate students have carried forward these studies and had the unique opportunity to work with supercomputers and massively parallel programming environments. The following list describes the results and their impact:

We have provided the first ab initio description of the open-shell $^4$He, $^6$He, $^8$Be, and $^{14}$N nuclei [2]. Such nuclei in the intermediate-mass region are key to further understanding the production of heavy elements, and especially X-ray burst nucleosynthesis. Of special interest are short-lived nuclei (such as $^{12}$C) that are difficult or impossible to study experimentally (see Fig. 2).

We have studied emergent phenomena from first principle in Mg isotopes and their mirror nuclei ($^{24}$Mg and $^{26}$F) and $^{22}$Ne, with work in progress on $^{23}$Mg and $^{24}$Mg. While enhanced deformation and cluster substructures are difficult to describe from first principles, the BW system has allowed us to achieve ab initio descriptions using chiral interneutron interactions [5]. This is important for providing accurate predictions for deformed and, in the future, heavy nuclei of interest to understanding the r-process nucleosynthesis.


Large investments have been made in new generations of radioactive beam facilities to enable important discoveries in nuclear science. While the above-mentioned applications focus on specific important questions, the concurrent new developments and dramatic improvements of the LSU3shell computer code, carried forward as part of the BW PAID program, may have wider impact, as multi-physics simulations in the areas of nuclear energy and national security have similar needs.

WHY BLUE WATERS

The ab initio nuclear structure studies represent an extremely computationally intensive endeavor. To illustrate the level of complexity, applications to medium-mass nuclei require in excess of hundreds of exabytes of memory to store the Hamiltonian matrix. The SA-NCSM drastically reduces the size of the problem and the associated memory requirement down to hundreds of terabytes and petabytes, but this comes at the cost of a major increase in computing intensity. As a result, SA-NCSM investigations of the intermediate-mass region are beyond the scale of available academic HPC systems. Currently, only the BW system provides resources required for the ab initio SA-NCSM studies of medium-mass isotopes. In order to capitalize on this opportunity, we drew from the experience and expertise of the Blue Waters staff and managed to improve scalability of our code. As a result, our largest production runs efficiently utilized 717,600 concurrent threads running on 22,425 Cray XE6 compute nodes to solve the nuclear eigenvalue problem with Hamiltonian matrices that occupy up to 400 TB of memory. Clearly, the BW system represents a unique computational platform that already plays a crucial role in advancing ab initio nuclear theory.

PUBLICATIONS AND DATA SETS


Figure 2: Effect on the abundance pattern from X-ray burst (XRB) nucleosynthesis (see Fig. 1).
DNS OF PRESSURE FLUCTUATIONS INDUCED BY SUPersonic TURBULENT BOUNDARY LAYERS

Allocation: NSF PRAC/L-100 Koh
PI: Linn Diez
Collaborator: Moshan Choudhari

Executive Summary

Pressure fluctuations are an important ingredient in wall-bounded turbulence as they are strongly correlated with turbulent vorticity dynamics and noise generation. Most existing analyses of boundary-layer-induced pressure fluctuations are based on the Poisson equation in the context of incompressible boundary layers. The pressure fluctuations induced by a supersonic turbulent boundary layer are, however, governed by the wave equation and are fundamentally more complicated than the low-speed counterpart. The objective of the research is to conduct direct numerical simulations (DNS) to advance fundamental understanding of the generic statistical and spectral features of boundary-layer-induced pressure fluctuations, including the freestream acoustic radiation at supersonic speeds and their dependence on boundary-layer parameters such as the Reynolds number. Current work with Blue Waters includes the analysis of boundary-layer data at modest Reynolds numbers, and conducting new simulations in Reynolds-number regimes difficult to reach without the large allocation provided on Blue Waters.

Research Challenge

Understanding the physics of the pressure fluctuations induced by turbulent boundary layers is of major theoretical and practical importance. From a practical point of view, the fluctuating pressure on aerodynamic surfaces of flight vehicles plays an important role in vibrational loading and often leads to damaging effects such as fatigue and flutter. The freestream pressure fluctuations radiated from the turbulent boundary layer on the nozzle wall are responsible for the genesis of freestream acoustic noise in supersonic wind tunnels. Therefore, the characterization of turbulent boundary-layer-induced pressure fluctuations and the relative importance of the two modes in the near duct is of fundamental importance to understanding the turbulent vorticity dynamics and to modeling the pressure-strain terms in the Reynolds stress closure.

One unique feature of boundary-layer-induced pressure fluctuations at supersonic speeds is that acoustic mode fluctuations emerge in the form of eddy Mach waves [1]. The pressure fluctuations thus include contributions from both vortical and acoustic modes. The characteristics of the acoustic pressure fluctuations and the relative importance of the two modes in different regions of the boundary layer are largely unknown. The current work aims to use the cutting-edge computational power of Blue Waters to provide the basis for an in-depth understanding of the global pressure field induced by supersonic turbulent boundary layers across a wide range of boundary-layer parameters. Such an understanding will advance the state-of-the-art knowledge of wall-bounded turbulence and boundary-layer-induced noise.

Methods & Codes

DNS are conducted using HyperWENO, an in-house high-order finite-difference solver that solves the compressible Navier-Stokes equations in generalized curvilinear coordinates describing the evolution of the density, momentum, and total energy of the flow. An optimized, high-order (up to 7th order) finite-difference WENO (weighted essentially non-oscillatory) scheme [2,3] is used to compute the convective flux. The WENO scheme combines a high order of accuracy with relatively low dissipation, making it suitable for simulations of compressible turbulent flows. A 4th-order central difference scheme is used for the viscous flux terms, and a 3rd-order low-storage Runge-Kutta scheme [4] is employed for time integration, which significantly reduces the memory requirement and is well suited for time-accurate simulations such as DNS. The turbulent inflow can be generated using either a recycling/rescaling method [5] or a digital filtering method [6]. On the wall, no-slip conditions are applied for the three velocity components, and an isotropic condition is used. The pressure at the top and outlet boundaries, unless nonreflecting boundary conditions are imposed. Periodic boundary conditions are used in the spanwise direction.

Results & Impact

We have used the computational resources of Blue Waters to investigate the characteristics of pressure fluctuations generated by supersonic cold-wall turbulent boundary layers with a nominal Mach number of 6 and a Karman Reynolds number of Reτ ≈ 450; results are published in the Journal of Fluid Mechanics. The DNS have overcome a number of difficulties encountered during experimental measurements of broadband pressure fluctuations and have provided access to quantities that cannot be measured easily, including multi-variate pressure statistics and large-scale structures in the pressure field. As such, the proposed computational efforts cannot be done without the world-class computing capabilities of Blue Waters.

Why Blue Waters

Direct numerical simulations of high-Reynolds-number turbulent boundary layers will be used to study the broadband fluctuating pressure field induced by the boundary layer, with the targeted Reynolds numbers significantly higher than the state of the art. In such simulations, extremely fine meshes are required to fully resolve all the turbulence scales in order to obtain the pressure spectra in the high-frequency/large-wave-number range. In the meantime, the simulations need large-domain sizes to locate very-large-scale coherent structures in the pressure field as well as to accommodate the eddy decorrelation length and to minimize inlet transience as a result of inflow boundary conditions. A large number of time steps are also required for the study of the low-frequency behavior of the pressure spectrum. As such, the proposed computational efforts cannot be done without the world-class computing capabilities of Blue Waters.

Publications and Data Sets

The objective of our research is to enhance the understanding of the four-way coupling effects in droplet-laden isotropic turbulence. (Four-way coupling means there is a two-way coupling between the droplets and turbulence, in addition to a two-way coupling between the droplets themselves.) The deformable droplets are fully resolved in 3D space and time.

The droplet-laden turbulence is simulated by directly solving the Navier–Stokes equations via a variable-density projection method. Most of the computation time is spent in solving the variable-coefficient Poisson equation. In order to improve the computational performance, we have developed an efficient Message Passing Interface (MPI)-based multigrid solver, which is further accelerated with multiple graphics processing units (GPUs).

Further, we examine the kinetic energy transfer between the droplets and turbulence. We also compare the dispersion characteristics of deformable droplets and solid particles with identical diameters and density ratios in isotropic turbulence.

Liquid fuel combustion devices typically atomize fuel into sprays of fine droplets. The fuel droplets disperse in the surrounding turbulent air, vaporize, and mix with it. The resulting chemical reaction transforms the chemical bonding energy into thermal energy, followed by gas expansion, which provides the desired mechanical energy. Understanding the four-way coupling effects between droplets and carrier flow is a necessary prerequisite for the efficient optimization of the energy conversion process.

The method of direct numerical simulation (DNS) of the fully resolved droplet-laden turbulence should be highly efficient in order to allow for the use of the appropriate fine mesh and timestep resolutions on Blue Waters. Our multigrid (MG) solver is essential for the efficient solution of Navier–Stokes equations.

Our recently developed multigrid solver has been efficiently parallelized as shown in Fig. 1. It has the potential to be accelerated with GPU technology. The XK7 GPUs can serve as a good platform for us to further develop our multigrid solver. Furthermore, the Blue Waters staff have always provided us with valuable assistance in postprocessing and profiling. This assistance is essential for our project.

In order to accurately resolve the turbulence length- and timescales as well as the dynamical properties of the deformable droplets, the mesh size of our DNS should be $1024^3$ – $2048^3$. Accordingly, the required number of high-performance computing cores ranges from $33 \times 10^3$ to $260 \times 10^3$ to solve the problem efficiently. Blue Waters’ XK6 nodes can accommodate this requirement. No other facility available to us can provide such a large computational resource.

In postprocessing and profiling. This assistance is essential for our project.
QMCBD: A LIVING DATABASE TO ACCELERATE WORLDWIDE DEVELOPMENT AND USE OF QUANTUM MONTE CARLO METHODS

EXECUTIVE SUMMARY

Blue Waters has enabled us to carry out automated, high-throughput quantum Monte Carlo calculations of condensed matter systems, and to create a database titled QMCBD (Quantum Monte Carlo DataBase) to share these results. We have been able to simulate approximately fifteen different classical semiconductors, which represents the largest set of QMC condensed matter simulations carried out to date. The database contains formation enthalpies and optical excitation energies, calculated via the highly accurate quantum Monte Carlo technique. QMCBD will serve as a platform for easy, searchable data exchange to accelerate the knowledge base around the use of QMC for materials modeling. This work would not be possible without Blue Waters, which allows the calculation of a large class of materials ranging from classic to exotic semiconductor materials, photovoltaics, thermoelectrics, and metallic systems; we can take advantage of the near-linear scaling of our code up to several thousand nodes.

RESEARCH CHALLENGE

Quantum Monte Carlo (QMC) methods are a suite of tools for direct stochastic solution of the many-body interacting Schrodinger Equation. Although QMC methods are considered to be one of the highest-accuracy simulation of complex materials, correlated systems, high-Tc superconductors, and other historically challenging materials. The implementation of an automated workflow (Fig. 1a) for carrying out QMC calculations of solid materials on Blue Waters. The workflow was used together with Blue Waters to carry out our automated quantum Monte Carlo calculations of semiconductor materials. It is now available to the public via our GitHub pages.

METHODS & CODES

We have developed an automated workflow that enables high-throughput simulations of solid-state materials using variational and diffusion Monte Carlo. Our open-source framework is available on our GitHub pages [1], and currently utilizes CRYSTAL [2] to generate trial wave functions and QWalk [3] for quantum Monte Carlo simulations.

RESULTS & IMPACT

The prospect of materials design using high performance computation is one of the most exciting for future technologies. However, the properties of modern materials are rather complex, and current numerical methods can fail to describe them quantitatively. Today, QMC is a state-of-the-art suite of tools for high-accuracy ab initio modeling. It shows great promise for high-accuracy materials modeling and is already well established in the physics community, but its application to real materials with chemical identity remains fairly young. Our goal is to accelerate the development of the QMC community’s collective knowledge base around the use of this method for predictive modeling of real materials. Ultimately, the impact will be to enable quantum Monte Carlo methods to emerge as a standard component of the computational materials modeling toolkit, enabling unprecedented high-accuracy simulation of complex materials, correlated systems, high-Tc superconductors, and other historically challenging materials.

Accomplishments:

- Autogen: The implementation of an automated workflow (Fig. 1a) for carrying out QMC calculations of solid materials on Blue Waters. The workflow was used together with Blue Waters to carry out our automated quantum Monte Carlo calculations of semiconductor materials. It is now available to the public via our GitHub pages.

- Calculation of Thermodynamic Properties of Wide Class of Challenging Materials: We were able to use the Autogen tools to carry out a large, systematic calculation of the thermodynamic properties and formation enthalpies of a large and varied class of materials (Fig. 1b). Our efforts thus far have focused on the calculation of thermodynamic properties (total energies and formation enthalpies) and the band gaps of a wide class of semiconductor materials spanning from conventional (silicon, etc.) to more exotic materials (wide band gap oxides, correlated systems, photovoltaic thin film materials).

- The establishment of QMCBD: Our Mongo database “QMCBD,” which includes results and provenance, is described in detail on the cached National Data Service Project wiki page (http://bit.ly/2fgoZB4). This database will serve as a vehicle to quickly overcome the current expertise hurdle and bring the QMC methodology into the standard computational modeling toolkit. While today there are several competing materials databases for density functional theory, there is no existing QMC database. Thus, Blue Waters has provided us the opportunity to establish this tool for the worldwide QMC community. To our knowledge, this comprises the largest QMC data set that has been established to date.

WHY BLUE WATERS

Blue Waters is the key to carrying out the comprehensive set of QMC materials calculations that populate our database. QMC methods exhibit near-linear scaling on the entire Blue Waters platform, which has allowed us to calculate properties of an extensive set of materials that would otherwise not be possible. Historically, the wide scale adoption of QMC as a method for materials modeling has been limited by its large computational cost, but Blue Waters allows us a first key opportunity to overcome the computational barrier.

PUBLICATIONS AND DATA SETS


Quantum Monte Carlo Database (QMCBD).
EXECUTIVE SUMMARY

Complex turbulent flows are one of the most important types of flows and are prevalent in both natural and built environments. We describe here simulations of two cases of turbulent flows under complex boundary conditions. The first addresses flow over biofilms and similar natural wall roughness. The second simulates sediment transport in bifurcating rivers and channels. These simulations are based on the scalable open source code Nek5000, which employs minimally dispersive spectral element discretization. Results from the biofilm simulations shed light on the interaction of the biofilm with the flow. Results from the flow at the bifurcation shed light on the dynamics of the flow near the bottom of a channel.

RESEARCH CHALLENGE

Turbulent transport is the principal driver for many processes in physics, engineering, geosciences, and biology. Examples include the in-fall of matter into black holes, combustion in automotive and aerospace applications, sediment and pollutant transport in rivers and oceans, and atherogenesis (formation of fatty deposits on arterial walls) in arterial blood flow. Our objective is to address our research questions through direct numerical and large-eddy simulation of turbulent flow by solving the governing Navier–Stokes and associated transport equations. The open problems are as varied as the associated geometries and are challenging because of the range of scales present in turbulent flows at high Reynolds numbers (i.e., high speeds).

The first project studies flow and fine particle transport over biofilms. Biofilms, in the form of microbial communities, serve as a key component in controlling carbon and nutrient cycling in freshwater systems. These microbial communities function as the coupling between physical and biological processes. They have a significant impact on a stream’s hydrodynamics and influence the amount and ability of carbon exported downstream. Most research efforts to date have relied on the use of experimental analysis to understand how biofilm growth affects the flow hydrodynamics and fine particle transport [1]. None of the studies, however, had the spatial and temporal resolution to unravel the interaction between the biofilm structure and the flow, and its effect on fine particle transport. Thus, we conducted Direct Numerical Simulation (DNS) of the flow over biofilms, with the structure of the biofilm provided by experiments that measured the benthic biofilm using an Optical Coherence Tomography microscope. Due to the highly irregular structure of the biofilm bathymetry, new methods for generating the computational mesh were developed.

The second project addresses transport in bifurcating rivers. It has been observed that when a stream divides between a main branch and a side channel, a disproportionate amount of the near-bed sediment is often directed into the side channel, which can ultimately alter the flow dynamics [2,3]. High-resolution Large Eddy Simulations were conducted for flow and sediment transport at idealized experimental scale bifurcations.

METHODS & CODES

The turbulence simulations were based on the open-source spectral element code Nek5000 [4]. The spectral element method (SEM) is a domain-decomposition approach in which the solution is represented by tensor–product polynomials on individual bricks that are assembled to cover the entire domain. The bricks are typically curvilinear, which allows accurate representation of the geometry [5]. The local tensor–product structure allows low-cost and low-storage matrix–matrix product-based operator evaluation so that high-order polynomials may be used with almost no overhead. The SEM thus yields minimal numerical dissipation and dispersion at low cost, which is ideal for simulation of turbulent flows in complex domains. In the simulations with sediment (or fine particle) transport, the particles have been modeled as Lagrangian point particles using a novel semi-implicit timestepping scheme developed to simulate polydisperse particles accurately [6]. For the biofilm problem, the measured biofilm bathymetry data were smoothed while keeping the general structure intact, which was then used to generate the computational mesh. For the first time, spectrally accurate DNS simulations were conducted for a channel with complex natural roughness, and this was possible due to development of sophisticated mesh smoothing algorithms for Nek5000 [7].

RESULTS & IMPACT

DNS of the flow over the biofilm, with a bulk Reynolds number of 8,000, was conducted using about 200 million computational points. Initial results show the interaction of the flow with the biofilm patches, resulting in a higher number of regions of low-velocity magnitude (streaks)—see Fig. 1. One can also observe the effect of the biofilm on the flow, with a relative increase in vortex shedding in the wall-normal direction. As we are conducting the study with “real roughness,” the results will also have an impact on the study of flow over different kinds of roughness, which is important for mechanical and aerospace engineering applications.

The channel bifurcation cases have been simulated with about 250 million computational points; the Reynolds number of the flow is 25,000. One can observe that most of the flow near the bottom enters the side channel even though the total flow is equally divided between the two channels. This clearly hints at the mechanism for the sediment near the bottom to preferentially enter the side channel.

WHY BLUE WATERS

The study of flow in environmental fluid mechanics requires computational resources with sustained computing power at an unprecedented scale. Simulations have been conducted for up to 296 million computational points, with the code scaling strongly up to 32,768 MPI (message passing interface) ranks. Without access to Blue Waters, completing the study within a realistic timeframe would be impossible. In addition, since visualization of a phenomenon is an effective way to understand and explain its mechanics, we will work with Blue Waters project staff to create animations of the phenomena using data from the simulations.

PUBLICATIONS AND DATA SETS

EXECUTIVE SUMMARY

Oscillatory boundary layer flows play an important role on coastal and offshore engineering and the sediment transport mechanics in coastal environments. The present work will be the first computational effort to simulate the effects of turbulent structures and bed roughness on the maximum bed shear stress and the free-stream flow for the case of turbulent oscillatory and momentum exchange between the pore-scale flow and the first numerical study that will examine the mixing layer phase difference with respect to the maximum free-stream velocity and the first numerical study that will examine the mixing layer and momentum exchange between the pore-scale flow and the oscillatory free-stream flow for the case of turbulent oscillatory flow over porous bed. The proposed work combines the expertise of Prof. Marcelo García’s group (Ven Te Chow Hydrosystems Laboratory, Civil and Environmental Engineering) and Prof. Paul Fischer’s group (Computer Science and Mechanical Science & Engineering) with the leading-edge petascale computing resources of Blue Waters available at the University of Illinois at Urbana-Champaign and aims to become one of the most comprehensive studies on the effect of turbulent structures on the oscillatory boundary layer flows and sediment transport.

RESEARCH CHALLENGE

The growing human needs as well as global economic development have resulted in a rapid increase in marine activities. Coastal areas are usually involved in these activities in different ways, from hosting the foundation of offshore structures and breakwaters to accepting the residue of these activities in form of pollution or disturbance of the ecology and sediment transport. Numerical and theoretical models are being used by scientists, engineers, and decision-makers to design infrastructures and study the effect of marine activities on coastal environments and the sediment transport processes. However, most of the current state-of-the-art models fail to accurately predict the interaction of oscillatory flow with sediment transport, highlighting the existing knowledge gaps regarding the complex interactions between the oceanic flow, the coastal bottom, and sedimentation processes.

METHODS & CODES

Using advanced experimental techniques, extensive experiments have been conducted in the Large Oscillatory Water-Sediment Tunnel at the Ven Te Chow Hydrosystems Laboratory. Those experiments suggest the presence of a phase-lag between the maximum bed shear stress and the maximum free-stream velocity in the case of a smooth flat bed [1]. This observation is important for the field of environmental fluid mechanics and coastal sediment transport, as this study is the first one in the literature that supports the hypothesis that the maximum shear stress is lagging instead of leading the maximum free-stream velocity over the period of each oscillation. Nevertheless, due to the limitation of the applied pointwise experimental technique used in the experiments (Laser Doppler Velocimetry), it was not possible to explicitly associate the presence of the bed shear phase lag with the development of the three-dimensional turbulence structures of the oscillatory wave flow, usually referred to as turbulence coherent structures. Thus, turbulence in complex three-dimensional flows is done mainly numerically using a combination of high-order numerical methods with backing from accurate experimental observations from the lab. In the proposed work, we developed a Direct Numerical Simulation model capable of simulating the complex oscillatory boundary layer flow and sediment transport using the Spectral Element Method framework provided by the highly scalable open-source code Nek5000 [2]. Except for the analysis of turbulence characteristics of OBL flow over different bed conditions representative of the coastal bottom, the present work requires use of a proper model for the simulation of the suspended sediment using an Eulerian approach and proper boundary conditions for the sediment mass exchange between the coastal bed and the free-stream flow (e.g., [3-4]).

RESULTS & IMPACT

Preliminary results show good agreement between the experimental observation [1] and numerical results for the case of the transitional/intermittent turbulent regime (Fig. 1A) [5]. Quadrant analysis of velocity fluctuations shows zones of sweeps and ejections (Fig. 1B) that dominate during the whole cycle of the oscillation. Also, velocity fluctuations seem to follow a normal distribution. These turbulence-related events (e.g., turbulent bursting and associated ejections and sweeps) are known as mechanisms that interact with sediment particles [9].

Turbulence coherent structures (Fig. 2) are also examined for the case of oscillatory flow, confirming the presence of turbulence spots in the transitional regime of oscillatory boundary layer flows study the effect of marine activities on coastal environments and the sediment transport processes. However, most of the current state-of-the-art models fail to accurately predict the interaction of oscillatory flow with sediment transport, highlighting the existing knowledge gaps regarding the complex interactions between the oceanic flow, the coastal bottom, and sedimentation processes.

WHY BLUE WATERS

The present work pushes the limit of the turbulent-resolving flow modeling of oscillatory flows. The dimensions of the computational domain are chosen based on prior knowledge of experimental observation of “turbulent spots” [6-7] to ensure that the computational domain is big enough to allow these turbulent structures to develop. A sensitivity analysis has been conducted within the first quarter of the proposed project to ensure that the computational domain size is adequate for the formation of the coherent turbulent structures. The size of the used domain, which is larger than the previous studies’ domains reported in the literature, together with the increased number of computational points (order 0.8 billion points), make the present work the first of its kind in terms of its turbulence/scale resolving results as well as its computational cost. The above, together with the requirement for the solution of several flow cycles to guarantee period-independent results, lead to increased computational requirements, making the use of a leading-edge petascale high-performance computing system like Blue Waters necessary for the success of the present work.

Figure 1A, B: Comparison between numerical results and experimental observations. Distribution and quadrant analysis of velocity fluctuations.
EXECUTIVE SUMMARY
Disastrous earthquakes have revealed the extreme vulnerability to seismic hazards of highly populated areas. This exploratory project contributed toward a reliable seismic risk evaluation, substantiating a future proposal for the development of site-specific seismic scenarios through physics-based 3D models of seismic wave propagation. To be reliable, those models require a high computational burden, with a significant level of parallelization and high-performance computing resources. That is why the leading-edge capability of Blue Waters is necessary. The exploratory project and the requested resources have been used to: 1) evaluate and tune the SPEED (Spectral Elements in Elastodynamics with Discontinuous Galerkin) software for Blue Waters, and 2) demonstrate its readiness for use in a future proposal submission. The next-generation work will be a proposal for a general allocation aimed at developing 3D probabilistic physics-based hazard maps.

RESULTS & IMPACT
The allocated resources have been used to: 1) set up the SPEED software on Blue Waters, 2) evaluate/tune the SPEED code with short tests (tutorials), and 3) evaluate/tune the SPEED code with tests at full machine scale.

METHODS & CODES
The numerical simulations tested on Blue Waters within the allocated exploratory project have been carried on with SPEED [1], an existing open-source high-performance software package. The code belongs to the family of the spectral element method (SEM), a powerful, well-established, numerical technique naturally suited for three-dimensional seismic wave propagation analyses. SPEED allows seismic wave propagation modelling through viscoelastic heterogeneous three-dimensional media, both on the local and regional scale. SPEED reproduces the propagation path of the seismic wave through complex geological structures and localized superficial irregularities such as alluvial basins and civil engineering infrastructures. The code is written in Fortran90 using its pseudo-object-oriented features. It takes advantage of the hybrid parallel programming based on the Message Passing Interface (MPI) library and relies on the domain decomposition paradigm and the OpenMP library for multi-threading operations on shared memory.

WHY BLUE WATERS
Among the most appealing features of physics-based 3D models of seismic wave propagation is the possibility of capturing in its entirety the complex coupling of: 1) directivity pulses; 2) 3D basin effects (Fig. 1); 3) topographic effects; 4) wave scattering; and 5) nonlinear soil response, especially near the source of an earthquake. Accounting for all these features within a single computational model requires a high computational burden (in terms of CPU time and RAM usage), with a significant level of parallelization and high-performance computing resources. That is why the leading-edge capability of Blue Waters was necessary to address this research. To perform our 3D seismic wave propagation simulation would have been infeasible without Blue Waters’ resources. Blue Waters, allowing the running of hundreds of simulations, is essential in a future work to generate the predictions needed for 3D probabilistic physics-based hazard maps.

PUBLICATIONS AND DATA SETS

Figure 1: Spatial variability of peak ground velocity as estimated by 3D numerical simulation referring to the February 22, 2011, Christchurch, New Zealand, earthquake (Mw 6.2).
OPTIMAL BIO-LOCOMOTION STRATEGIES IN FLUIDS

EXECUTIVE SUMMARY

Is the familiar goldfish hiding a technological treasure? Seemingly simple, its swimming around a bowl involves complex interactions among its senses, its body, and the surrounding water. More generally, during millions of years of evolution animals have refined their shapes, gaits, and collective behaviors to master the complex interplay among their bodies, their senses, and the environment. How can we understand the biophysical mechanisms at play and translate them into rational design principles? To inquire into the fundamental mechanisms underlying swimming and flying we have developed an automated inverse design method based on large-scale numerical simulations and artificial intelligence techniques. This allows us to set a desired trait—for example, energy efficiency—and to reverse engineer corresponding optimal solutions. Subsequent computational analysis guides our theoretical intuition toward the identification of broader design principles.

RESEARCH CHALLENGE

All animals that swim or fly are constantly confronted with the physics of the surrounding media. Whether natural creatures can take advantage of physics to push their performance limits depends on their biological strategies, from materials, morphologies, and gaits to collective behaviors. Their struggle for survival has produced a rich array of solutions that often outperform human engineered designs and that work in ways we still do not fully understand. They bear a great potential for technological innovation, with applications ranging from robotics to energy-harvesting devices.

In this context, bioinspired approaches rely on mimicking existing natural solutions to enhance the performance of current engineered designs. Nevertheless, one may question to what extent engineers should follow biomimicry. Indeed, natural creatures have not evolved to optimize engineering objectives, and we have limited information to determine whether a particular solution optimally serves a given function. Moreover, today's materials and components pose constraints and enable opportunities that may differ from their biological counterparts. Hence, we suggest that man-made solutions obtained through inverse design based on an automated optimization process may outperform pure biomimicry.

METHODS & CODES

The characterization of biopropulsion, from the standpoint of optimality, demands accurate, robust, fast, and flexible numerics for flow-structure interaction problems. We have been developing and implementing novel schemes for the direct numerical simulation of individual and multiple swimming bodies. Our algorithms rely on remeshed vortex methods enhanced with projection approaches to capture the effects of the fluid on the body, and with a penalization technique, to capture the effects of the body on the fluid [1]. These techniques enable wavelet-based multiresolution discretization [2], effective mapping of the body on the fluid [1]. These techniques enable wavelet-based multiresolution discretization [2], effective mapping of the body on the fluid [1]. Effective mapping of the body on the fluid [1]. These techniques enable wavelet-based multiresolution discretization [2], effective mapping of the body on the fluid [1].

RESULTS & IMPACT

The coupling of realistic numerical simulations with artificial intelligence techniques is one of the frontiers of Computational Fluid Dynamics and is a unique aspect of this project. We have successfully demonstrated the predictive power of this approach and its ability to provide biophysical insight in the context of rational design of artificial swimmers. For example, we showed that the C-start mechanism (Fig. 1), a widespread escape response among living fish, is optimal as it maximizes the swimmers’ ability to channel water displacement into forward acceleration [5]. Furthermore, we showed that artificial swimmers have the potential to outperform biological ones [6]. We are currently focusing on the design of swimmers able to “sculpt” the surrounding flow in order to achieve multi-tasking behavior (Fig. 2).

WHY BLUE WATERS

Our inverse design process entails thousands of flow-structure interaction simulations, each one characterized by billions of computational elements. Without the sheer size and cutting-edge technology of Blue Waters, these investigations simply would not be possible.
EXECUTIVE SUMMARY

The Advanced Reactors and Fuel Cycles Group (ARFC) conducts modeling and simulation in the context of nuclear reactors and fuel cycles toward the improved safety and sustainability of nuclear power. In the context of high-performance computing, this work requires the coupling of multiple physics models at multiple scales to model and simulate the design, safety, and performance of advanced nuclear reactors. In particular, thermal-hydraulic, heat transfer, mass transport, and nuclear reactor safety phenomena, need to be appropriately modeled and simulated on the most advanced supercomputers in the world. These simulations require greater capability per simulation, are on the horizon for future Moltres work includes generating a high-fidelity, 3D model as well as investigating various transient accident scenarios, additional reactor configurations, and numerous design concepts.

WHY BLUE WATERS

To assess nuclear reactor performance under a variety of conditions and dynamic transients, the ARFC group must conduct a broad range of simulations using the MOOSE framework and our in-house developed modules. Such simulations commonly occupy tens of thousands of CPU cores at a time and vary in completion time. The MOOSE framework has been shown to scale very well up to 10,000 cores. The ARFC group has demonstrated appropriate scaling for MSR simulation above 20,000 CPU cores (600 Blue Waters nodes). Transient and multiscale simulations, which require greater capability per simulation, are on the horizon for our work. These may occupy up to 100,000 CPU cores at a time. Only a few of those larger simulations will be necessary to enable better understanding of the dynamics in these reactor systems.

RESULTS & IMPACT

Blue Waters has enabled ARFC to develop and test a first-of-its-kind, scalable, finite element model of the transient neutronics and thermal hydraulics in a liquid-fueled molten salt reactor design. Moltres is a physics application for multiphysics modeling of fluid-fueled molten salt reactor designs. It couples equations for neutron diffusion, thermal hydraulics, and delayed neutron precursors to transport. Neutron diffusion and precursor transport equations are set up using an action system that allows the user to employ an arbitrary number of neutron energy and precursor groups respectively with minimal input changes. Moltres solves arbitrary-group neutron diffusion, temperature, and precursor governing equations in anywhere from one to three dimensions and can be deployed on an arbitrary number of computational processing units. Moltres is devoted to previously unmatched fidelity in coupled neutronics and thermal hydraulics simulations. MOLTRES is a physics application for multiphysics modeling of fluid-fueled molten salt reactor designs. It couples equations for neutron diffusion, thermal hydraulics, and delayed neutron precursors to transport. Neutron diffusion and precursor transport equations are set up using an action system that allows the user to employ an arbitrary number of neutron energy and precursor groups respectively with minimal input changes. Moltres solves arbitrary-group neutron diffusion, temperature, and precursor governing equations in anywhere from one to three dimensions and can be deployed on an arbitrary number of computational processing units. Moltres is devoted to previously unmatched fidelity in coupled neutronics and thermal hydraulics simulations.

MOLTRES is a physics application for multiphysics modeling of fluid-fueled molten salt reactor designs. It couples equations for neutron diffusion, thermal hydraulics, and delayed neutron precursors to transport. Neutron diffusion and precursor transport equations are set up using an action system that allows the user to employ an arbitrary number of neutron energy and precursor groups respectively with minimal input changes. Moltres solves arbitrary-group neutron diffusion, temperature, and precursor governing equations in anywhere from one to three dimensions and can be deployed on an arbitrary number of computational processing units. Moltres is devoted to previously unmatched fidelity in coupled neutronics and thermal hydraulics simulations.

PUBLICATIONS AND DATA SETS


UNDERSTANDING HYDROGEN STORAGE IN METAL ORGANIC FRAMEWORKS USING MASSIVELY-PARALLEL ELECTRONIC STRUCTURE CALCULATIONS

EXECUTIVE SUMMARY

Hydrogen has the potential to be a major green energy source, but, for use as a fuel, it requires efficient storage materials for retaining and releasing hydrogen in large quantities. Metal–organic frameworks (MOFs) are potential storage materials, and their hydrogen storage potential and microscopic properties need further investigation. We perform first-principles quantum-mechanical molecular dynamics calculations to understand the behavior of hydrogen inside MOFs. We include the effect of quantum nuclear motions, which are critical for the properties of hydrogen, the lightest atomic element. These challenging calculations are only possible on large-scale and tightly coupled petascale computer systems. The effects of quantum mechanics on density functional theory. In addition to large-scale molecular dynamics simulations of the diffusion and dynamics of hydrogen inside MOFs, we describe the atomic nuclei in the entire system quantum mechanically via a path integral formalism. The effects due to the quantum fluctuations of the nuclear degrees of freedom are critical for understanding the binding and dynamics of light elements such as hydrogen.

METHODS & CODES

Answering the above questions requires an accurate quantum-mechanical simulation method at finite temperature. The method must work beyond the harmonic approximation as hydrogen motion can be diffusive and highly anharmonic. Due to the use of complex ligands chelated to metal ions and the metal ions themselves, standard pairwise force fields can be problematic in describing the energetics, and parameterizing a more transferable model in both-time-consuming and difficult. We employ the Path Integral Car–Parinello Molecular Dynamics simulation technique (PI-CPAIMD), which allows the nuclei to move on the Born–Oppenheimer energy surface provided by plane wave-based Density Function Theory (DFT) and includes quantum effects via Feynman Path Integrals. By including the (valence) electronic degrees of freedom explicitly, we bypass force-field difficulties; by employing path integrals, nuclear quantum effects are treated in a completely general way and are converged with a single parameter.

The use of the PI-CPAIMD method and the high-quality results it delivers requires highly scalable software, developed by us under NSF support (NSF SHI grants ACI-1339804 and ACI-1339715), and named OpenAtom. CPAIMD, which simulates classical mechanics of the nuclei moving on the DFT-derived energy surface, has a computational workload scaling as $N^3$ where $N$ is the number of atoms in the simulation. Adding Path integrals, the computational cost increases by the factor $P/P_0$, where $P$ is the number of “beads” used to discretize the Feynman path integral. Typically, $P$ ranges between 20 and 80, which is a significant increase in computational requirements. As one can imagine, the computational cost requires petascale resources, and a supercomputer such as Blue Waters is needed to carry out this work. Work on Blue Waters is supported by NSF PRAC grant OAC-1614491.

RESULTS & IMPACT

In the first year of our PRAC project on Blue Waters, we have spent the majority of the time studying a model MOF system with hydrogen to understand the behavior of the system and its various timescales. Fig. 1 illustrates the unit cell of the MOF crystal with hydrogen molecules inside. A smaller version of the MOF was carefully studied with CPAIMD long simulations, which permitted us to obtain preliminary results. Fig. 2 shows the results of the CPAIMD simulations (without quantum effects) for the hydrogen probability distribution inside the MOF when the system is at 300 K. The hydrogen molecules spend most of their times in the “voids” (corner regions) away from the MOF backbone, signaling weak binding to the MOF backbone at these temperatures and potentially rapid diffusion. Other measures of diffusion (e.g., the hydrogen diffusion coefficient) are in good agreement with prior results, giving us confidence that the benchmarking has been successful in producing physically accurate and important results.

The next step is to perform CPAIMD simulations with the full larger MOF system and then to turn on quantum nuclear effects via the path integral approach to understand how the quantum effects modify the behavior of hydrogen inside the MOF.

SUCCESSFUL modeling of the behavior of hydrogen inside MOFs will provide valuable information to the community of materials and energy researchers, in academia as well as at national and industrial labs, regarding the underlying reasons for the performance of these materials. Chemical modifications of the MOF structure can then be explored using such simulations to help design improved hydrogen storage performance.

WHY BLUE WATERS

Massively parallel electronic structure calculations require tightly coupled computing nodes due to intense communication loads. Electron waves are delocalized over the entire system so all parts of the system end up interacting with each other. For the MOF system of interest, the CPAIMD simulations already require a massively parallel calculation with many hundreds of nodes. The inclusion of nuclear quantum effects, however, means that only a tightly coupled petascale computer such as Blue Waters is capable of delivering results on a reasonable (one- to two-year) timescale.
ATOMIC MODELING OF TRANSFORMATIONS IN NANOCRYSTALS

Allocation: Exploratory/30 Kwh
PI: Prashant K. Jariwala
Co-PI: Sudhakar Pamidighantam

1University of Illinois at Urbana-Champaign
2National Center for Supercomputing Applications

EXECUTIVE SUMMARY

Discovery of materials drives renewable energy and optoelectronic technologies. While using an unconventional method of nanostucture synthesis in our lab, we encountered a novel form of mercury–cadmium selenide. The wurtzite structure of this crystal distinguishes it from the natural zincblende phase. The properties of the material were unknown, motivating our computational investigations of its electronic structure. These computations revealed how subtle differences in crystallographic symmetry combine with the relativistic nature of electrons to result in novel behavior. Natural mercury selenide is a zero-band-gap semiconductor. However, the wurtzite form has a band gap and is a three-dimensional topological insulator, expected to exhibit spin-protected conduction on the surface, while being electrically insulating in the bulk, a property useful for next-generation logic devices and the discovery of exotic quasi-particles. The study also elucidates how crystal structure and chemical composition can be paired to tune relativistic effects and topology of electron motion.

RESEARCH CHALLENGE

Engineered nanocrystals are often utilized for making new functional electronic and optical materials such as superionic solids and battery electrodes. Our laboratory makes use of unconventional methods that enable manipulation of the chemical composition and crystal structure of nanocrystals. These techniques often produce novel compositions and crystal phases that are often not found in the bulk phase diagram. Computational electronic structure investigations are allowing us to explore the properties of these new, unconventional materials that have been created for the first time. Alongside, we are also elucidating chemical trends in heterostructures and alloys and developing solid-state physics principles from these trends. The results from our investigations will enable the rational design of new phases and compositions with targeted applications for resolving longstanding challenges of energy storage and device efficiency.

With advances in nanotechnology and chemical synthesis, materials are becoming ever more complex. Computations can uncover chemical principles that will ultimately allow prediction of the properties of tomorrow’s indispensable materials, an existing Grand Challenge. However, these studies require extensive calculation spanning a range of physicochemical parameters. As opposed to a single large and expensive calculation, our work requires a library of moderately expensive calculations. The net cost for generating such a library of data is feasible only with a resource such as Blue Waters, with considerable payoff for future scientific advances. Dissemination of solid-state physics principles, like those resulting from our project, will accelerate this process.

METHODS & CODES

We used the open source Quantum Espresso software suite [1] to run our electronic structure calculations. In order to study the effects of chemical composition and structure on the alloy’s electronic properties, numerous calculations were run. Each calculation is distinguished from the others in the crystal geometry, chemical formula, or both. Comparisons of energies, electron distribution, and band structure across all the calculations allowed us to make predictions about novel behavior of the wurtzite polymorph of the alloy and also understand the role of relativistic effects and symmetry breaking.

RESULTS & IMPACT

Our electronic structure calculations demonstrated that band elongation in a novel polymorph of HgSe and HgCd0.5Se0.5 alloys is responsible for the opening of a band gap (Fig. 1). The presence of this band gap is of significance because it, when combined with the inverted nature of bands in HgSe, qualifies these materials as potential three-dimensional (3D) topological insulators (TIs). 3D TIs are of interest because electrons at their surface states are spin-protected from back scattering. This protection allows 3D TI materials to conduct electrons along their surface without resistance, that is, with minimal loss of energy as wasteful heat. For this reason, 3D TIs are garnering interest as components of energy-efficient logic devices that can operate at high capacity while simultaneously not contributing to the generation of heat. Such devices would reduce the need for coolants and could permit the design of more powerful supercomputers by overcoming problems caused by overheating. TI materials are also expected to harbor exotic quasi-particle states of importance in particle physics. However, only a handful of TI materials have been identified, and through our contribution we are adding this alloy system to the list of 3D TIs. In addition, our work unveils crystallographic anisotropy as a powerful synthetic handle for tuning band topology.

Our study further contributed two new principles governing band structure and topology in HgCd0.5Se0.5 and similar alloys of a strongly relativistic and a weakly relativistic metal. The first of these new insights involves the effect of mixing electron character (Fig. 2a–2c). HgSe and CdSe have different energy ordering of their bands; the band order is inverted in HgSe with respect to the conventional CdSe case. This inversion is due to the greater relativistic mass of the Hg valence “s” electrons resulting in the contraction and energetic downshift of their orbitals. Thus, in HgCd0.5Se0.5 alloys, wherein the electronic character of the two metals is mixed, the energy of the “s”-like band, the resulting band ordering, and the band gap depend on the relative abundance of Hg and Cd. At compositions of the wurtzite alloy, where the “s”-like band of spherical symmetry becomes close to the valence band edge, the light hole and heavy hole energies become less sensitive to crystallographic anisotropy. As a result, a smaller gap is exhibited between these two bands, negating the effect of the symmetry breaking caused by crystalline anisotropy. The second finding relates to the effect of a layered structure of the crystal (Fig. 2d). Ordering of the two metals Hg2+ and Cd2+ in alternate layers mimics the effect of crystallographic anisotropy and provides a second means by which a band gap can be opened. These principles gained from our study translate readily to existing TI systems and also enable the design of new alloys with TI behavior.

WHY BLUE WATERS

Calculations of crystal surfaces and nanocrystals may be accomplished on computational resources other than Blue Waters. However, a large number of single calculations of these structures with varying elemental composition is required to study chemical trends. The computational expense of such an effort would be prohibitive for us were it not for a Blue Waters allocation. Furthermore, the specialized hardware of Blue Waters allows the Quantum Espresso code to run even more efficiently. This is because Quantum Espresso’s parallelization schemes involve sizable and frequent communication among CPUs, which rely on the speed of the Blue Waters communication hardware. Furthermore, our ability to work with Sudhakar Pamidighantam of NCSA as a co-PI has allowed us to greatly expedite code preparation and troubleshooting with use of the SeaGrid portal [2].

PUBLICATIONS AND DATA SETS


Figure 1: In the novel wurtzite form of HgSe, anisotropy elongation of Hg-Se bonds results in the opening of a band gap between the highest energy valence (green curve) and lowest energy conduction (red curve) bands. Hg atoms are shown in purple and the electron density is shown in red.

Figure 2: Electronic density (red) maps elucidating trends in HgCd0.5Se0.5 (a) Hg, purple; Cd: grey; Se: yellow). Unlike CdSe (b) and HgSe (c), alloys (b) show significant mixing of the states, which weakens the effect of bond elongation. (d) Creation of alternating Cd and Hg layers has a similar effect as bond elongation.
NUMERICAL SIMULATIONS OF COLLAPSING CAVITATION BUBBLES ON BLUE WATERS

Allocation: GLPC/880 Kneh
PI: Eric Johnson¹
Co-PIs: Shahaboddin Alahyari Reig¹, Mauro Rodrigues²
¹University of Michigan

EXECUTIVE SUMMARY

The collapse of cavitation bubbles is known to cause damage, ranging from the erosion of naval propellers to soft tissue ablation. While significant attention has been dedicated to investigating this phenomenon in the case of hard materials such as metals, less is known about cavitation-induced damage to soft materials. In order to elucidate these damage mechanisms, we perform high-fidelity numerical simulations of the collapse of individual and multiple vapor bubbles near both rigid and compliant surfaces. We have developed a computational framework to conduct massively parallel simulations of the three-dimensional compressible Navier–Stokes equations for gas–liquid flows. These simulations provide a clearer image of the detailed nonspherical bubble dynamics, pressure, and temperature fields, and stresses/deformations of the neighboring solid. This knowledge and data enable us to develop numerical models for the collapse of bubble clouds that can be used in biomedical or naval hydrodynamic applications of interest.

RESEARCH CHALLENGE

Cavitation occurs in a wide range of hydraulic applications, such as naval engineering, turbomachinery, and biomedical ultrasound. In contrast with boiling, in which local temperature reductions of a liquid lead to the formation of vapor bubbles, these cavitation bubbles dramatically respond to pressure changes, growing to sizes far greater than their equilibrium radius and undergoing a violent inertially dominated collapse [1]. As a result, shock waves and high-pressure regions are produced, which may damage neighboring solid objects such as propellers or soft tissues [2–5]. While relatively well understood in the context of hard materials (e.g., metals), cavitation erosion to soft matters is not well known [3]. Given the prevalence of cavitating flows in a vast variety of applications, there is an essential need to investigate the detailed bubble dynamics and to predict pressures, temperatures and deformations produced. However, compressibility effects, including the propagation of shock and rarefaction waves in a multiphase medium, result in a complicated nonlinear multi-scale and multi-physics problem that is challenging to solve. Moreover, owing to the wide range of temporal and spatial scales of these flows, precise and accurate measurements are nearly impossible to obtain experimentally. Numerical simulations of effects not available and/or feasible via experimentation have therefore emerged as a powerful tool to complement and enhance our fundamental understanding of these flows [6].

METHODS & CODES

In order to perform high-resolution simulations of the three-dimensional compressible Navier–Stokes equations for a gas–liquid system, we have developed a novel parallel algorithmic framework [7]. We used an accurate model of compressible multiphase flows capable of resolving flows around cavitating bubbles to correctly compute pressures and temperature across the material interfaces. The algorithm employs a third-order accurate explicit strong-stability-preserving Runge–Kutta scheme to march in time [8]. For the spatial discretization, we proposed a solution-adaptive high-order accurate, central difference/discontinuity-capturing method. This method can represent both broadband flow motions and discontinuities accurately and efficiently. The basic idea is that nondissipative methods are used where the solution is smooth, while the more dissipative and computationally expensive capturing schemes are applied near discontinuous regions. For this purpose, a discontinuity sensor discriminates between smooth and discontinuous (shocks, contacts, and interfaces) regions, which all require a different treatment. For large-scale parallel calculations, our in-house petascale production code achieves parallel efficiency over 85% on 4,096 processors using MPI. We have also been exploring alternative approaches such as GPU acceleration. Our code also uses the parallel HDF5 library to manage large and complex data collections.

RESULTS & IMPACT

The current project focuses on two specific problems: the collapse of a single bubble near solid/soft media and the collapse of multiple bubbles near solid/soft media. The first problem provides insight into the detailed dynamics of the collapse, including the nonspherical behavior of the bubble, high-velocity jet formation, propagation of shock waves, and the vortex ring convecting toward the boundary. The simulations provide the flow field pressure and temperature distribution throughout the collapse, which can be used to model cavitation erosion (Fig. 1). Although studying single-bubble collapse is valuable specifically in exploring the flow physics, the disruptive effects of cavitation erosion are generally caused by the collapse of bubble clouds containing tens of thousands of bubbles. However, resolving every bubble is computationally prohibitive, such that a model representing the collective behavior of the bubble cloud is necessary. Currently, such models are rudimentary [9], as they only account for spherical bubble dynamics. By conducting resolved simulations of individual bubbles, we are investigating the complex interactions within the collapsing bubbles in order to establish a numerical cloud model that includes the bubble–bubble interactions and the nonspherical effects of the collapse.

This project will deepen our knowledge and understanding of the nonspherical behavior of the bubbles in cavitating flows. Developing a comprehensive model for bubble clouds will result in more precise numerical simulations of the collapse of bubble clusters. This will help to reduce the disruptive effects of cavitation erosion in naval applications and turbomachinery. Reducing cavitation erosion will significantly strengthen the structures exposed to deterioration caused by cavitating flows, extend the lifetime of machineries, and has the potential to save billions of dollars each year. Moreover, a well-known treatment in biomedical utilities structural damage induced by collapsing bubbles to break apart kidney stones. The same idea has been applied to destroying malignant cells and soft tissue. However, a negative side effect of these medical advancements is the unwanted damage to neighboring cells and tissues. This study delves further into these matters to provide insight into damage mechanisms to better control them in order to minimize the impairment of healthy cells and tissues.

WHY BLUE WATERS

This project utilizes two different in-house codes: our petascale production code for the large-scale simulations (based on MPI) and our next-generation code that enables larger-scale heterogeneous architectures (based on MPI and GPUs). These codes solve the compressible Navier–Stokes equations for multiple gases and liquids. The foundations of both codes are high-order accurate algorithms, explicit in time and in space, thus naturally lending itself to massive parallelization. To carry out accurate simulations of cavitating flows that effectively resolve the small-scale features, extremely high spatial resolution is essential to run even a single simulation over long compute times, which is difficult to achieve on any other NSF computing resource. Given its speed and available computation power, the Blue Waters supercomputer is capable of providing us with this opportunity.
LEADING FUTURE ELECTRONICS INTO THE NANO REGIME USING QUANTUM ATOMIC SIMULATIONS IN NEMOS

Allocation: NSF FAAC/1,239 Ish
PI: Gerhard Klimeck
Co-PI: Tillmann Klusbauer

Purdue University

EXECUTIVE SUMMARY

With a revenue of $338.9 billion in 2016, and as an enabler for larger economy chains (i.e., electronic systems), the semiconductor industry influences approximately 10% of the world GDP [1]. The transistor is at the heart of this enormous industry and continuous improvements of transistors in terms of speed and power consumption are essential for the stability and growth of the semiconductor industry as well as the dependent product chains and economies. Needed improvements in transistors’ performance have driven the semiconductor industry to push for smaller transistors, reaching 14-nanometer (nm) in the latest technology node, while development is ongoing for 10-nm technology and beyond. Such aggressive downscaling into a countable number of atoms in the critical dimensions makes atomic simulations necessary pathfinders in the quantum regime. NEMOS is designed to comprehend the critical multiscale, multi-physics phenomena of nano-atomic dimensions that occur in transistors, and enables quantitative study of new generations of nano-electronic devices even beyond transistors [2–3].

RESEARCH CHALLENGE

The U.S. has always been a world leader in the semiconductor industry with 40% of the worldwide semiconductor device-related patents originating in the U.S. [4]. The U.S. semiconductor industry is one of the nation’s largest and most strategic industries, and the U.S. holds one-third of the global semiconductor device market worth over $300 billion per year. Simultaneously, a relentless pursuit of nanoscale technology through efficient computational approaches, and enables quantitative study of new generations of nano-electronic devices even beyond transistors [2–3].

METHODS & CODES

NEMOS is designed to comprehend the critical multiscale, multi-physics phenomena of nano-atomic dimensions that occur in transistors, and enables quantitative study of new generations of nano-electronic devices even beyond transistors [2–3].

RESULTS & IMPACT

For nitride devices, the I–V characteristics produced from these simulations agree quantitatively with experimental measurements. The simulations have been used to suggest improvements in the multi-quantum-well nitride-based light-emitting diode. In addition to this, a new alloy engineered TFET is proposed as a novel low-power transistor design. Regardong transistor simulations, the tunneling from the valence to conduction band has the potential to yield significantly improved subthreshold slopes to allow lower supply voltages and much-needed lower power consumption compared to MOSFETs. For nitride devices, the I–V characteristics produced from these simulations agree quantitatively with experimental measurements. The simulations have been used to suggest improvements in the multi-quantum-well nitride-based light-emitting diode. In addition to this, a new alloy engineered TFET is proposed as a novel low-power transistor design.

WHY BLUE WATERS

Quantum transport simulations are very computationally expensive and memory demanding due to the high degree of complexity of the equations used, especially if incoherent scattering of particles is needed. A toy quantum transport calculation of a 50 nm long wire with a 3 nm diameter requires around 1 teraflop for a single energy point where more than 1,000 energy points are needed. And this calculation must be repeated perhaps hundreds of times for a full current-voltage sweep. The treatment of a realistic device would require an atomistic resolution of a device with a cross section of more than 200 μm², which includes the core semiconductor and the surrounding gate material. Such devices of larger sizes are especially an issue due to the O(e^6) scaling of matrix operation time-to-solution and O(e^7) scaling of memory. Blue Waters was used for running such simulations on up to 16,384 cores per simulation. In many cases the work could not be accomplished in a reasonable amount of time without Blue Waters, and for the larger simulations the work could not be accomplished on other available systems. Blue Waters staff provide exemplary support and user outreach to guide system usage, help with issues as they arise, and assistance with code performance and scaling.

PUBLICATIONS AND DATA SETS


Figure 1: A 3D visualization of a typical sub-10 nm novel transistor design made from bilayer phosphorene that is being investigated using NEMOS. (Credit: Tarek Ameen)

Figure 2: Energy resolved electrons, hole density of states (contour lines) in blue with electrons and holes (color contours) of energy efficient multi-quantum-well light-emitting diode simulated by NEMOS. (Credit: Junzhe Gong, Gerhard Klimeck)
KINETIC SIMULATIONS OF UNSTEADY SHOCK-BOUNDARY LAYER INTERACTIONS

EXECUTIVE SUMMARY

Hypersonic flow over configurations such as a double wedge at continuum-like free stream conditions has been a challenging problem because of the multiple shock–shock and shock–boundary layer interaction, separated flows near the hinge, shear layer, and three-dimensional effects. These conditions generate a mesh that is highly nonuniform because of very high levels of refinement near the surface due to extremely high flow gradients in temperature and pressure. The octree cells lying in this region are highly refined as compared to those in the free stream and inside the geometry leading to a high degree of load imbalance among the processors, which can cause significant increases in communication time. In addition, these flows are unsteady in nature but provide an opportunity to study flow stability mechanisms and identify near-transition behavior. However, large sizes of additional particle data must be collected in order to analyze the time-dependent signals. The multiscale nature of these gradients in fluid density, high fluid temperatures, and a large degree of nonequilibrium. The time-accurate, large-scale DSMC simulations performed to obtain the results shown in Fig. 2 require on the order of 100,000 node-hours to reach a steady state of 1 millisecond. These are the first simulations that enable the hypersonic laminar shock boundary layer interaction community to understand the role of slip in a fully 3D simulation. In addition, we have made extensive use of the CPMAT and Perf-tools profiler on Blue Waters for testing these algorithmic improvements.

RESEARCH CHALLENGE

Hypersonic compressible flows are characterized by significant gradients in fluid density, high fluid temperatures, and a large degree of nonequilibrium. The multiscale nature of these problems makes it a challenge in terms of flow physics and computational intensity. This work was motivated by the study of laminar shock wave boundary layer interaction (SWBLI) problems, which are especially difficult to accurately model due to complex shock interactions, flow separation, three-dimensional effects, and shear layers. All these effects play a significant role on aerothermodynamic quantities such as heat transfer, skin friction, and pressure loads over different angular sections. The particle-based Direct Simulation Monte Carlo (DSMC) method is a good candidate to simulate such flows because it solves the Boltzmann transport equation exactly; hence, it is valid even in strong shock regions where the continuum assumptions inherent in the derivation of Navier–Stokes equations fail. However, the extent of its applicability to such problems has been limited by the high computational requirements arising from the need to simulate a large number of particles to satisfy DSMC requirements such as having the cell size comparable to the local mean free path, a time step that is smaller than the local mean collision time, and sufficient numbers of particles per cell to perform collisions.

METHODS & CODES

We have developed a three-dimensional MPI-parallelized DSMC code known as Scalable Unstructured Gas-dynamic Adaptive mesh Refinement (SUGAR). The code uses new techniques to simulate hypersonic, compressible flows such as an octree-based adaptive mesh refinement (AMR) implementation for capturing multiscale physics, linearized representation of an unstructured grid using Morton-Z space filling curve for efficient access of computational cells, an accurate cut-cell algorithm to compute correct volume of intersected computational cells, algorithmic improvements for efficient gas–surface interactions, and array-based data structures for optimal use of cache memory utilization [1].

A great deal of work has been done to improve the scalability of the code. In brief, the improvement in the strong scaling, as compared to the original SUGAR version presented at the Blue Waters conference approximately two years ago, is shown in Fig. 1, where a near-ideal scaling is now obtained for a 128 times increase in the number of processors for hypersonic flow over a hemisphere with 96 million computational particles (with X=32 for the blue and purple lines). The result presented two years ago (green line with X=128) achieved poor parallelization, and exhibited very long run times well. In addition, 87% weak scaling was obtained for 8,192 processors for a hemisphere flow with 24 billion particles.

RESULTS & IMPACT

We have recently simulated an even more computationally challenging flow over a double wedge using 768 nodes, with 24 billion particles and 1.8 billion computational cells. The double wedge geometry has the forward and aft angles of 30° and 55°, with a length of 0.08 mm for the first wedge. The flow conditions corresponded to a free stream Mach number, velocity, static pressure, and unit Reynolds number of 7.14, 3.812 m/s, 0.79 kPa, and 52,200, respectively. For this simulation shown in Fig. 2, the VHS model and the majorant frequency scheme were used and gas–surface interactions were modeled using the Maxwell’s model with full momentum and energy accommodation.

Fig. 2 shows the resolution of the highly complex shock physics that occurs in this domain of hypersonic shock–boundary layer interaction flows. An oblique shock is formed at the leading edge of the lower wedge and a bow shock forms in front of the upper wedge. As the flow moves up the lower wedge, an adverse pressure gradient is formed and the boundary layer separates, which generates a separation shock. These three shocks meet at the triple point, and with their interaction a transmitted shock (also known as the reattachment shock) is formed. The reattachment shock impinges on the upper wedge and the thickness of the boundary layer reduces at the impingement point. This point is usually where the boundary layer reattaches to the surface. The zone isolated among the hinge, separation point, and the reattachment point is the separation zone where recirculation is observed. Note that the flow downstream of the bow shock is subsonic, but the flow going through the oblique, separation, and reattachment shocks remains supersonic. Therefore, a contact surface, also known as the shear layer, is formed across which the pressure is constant; however, the velocities and temperatures are different. The reattachment shock, after reflecting off the boundary layer, forms expansion waves, which reflect at the contact surface and coalesce to form a compression wave. Even though this flow is still laminar, it is unsteady, and future calculations on Blue Waters will enable us to use global linear stability analysis to understand the different excitation and decaying eigenmodes as well as the nature of the disturbance in terms of acoustic, entropy, or vortex waves [2].

WHY BLUE WATERS

The times-accurate, large-scale DSMC simulations performed to obtain the results shown in Fig. 2 require on the order of 100,000 node-hours to reach a steady state of 1 millisecond. These are the first simulations that enable the hypersonic laminar shock boundary layer interaction community to understand the role of slip in a fully 3D simulation. In addition, we have made extensive use of the CPMAT and Perf-tools profiler on Blue Waters for testing these algorithmic improvements.

PUBLICATIONS AND DATA SETS


HIGH ENERGY PHYSICS ON BLUE WATERS

Allocation: NSF PRAC/21,780 Knsh
PI: Paul Mackenzie4
Co-PIs: Norman Christ, Robert Sugar
Collaborators: Alexei Bazavov, Peter Boyle, Carleton DeTar, Steven Gottlieb, Doug Toussaint, Chulwoo Jung, Yuzhi Liu

1Fermi National Accelerator Laboratory
2Columbia University
3University of California, Santa Barbara
4Michigan State University
5University of Edinburgh
6University of Utah
7Indiana University
8University of Arizona
9Brookhaven National Laboratory

EXECUTIVE SUMMARY

The goal of this project is to carry out groundbreaking studies of the standard model of high-energy physics. The calculations we have in progress address basic questions in high-energy physics and are directly supportive of the very large, worldwide experimental studies in this field. This project is a joint effort of the Fermilab Lattice, MILC, and RBC Collaborations, which among them contain almost all of the high-energy physicists in the United States working on the numerical study of the standard model. The advent of petascale computers such as Blue Waters has had a transformational impact on our field. Members of our three groups have played a leading role in this transformation through the development of algorithms and community codes, and by carrying out petascale simulations. We are using Blue Waters to build upon these advances in order to make major progress in understanding the fundamental laws of physics.

RESEARCH CHALLENGE

The standard model of high-energy physics encompasses our current knowledge of the fundamental interactions of nature. It consists of two quantum field theories: the Weinberg–Salam theory of electromagnetism and weak interactions, and quantum chromodynamics (QCD), which is the theory of the strong interactions. The standard model has been enormously successful in explaining a wealth of data over the past forty years; however, high-energy physicists believe that a more general theory will be needed to explain physics at the shortest distances or highest energies. The research in which we are engaged aims at obtaining a deeper understanding of the standard model and at searching for physical phenomena that go beyond it.

METHODS & CODES

QCD is formulated in the four-dimensional space–time continuum; however, in order to carry out numerical calculations, one must reformulate it on a lattice or grid. To obtain physical results, one carries out calculations for a range of small lattice spacings and then performs extrapolations to the zero lattice spacing (continuum) limit. This continuum extrapolation is one of the major sources of errors in lattice QCD calculations. Another important source of systematic errors arises because the calculations must take place in a finite box, whose physical size must be much larger than the largest lengths in the problem. Keeping both the continuum extrapolation and finite size effects under control requires working on very large lattices. The power of petascale computers is critical for enabling us to do this.

A number of different formulations of quarks, an elementary particle of matter in QCD, on the lattice are currently being used in numerical studies of QCD, all of which are expected to yield the same results in the continuum limit. We are using the two formulations most widely employed in the study of high-energy physics: domain wall fermions (DWF) and highly improved staggered quarks (HISQ). The DWF and HISQ actions each have important advantages for different aspects of our work. DWFs have nearly exact chiral symmetry at finite lattice spacings. This high degree of chiral symmetry is required for key studies of kaon decays that lead to precise tests of the standard model. On the other hand, staggered quarks are essential for studies of the decays and mixings of particles with heavy quarks for which chiral symmetry plays a less important role. However, the large lattice volumes, made accessible by the lower computational cost of staggered fermions, are necessary for accurate control of finite lattice spacing errors. The HISQ calculations are aimed at precise determination of some of the least well-known parameters of the standard model, and at making further precise tests of it.

Lattice QCD calculations proceed in two steps. In the first, one uses importance-sampling techniques to generate gauge configurations, which are representative samples from the Feynman path integrals that define QCD. These configurations are saved, and in the second step they are used to calculate a wide variety of physical quantities. Generating gauge configurations is the rate-limiting step and requires the most capable supercomputers available.

RESULTS & IMPACT

During the first year of our PRAC grant, we have used our allocation to generate DWF and HISQ gauge configurations that are among the most challenging produced to date. The initial applications of the DWF configurations will be to study two processes that are highly suppressed in the standard model and therefore serve as places where physics beyond the standard model may emerge. The first of these is the direct violation of CP (charge conjugation parity) symmetry in the decay of neutral kaons. The second is the determination of the mass difference between the two neutral kaon-decay eigenstates, which is the smallest particle mass difference ever measured.

The first application of the HISQ gauge configurations generated under this allocation has been to enhance the determination of quark masses and leptonic decay constants of unprecedented precision.

One particularly striking result is the determination of the leptonic decay constant of the B meson to a precision of 0.4%, a factor-of-five reduction of uncertainty from the world average. This calculation is illustrated in Fig. 1. A preliminary report on this work was published in [1]. Near-final results were presented at the Lattice 2017 conference, and a journal article is in progress.

WHY BLUE WATERS

Work on lattice QCD calculations has made major progress in the last few years, with a limited number of calculations reaching precision of a fraction of a percent, and techniques in place to determine many more quantities to this level of accuracy. Such precision is needed to test the standard model and to obtain a quantitative understanding of physical phenomena controlled by the strong interactions. This progress has been enabled by the advent of petascale computers, such as Blue Waters, and could not have been made without them.

PUBLICATIONS AND DATA SETS

FIRST-PRINCIPLES COMPUTER SIMULATIONS OF HYDROCARBONS UNDER FUSION CONDITIONS

Allocation: NSF PRAC/9,200 Koh
PI: Burkhard Miltner1
Co-PIs: Shuai Zhang1,2, Francois Soubiran1, Kevin Driver1,2
1University of California, Berkeley
2Lawrence Livermore National Laboratory

EXECUTIVE SUMMARY

Matter in the interiors of giant planets and stars is exposed to conditions of extreme temperature and pressure. In addition to the astrophysical relevance, a rigorous and consistent theoretical description of materials properties in the warm dense matter and dense plasma regimes has been identified as a central goal in the development of key energy technologies such as advanced nuclear reactors and inertial confined fusion, shock physics, plasma science, and stockpile stewardship. The optimal design of new plasma experiments relies on computational models of the equation of state, transport, and optical properties in order to achieve desired pressures and temperatures and to make the first, key measurements. Here we use first-principles computer simulations to study the properties of hydrocarbons at extreme conditions.

RESEARCH CHALLENGE

The development of a first-principles methodology for warm dense matter (WDM) applications that treat temperature effects consistently is a key component of the stewardship of plasma science [1]. Indeed, technological progress in high-energy-density physics (HEDP) applications, such as fusion energy [2], shock-wave physics [3], astrophysical processes [4], and planetary [5] and stellar interiors, relies on simulations for input and guidance. WDM is broadly described as the HEDP regime between condensed matter and ideal plasmas, where strong electron correlation, and quantum and ionization effects, are all important.

Hydrocarbons are the primary materials used for the ablator in inertial confinement fusion target capsules. It follows that the treatment of the correct equation of state (EOS) of hydrocarbon ablators is important in order to optimize experimental designs to achieve desired density and temperature conditions. In this work, we perform state-of-the-art, benchmark-quality EOS calculations by combining high-temperature path integral Monte Carlo (PIMC) data and low-temperature density functional theory (DFT) data to construct coherent EOSs for several hydrocarbon materials over a wide range of densities and temperatures. At a commonly accessible temperature of 10^5 K, we find PIMC and DFT-MD predict consistent internal energies and pressures, validating their accuracy. While there have been several previous simulations of CH reaching into the WDM regime based on DFT-MD methods alone, our calculations provide the first set of first-principles data across the entire WDM regime that properly treats the many-body and shell-structure ionization effects. These processes affect the structure of the Hugoniot curve, which represents the density that can be achieved with various shock wave experiments. Our simulation results provide a benchmark for widely used EOS tables such as SESAME and QEOS as well as orbital-free DFT and average-atom methods.

METHODS & CODES

Since we need to cover a large temperature interval that spans five orders of magnitude (10^10°K), we rely on two different first-principles simulation methods (see Fig. 1). To study the high-temperature regime, we focus on the development of the PIMC method [6], which naturally incorporates finite temperature quantum effects by working within the many-body thermal density matrix formalism. The combination of Feynman’s imaginary time path integrals and efficient Monte Carlo sampling techniques makes this approach one of the most appropriate first-principles simulation techniques for quantum systems at high temperature. Since the length of the path scales like 1/T, the method becomes increasingly efficient for high temperatures. Electrons and nuclei are often treated equally as paths, but here we treat the nuclei classically because their zero-point motion is negligible for the temperatures under consideration. All PIMC simulations were performed with our own code, CUPID [7].

For the low-temperature part of the WDM regime, DFT-MD is an accurate and efficient first-principles simulation method for these conditions. The thermal occupation of electronic states is treated as a perturbation of the ground state by Fermi–Dirac smearing. The main drawback of this method is that it becomes computationally infeasible as electrons occupy a large number of bands at high temperature, which is why we switch to PIMC simulation at high temperature (see Fig. 1). All DFT-MD simulations were performed with the VASP code [8].

RESULTS & IMPACT

We performed an entirely first-principles determination of hydrogen mixtures in the WDM regime by including all nonideal effects. Based on PIMC and DFT-MD, we obtained coherent sets of EOSs over a wide range of density and temperature conditions and derived the shock Hugoniot curves of a series of hydrocarbon materials [9]. For polystyrene, we predict a maximum shock compression ratio of 4.7, whereas earlier estimates range from 4.3–4.7. Our calculated shock Hugoniot curve (Fig. 1) agrees very well with experimental measurements and provides guidance for the interpretation of experiments on the Gbar platform at NIF. We observe a single compression maximum for hydrocarbon materials while there are two compression maxima in the Hugoniot curve of nitrogen, oxygen, and neon. We have shown that this difference is related to the properties of the L-shell ionization, which is much more gradual for carbon. We found that the linear isobaric–isothermal mixing approximation works very well, resulting in a discrepancy in the density of CH of 1% or less under stellar core conditions. This implies that it is sufficient to derive only accurate EOS tables for the end members in order to provide a thermodynamic description of deep stellar interiors. In the past, models for stellar and giant planetary interiors relied on semi-analytical models and experimental results for a few key conditions. Through high-performance computation, it has become possible to provide a more rigorous theoretical description of matter at extreme conditions. While, in the past, the characterization of one material has taken a Ph.D. thesis [5], on HPC systems like Blue Waters many materials can now be characterized in parallel within a single year.

WHY BLUE WATERS

Our Blue Waters allocation is one order of magnitude larger than any other allocation that we have obtained elsewhere. This enabled us to perform simulations with more accurate fermion nodes that are more realistic but also much more expensive. Besides the hydrocarbons discussed in this report, we performed simulations for a variety of other materials, including sodium, lithium fluorides, and silicon. These investigations are governed by myriad short-lived chemical bonds that are mitigated by hot electrons. Their density is illustrated by the yellow line.

PUBLICATIONS AND DATA SETS


Figure 2: CH plastic at extreme pressure–temperature conditions studied with atomistic simulations. The C and H atoms are colored blue and pink spheres. Their interactions are governed by myriad short-lived chemical bonds that are mitigated by hot electrons. Their density is illustrated by the yellow line.
DIRECT NUMERICAL SIMULATIONS OF THE RELATIVE MOTION OF HIGH-INAERTIA PARTICLES IN ISOTROPIC TURBULENCE

Allocation: NSF PRAC/2017 Koh
PI: Neama L. Rani
Collaborator: Rohit Dhurawal

University of Alabama in Huntsville

EXECUTIVE SUMMARY

The overall objective of our research is to investigate the role of turbulence in driving the relative velocities and positions of inertial particles in isotropic turbulence. During the second year of this PRAC grant, we investigated the relative motion of high-inertia particle pairs in isotropic turbulence. We performed direct numerical simulations (DNS), as well as Langevin simulations (LS) based on a probability density function (PDF) kinetic model for pair relative motion. We developed a stochastic theory that involved deriving closures in the limit of high Stokes numbers for the diffusivity tensor in the PDF equation for particle pairs. Quantitative analysis of the stochastic theory was performed through a comparison of the particle pair statistics obtained using LS with those from DNS. The high-performance computing (HPC) resources of the Blue Waters system were invaluable and indispensable in performing the DNS and LS runs needed to validate the stochastic theory.

RESEARCH CHALLENGE

Turbulence-driven relative motion of high-inertia particles is relevant in astrophysical scenarios, such as the interstellar medium, protoplanetary disks, and the atmospheres of planets and dwarf stars. Specifically, the “sticking” of dust particles in protoplanetary disks is believed to be the mechanism for planetesimal formation. An intriguing question that astrophysicists are investigating is believed to be the mechanism for planetesimal formation. Specifically, the “sticking” of dust particles in protoplanetary disks, and the atmospheres of planets and dwarf stars. Recently, we developed a stochastic theory for the relative motion of inertial particles in a turbulent flow. During the second year of this PRAC grant, we investigated the relative motion of high-inertia particle pairs in isotropic turbulence. During the second year of this PRAC grant, we investigated the relative motion of high-inertia particle pairs in isotropic turbulence. The theory involves deriving closures in the limit of high Stokes numbers for the diffusivity tensor in the PDF equation for particle pairs. The current DNS study, as well as the stochastic theory, proved to be an invaluable resource in computing key inputs to evaluating the nonlinear convective terms in the Navier–Stokes equations is extremely computationally intensive. Hence, a pseudospectral approach is adopted wherein the nonlinear terms are first computed in physical space and then transformed into the spectral space. The P3DFFT library [2] is used to carry out the transforms between physical and spectral spaces.

RESULTS & IMPACT

The RDF is a well-established measure of particle clustering. In fig. 1, the RDF is presented as a function of Sτ at four separations τ = 6, 12, 18, and 24 (τ is the Kolmogorov length scale). The results from LS are compared with the data from the DNS performed in the current study, the Février, et al. [3] DNS, and also with the results from the Zaichik and Alipchenkov [4] theory. The Février, et al. [3] data were for Reτ = 69, while the current DNS data are for Reτ = 76. There is excellent agreement between the LS RDF and the two sets of DNS RDFS at all four separations, particularly for Sτ > 10. The Zaichik and Alipchenkov [4] theory significantly overpredicts the RDFS for high Stokes numbers at all separations. The current DNS study, as well as the stochastic theory, provided the basis for the first-ever investigation of the validity of the Zaichik and Alipchenkov [4] theory in the high-Stokes-number limit.

WHY BLUE WATERS

Direct numerical simulation is the most accurate numerical approach to resolve all the temporal and length scales in a turbulent flow. However, DNS of particle-laden turbulent flows are computationally very intensive, since in DNS the cost of a simulation scales as Re3, where Re is the Reynolds number. The overall computational objective of this project is to simulate particle-laden isotropic turbulence at Taylor micro-scale Reynolds number Reτ = 600 with grid sizes ~ 20483. This will require running our code on tens of thousands of cores. Also, each DNS run is expected to generate several terabytes of data. Due to these central processing unit time and storage requirements, the Blue Waters supercomputer is the ideal platform to achieve our objective. It would be relevant to mention that during the last year, Blue Waters proved to be an invaluable resource in computing key inputs to our stochastic theory. For instance, the theory requires as an input the two-time Eulerian correlations of fluid relative velocities seen by particle pairs. Evaluation of the two-time correlation for nearly half a trillion pairs is a highly computationally intensive process. We were only able to compute this quantity because of the Blue Waters access, where we ran the code on 625 nodes using 20,000 cores.

PUBLICATIONS AND DATA SETS


Figure 1: Radial distribution function as a function of Sτ/η at Reτ = 69, while the current DNS data are for Reτ = 76. The two sets of DNS RDFS at all four separations, particularly for Sτ > 10. The Zaichik and Alipchenkov [4] theory significantly overpredicts the RDFS for high Stokes numbers at all separations. The current DNS study, as well as the stochastic theory, provided the basis for the first-ever investigation of the validity of the Zaichik and Alipchenkov [4] theory in the high-Stokes-number limit.
**EXECUTIVE SUMMARY**

The COMPASS experiment at CERN probes proton substructure by scattering high-energy pion and muon beams off of nuclear targets to measure the momentum and coordinate phase space of quarks inside the proton. Observing correlations between proton spin and the intrinsic transverse momentum of quarks will shed light on the quark dynamics inside the proton and will provide a critical test of fundamental predictions derived from quantum chromo dynamics, the quantum field theory describing the nuclear force. The measurements will produce 10 petabytes of experimental and simulated data. Blue Waters’ balance of processing capabilities with data storage and handling is well suited for the analysis of the large COMPASS data samples as these require significant algorithmic processing per pion/muon-proton scattering event. In addition to raw data processing and physics-level analysis, Blue Waters allows for detailed simulation of COMPASS detector properties and environmental effects.

**RESEARCH CHALLENGE**

Observation of the sign change of the Sivers quark distributions (“Sivers functions”) in Drell–Yan scattering compared to existing “Sivers functions” arise from correlations between proton spin and quark transverse momentum and thus appear connected to quark orbital motion inside the proton. The experimental measurement of these correlations will provide the first tomographic images of quarks in the proton in transverse momentum space (Fig. 1). With the 2016 and 2017 runs, COMPASS will add valuable observables in Deeply Virtual Compton Scattering (DVCS) to constrain spin-independent Generalized Parton Distributions in the so far unexplored kinematic domain between HERMES and the Jefferson Lab experiments on the one side, and the HERA experiments on the other side.

**METHODS & CODES**

For experimental data production, the raw COMPASS data collected at CERN are transferred to Blue Waters at average throughput speed of about 1 GB/s using the File Transfer System FTSM [3], a bulk data mover created to distribute globally the throughput speed of about 1 GB/s using the File Transfer System (FTS) [3], a bulk data mover created to distribute globally the Large Hadron Collider (LHC) data. For each triggered event in COMPASS, the information of the detectors is recorded by the Monte Carlo generator PYTHIA [6] adapted to pion and kaon beams versus the transverse momentum of the Drell–Yan particle pair. Different beam energies are indicated by different colors. Correspondence analysis Library (CORAL) software performs the transition from raw data to physical quantities. CORAL’s function is to reconstruct particle trajectories and momenta, and the position of vertices. The reconstructed information is stored in the form of Data Summary Trees, which are read and analyzed using the COMPASS Physics Analysis Software Tools (PHAST). The production of Monte Carlo data is performed in three steps: (1) The generation of signal and background events is carried out with event-generator packages. (2) For the simulation of the detector response to the physics event, a GEANT [4] toolkit is used based on the description of the background events is carried out with event-generator packages. (2) For the simulation of the detector response to the physics event, a GEANT [4] toolkit is used based on the description of the COMPASS apparatus. (3) Simulated hits are subjected to the same reconstruction CORAL and PHAST codes as experimental data.

**RESULTS & IMPACT**

COMPASS accumulates a raw experimental data set of about 0.75 petabytes per year. A first step in the data analysis is the conversion of raw data into the physical properties of the fundamental particles created in a collision event. This data production is an iterative process that requires two or three passes. Approximately 6.5 million CPU hours are needed for one data production pass, which requires usually about three calendar months given the available resources at CERN. In an exploratory mass production in 2016, 5% of a yearly COMPASS data set was processed in 15 hours [5]. Since the task is scalable, a full year can be processed in about the same time if sufficient computing resources are available. This project has recently been granted millions of node hours. As of the middle of 2017, the launching of full-scale data production on Blue Waters was in preparation. Apart from the processed experimental data, simulated Monte Carlo data are an essential ingredient to the data analysis. Simulations of the detectors play a central role in understanding subtle detector effects and removing background events from the data sample. With the available resources at CERN and collaborating institutions, the CPU-intensive part of the Monte Carlo—the simulation of the detector properties with GEANT—can often not be afforded for extensive studies, for example event pile-ups or detector efficiencies in a fine time binning. The production of simulated data for studies on a future fixed-target experiment at CERN using radio-frequency separated pion, kaon, and anti-proton beams has begun (Fig. 2). The Blue Waters project involves students and young postdocs and it will in the future attract more young physicists. It thus offers outstanding educational potential for a significant number of students and postdocs towards building a community capable of using petascale computing.

**WHY BLUE WATERS**

Requiring both substantial Monte Carlo data production and up to three data campaigns in parallel, a timely analysis of COMPASS data appears difficult. A delay of several years between the end of data-taking in November 2016 and publication of all COMPASS results appears likely. Given the present computing resources at CERN and collaborating institutions, COMPASS might lose out to competing experiments scheduled to run in the future at FNAL, RHIC, and Jefferson Laboratory. The current projected schedule for data analysis also provides challenges to sustaining support from funding agencies for extended periods of time and for on-schedule completion of doctoral theses based on COMPASS data. With the petascale resources of Blue Waters, COMPASS experimental and Monte Carlo data can be processed significantly faster, and in the case of simulated data, also can be generated in greater detail.
OPTICAL DETERMINATION OF CRYSTAL PHASE IN SEMICONDUCTOR NANOCRYSTALS

Allocation: Blue Waters Professor/190 Knh; Illinois/550 Knh
PI: André Schleife
Collaborators: Sung Jun Lim, Andrew Smith
1University of Illinois at Urbana-Champaign
2Daejeon Gyeongbuk Institute of Science and Technology

EXECUTIVE SUMMARY

Semiconductor nanocrystals are highly interesting and diverse material systems with unique electronic and optical properties that are even tunable as to size, shape, composition, and internal structure. Currently, the fundamental connection between crystal structure and optical properties is not well understood, hampering wide-spread and efficient use of these nanocrystals in scientifically important applications such as light-emitting diodes, solar cells, bioimaging, and consumer electronics. In order to establish a connection between structure and optical properties, we use a combination of cutting-edge experimentation and first-principles theoretical spectroscopy. Our work shows that by understanding the absorption spectrum at energies well above the absorption onset, it is possible to efficiently distinguish optically between two different crystal structures of CdSe materials, wurtzite and zincblende, both in bulk crystals and nanocrystals. This may rapidly accelerate development of these materials by allowing fast and accurate structural characterization with small quantities as well as samples in liquids.

RESEARCH CHALLENGE

Semiconductor nanocrystals are diverse material systems currently used in light-emitting devices, solar cells, bioimaging, and consumer electronics. Their unique electronic and optical properties and the fact that these properties are tunable by changing the size, shape, composition, and internal structure of the nanocrystals, make them highly interesting material systems to a broad community of researchers and diverse commercial sectors from optical devices to medical diagnostics. Among the most fundamental questions is the connection between crystal structure and optical properties, as answering that question would allow control of functionality through structural properties that are accessible to synthesis. Despite highly interesting implications for materials design of semiconductor nanocrystals, this connection is not fully understood today, hampering a more widespread and more efficient use of semiconductor nanocrystals for many of the above-mentioned applications. Establishing such a connection is extremely challenging since, in experiment, clarifying the crystal structure typically relies on X-ray crystallography. This technique, however, is not well-suited for samples in solution, for process analysis, for micro-scale reactions, or for high-throughput synthesis. In addition, X-ray diffraction patterns are less clear for very small or mixed-crystal nanocrystals. Application of fast and reliable optical techniques is highly desirable but is, so far, impossible since spectral features in the vicinity of the optical absorption onset are too similar across different crystal structures to allow reliable distinction.

METHODS & CODES

In order to extend the spectral range of interest beyond the absorption onset, we use a combination of cutting-edge experimentation and first-principles theoretical spectroscopy. This allows us to clarify important structure–property relationships for CdSe nanocrystals in two different crystal structures, wurtzite and zincblende. Here, we focus on the computational aspects of the project: We use many-body perturbation theory to explore the influence of quasiparticle effects on optical properties of CdSe. More specifically, single–quasiparticle effects are described using the HSE06 hybrid exchange–correlation functional. From the solution of the Bethe–Salpeter equation for the optical polarization function we conclude that excitonic effects are small in bulk CdSe due to the large screening of the electron–electron interaction. This allows us to resort to a computationally more affordable density–functional theory description to extend this study toward optical properties of much more intricate hydrogen-passivated CdSe nanocrystals in both crystal structures. We use density functional theory to study both crystal structures and mixed-crystal nanocrystals where the wurtzite and zinc-blende phase are separated by a stacking fault. All calculations described here are carried out using the Vienna Ab Initio Simulation Package [1–3] as well as the Bethe–Salpeter equation implementation discussed in [4] and [5].

RESULTS & IMPACT

We report the experimental and computational identification of unambiguous optical signatures of cubic and hexagonal phases in II–VI CdSe nanocrystals: The main result of this work is that it is possible to efficiently distinguish between wurtzite and zincblende CdSe, both in bulk and nanocrystal form, exclusively using optical absorption spectra. Both computation and cutting-edge experimentation clearly show that it is key not only to focus on the absorption onset: While the onsets look very similar between both polymorphs, our work shows that high-energy spectral features in the 4–6 eV energy range are suitable to rapidly identify phase even in nanocrystals as small as about 2 nm. We found that this not only allows distinguishing the crystal phase of bulk CdSe but even to semi-quantitatively analyze wurtzite and zincblende contributions to nanocrystals that contain both structure types, separated by a stacking fault. Furthermore, the first-principles calculations carried out on Blue Waters explain the band-structure origin of these spectral features for both crystal structures. Our results pave the way for applying accelerated experimentation, possibly even high-throughput experiments, to nanocrystal samples in solution. Experiments carried out for this work additionally show that important ligands studied here are not affecting these conclusions.

WHY BLUE WATERS

Computing converged optical–absorption spectra of semiconductor nanocrystals is a computationally expensive task: Each nanocrystal consists of about 300 atoms, including passivating hydrogen atoms, and needs to be surrounded by a sufficiently large region of vacuum. In addition to densely sampling the Brillouin zone, a large number of valence and conduction bands needs to be included in order to achieve convergence of the computed optical spectra. These requirements lead to large matrices for the Kohn–Sham wave functions and solving the Kohn–Sham equations for such large systems can only be achieved using well-parallelized codes on fast supercomputers with fast interconnects. Such calculations are not a use case for the hardware infrastructures provided by cloud computing and, hence, are not possible on such high-throughput resources. Instead, these calculations are exactly what Blue Waters is developed for: We benefited from the fast CPUs, the large amount of memory per compute core, and extremely fast interconnects between nodes. Continuous access to such machines is absolutely critical for our research and cannot be accomplished by cloud services in the foreseeable future. Furthermore, Blue Water staff, e.g., Victor Anisimov, helped us identify and fix performance bottlenecks when writing and reading large wave function files, as was required especially in the early stages of this project. Only this combination of hardware and staff resources enabled the computational part of this research project.

PUBLICATIONS AND DATA SETS

Jun Lim, S., A. Schleife, and A. M. Smith, Optical determination of crystal phase in semiconductor nanocrystals. *Nat. Commun.*, 8 (2017), DOI: 10.1038/ncomms14849.
ADVANCED DIGITAL TECHNOLOGY FOR MATERIALS AND MANUFACTURING

Allocation: Exploratory/50 Koh
PI: Ahmed Tahb
Co-PIs: Soil Kori, Sudharkar Parimi-Ghantam, Nazmoon Akura, Gabrielle Allen
Collaborators: Ashraf Ildiaziok, Shantamu Shahans, Ferehshtoh S. Sabet, Ethan Shapera

1National Center for Supercomputing Applications
2Indiana University

EXECUTIVE SUMMARY

This exploratory Blue Waters (BW) proposal provided computing resources to four graduate students funded by the National Center for Supercomputing Applications Materials & Manufacturing group, allowing the students to explore how their research can be furthered through the use of high-performance computing to address large-scale problem-solving. Two of the research activities that used the BW allocation, both involving analysis of bone, are reported here.

The first project, titled “Simulation of Reference Point Indentation on Cortical Bone,” was conducted by Ashraf Ildiaziok. He used two different instruments—BioDent and Osteoprobe—that utilize the Reference Point Indentation (RPI) technique. The second project, “Mechanics of materials with focus on accelerated design and structure-processing-property relations of materials via high scale computations” was conducted by Fereshteh A. Sabet. This project examined and compared the performance of implicit and explicit solvers for modeling trabecular bone using Abaqus.

Project 1 - Simulation of Reference Point Indentation on Cortical Bone

RESEARCH CHALLENGE

Bone has a hierarchical architecture ranging from atomistic to macroscopic scales. At the scale of one to a few millimeters, the bone tissue is composed of cortical and trabecular bone. Osteoporosis is a bone disease characterized by low bone density, which leads to an increased risk of fractures that occur mainly in trabecular bone. Trabecular bone is also the primary site for insertion of orthopedic implant systems. Thus, the mechanical properties of trabecular bone are of great clinical and research interest for prediction of age and disease-related fractures as well as for designing improved implant systems.

METHODS & CODES

Modeling of trabecular bone entails highly nonlinear mechanical behavior along with contacts. As a result, it is of considerable interest to assess the effectiveness and efficiency of explicit solution methods. In this project, we used the implicit and explicit solvers of Abaqus to analyze micro-computed tomography (micro-CT) finite element (FE) models of trabecular bone, and compared the performance of the two solvers.

RESULTS & IMPACT

Our results show that there is a good match between micro-CT FE model results using implicit and explicit solvers (see Fig. 1). We also observed that implicit and explicit solvers scale similarly, but the explicit solver performs five times faster.

WHY BLUE WATERS

We were able to successfully scale our simulations on eight to 12 nodes on Blue Waters with the explicit solver, which significantly saved computational time. Each of our models has many millions of degrees of freedom and nonlinearities, making such models impossible to solve without the use of a supercomputer.

Project 2 - Mechanics of materials with focus on accelerated design and structure–processing–property relations of materials via high-scale computations

RESEARCH CHALLENGE

Osteoporosis is responsible for two million broken bones at a cost of $19 billion, annually in the United States. According to the National Osteoporosis Foundation, by 2025 this disease is expected to lead to three million fractures at a cost of $25.3 billion per year. Assessing bone material properties in relation to its fracture resistance is important for the diagnosis and treatment of bone diseases. Using traditional materials testing approaches to measure the mechanical properties of bone, such as compression, tension, and three-point and four-point bending, are ex vivo and destructive.

METHODS & CODES

Cortical bone forms the outer hard shell of the whole bone. Therefore, understanding cortical bone fracture behavior is essential to evaluate fracture resistance of the complete bone. The Reference Point Indentation (RPI) technique was invented to allow in vivo evaluation of bone properties. Two different instruments use the RPI technique: BioDent and Osteoprobe. BioDent applies multiple indents at the same location on cortical bone, whereas Osteoprobe applies only one loading cycle at multiple neighboring locations on cortical bone. The relationship between RPI and bone properties has not been developed and is still an open topic.

In our research, we are focused on numerically relating both BioDent and Osteoprobe RPI instrument outputs to actual bone material mechanical properties. The cortical bone RPI simulation problem is highly nonlinear, where geometric nonlinearity, material nonlinearity, and contacts must be accounted for in order to preserve the accuracy of simulation results. We use Abaqus software to simulate cortical bone RPI.

RESULTS & IMPACT

Relating each of the 10 outputs of the BioDent RPI instrument to the bone material properties has been established using the finite element method [1]. Also, simulating bone fracture using the extended finite element method on a single-osteon cortical bone sample has been evaluated and published [2]. A study to relate Osteoprobe RPI output to bone material properties and its fracture resistance is currently being developed.

WHY BLUE WATERS

Completing this study is fully dependent on the numerical finite element method. The problem is highly nonlinear, and multiple iterations are needed to be able to relate Osteoprobe device output to different bone mechanical properties. Each Osteoprobe RPI simulation iteration demands high computational power and time. Therefore, completing such study using the multi-core BW system is essential.

PUBLICATIONS AND DATA SETS


Figure 1: Project 1—(a) Example of von Mises stress distribution in samples under uniaxial compression, (b) implicit and explicit solver results versus experimental stress–strain curves.

Figure 2: Project 2—Mechanics of materials with focus on accelerated design and structure–processing–property relations of materials via high-scale computations.
EXECUTIVE SUMMARY

The objective of this project is to develop computational models to simulate transient multiphase flows and related phenomena, to apply them to gain a deeper understanding of defect formation in continuous steel casting, and to find ways to further improve the process. Large-Eddy Simulations (LES) coupled with a Volume of Fluid (VOF) model were applied to track transient motion of the liquid mold flux/molten steel interface and slag entrainment into the molten steel pool during steady state continuous casting. In addition, the transport of argon bubbles in the molten steel and their capture into the solidifying steel shell were simulated using LES simulations coupled with a Lagrangian Discrete Phase Model for particle transport and particle capture criteria. Furthermore, the LES models were validated with facility measurements and applied to investigate optimal process conditions for the nozzle port angle, submergence depth of the nozzle, and Electro-Magnetic Braking (EMBr) field strength. These simulations on Blue Waters reveal deeper insights into defect formation during the continuous casting of steel and have enabled improved operations.

RESEARCH CHALLENGE

Continuous casting is the most widely employed solidification-process for steel manufacturing in the world [1], so even small improvements in this important process can lead to large benefits. Most defects in final products are related to transient multiphase flow phenomena in the mold region of the process (Fig. 1(a)). Severe instability at the liquid mold flux/molten steel interface can entrain some of the liquid mold flux (added on top of the molten steel pool in the mold to prevent steel oxidation) into the molten steel [2]. In addition, argon gas bubbles, injected to prevent nozzle clogging [3], can be trapped by the solidifying steel shell in the mold. To reduce these problems, transient multiphase flow phenomena should be understood, and process conditions should be optimized to reduce defect formation during continuous casting.

In this year’s project, LES of several different important aspects of multiphase flow were performed to quantify the transient liquid mold flux/molten steel interface, transport of argon gas bubbles, and bubble capture into the steel shell in the mold during nominally steady continuous casting of steel slabs for different process conditions. The modeling results have been validated with plant measurements and applied to find optimal process conditions, including nozzle port angle, nozzle submergence depth, and EMBr field strength.

METHODS & CODES

LES coupled with VOF were applied to model transient molten steel flow and to track the liquid mold flux/molten steel interface. These models were implemented into the commercial package ANSYS Fluent on Blue Waters (BW) XE nodes. To calculate bubble transport and capture into solidifying steel shells with and without EMBr, LES coupled with Lagrangian particle capture (based on a force balance on each particle at the solidification front) [5] and MagnetohydroDynamics models [8] using the GPU-based in-house code CUFFLOW were employed on BW XE nodes.

RESULTS & IMPACT

Turbulent swirl flow from the upward-angled nozzle ports produces jet wobbling in the mold [4]. Sometimes, the jet impinges onto the liquid mold flux/molten steel interface and drags some of the liquid mold flux into the steel pool, resulting in slag entrainment, as shown in Fig. 1(b). Most of the entrained slag becomes entrapped into the solidifying steel shell to form defects. Calculated interface instability reveals level variations greater than ~20 mm, especially at the meniscus region around its perimeter (Fig. 1(c)). These severe level instabilities can cause the liquid mold flux to touch the solidifying steel shell, and to be captured into the steel shell via meniscus hooks. However, the jet flow from the downward-angled nozzle ports with well-optimized casting conditions makes a classic double-roll pattern with less jet wobbling, resulting in better stability of the surface level and velocity in the mold. Velocity variations are smaller and the interface shows only ~2 mm fluctuations, so slag entrainment defects are drastically reduced.

Argon bubble motion is affected by turbulent jet flow in the mold, as shown in Fig. 2(a). Most bubbles larger than ~3-mm in diameter float up toward the top surface due to their large buoyancy. On the other hand, small bubbles move along with the jet flow and easily reach the narrow face to be carried deep into the mold cavity. Many of the small bubbles move below the dendrites to be captured into the steel, especially without EMBr. However, EMBr slows and deflects the jet flow [6–7], so more bubbles float upward near the nozzle and fewer bubbles are transported to the narrow face and deep into the mold. As shown in Fig. 2(b), the bubble capture rate increases with time (until ~15–16 sec after gas injection for EMBr off and ~15 sec for EMBr on, because the jet flow path is shorter with EMBr). Then, after the flow has stabilized at nominally steady casting, EMBr is observed to reduce bubble entrainment significantly.

Parametric studies with these multiphase LES models have enabled better understanding of the complex multiphysics phenomena related to defect formation, including slag entrainment, inclusion and bubble transport, and capture. This has led to suggestions of nozzle geometry/casting condition combinations that lead to fewer defects, and, consequently, to significant savings to the steel plants.

WHY BLUE WATERS

Blue Waters enabled high-resolution multiphase flow simulations of the continuous caster needed for accurate predictions. In particular, the transient transport of small volume secondary-phases (entrained liquid mold slag and argon bubbles) in the turbulent flow require very small cells (~1 mm3) in a huge domain, and simulations of over 50 seconds flow time (with >700-second time steps). Furthermore, Blue Waters resources (both ANSYS Fluent HPC on XE nodes and our in-house multi-GPU-CUFFLOW on XE nodes) showed speed-up breakthroughs (e.g., over 3000x with ANSYS Fluent HPC on BW) needed to provide this modeling capability for the steel continuous casting process.

PUBLICATIONS AND DATA SETS


Figure 1: Schematic of multiphase flow phenomena in the steel continuous casting mold and Effect of nozzle port angle on (a) mold flow patterns, including motion of the top liquid mold flux layer, and (b) level variations at the liquid mold flux/molten steel interface in the meniscus region.
EXECUTIVE SUMMARY

Aquatic vegetation provides a wide range of services to the ecosystem: improving water quality through nutrient uptake and oxygen production, providing a wide range of services from increasing biodiversity. While vegetation–flow interactions have been studied extensively for unidirectional flows, much less is known about oscillatory conditions. The current study is geared toward understanding the interactions among vegetation, flow, and sediment under oscillatory flows. Direct Numerical Simulations (DNS) and Large-Eddy Simulations (LES) through different arrays of idealized vegetation, represented as cylinders, are conducted using the higher-order spectral element-based computational fluid dynamics (CFD) solver Nek5000. Different arrangements and numbers of cylinders have been simulated in 2D and 3D, with the largest simulation having ~296 million computational points to model the whole domain is near 1.2 billion. While such simulations are still tractable on a petascale platform like Blue Waters, the computational cost is high, resulting in a reduction of the number of cases one can run, thus constraining the insights a broader range of parameters could yield. To increase variable space, a wide range of conditions are first simulated in 2D for the whole domain (~4 million computational points) to get an overview of the effect of different parameters such as Reynolds number, vegetation density, period and amplitude of oscillatory flow, and spatial heterogeneity. Once the effects of different parameters are well understood, 3D simulations will be conducted for partial domains, big enough to accurately capture the general dynamics, but within manageable computational costs. A better understanding of vegetation–flow dynamics will not only advance fundamental knowledge of physical processes but also will guide design efforts for scour protection and artificial wetlands. Part of the study is to characterize drag coefficients of the vegetation array, improving accuracy of reduced-order models of flow through vegetation. The study will also identify the ideal quantity and locations to place instrumentation in experiments through large arrays of cylinders.

METHODS & CODES

High-resolution LES and DNS of the flow at different configurations of the idealized vegetation were conducted using the open-source, spectral-element-based higher-order incompressible Navier–Stokes solver Nek5000 [5]. The Spectral Element Method combines the accuracy of spectral methods and the flexibility of Finite Elements Method [7]. In the planned simulations with sediment transport, sediment would be modeled as Lagrangian particles using a novel semi-implicit time-stepping scheme developed to simulate polydisperse sediment accurately.

RESULTS & IMPACT

2D simulations have been conducted for the full domain on different configurations. Two cases having the same vegetation density and Reynolds number, but different array configurations, are presented here. The velocity magnitude, along with the pressure field, is shown in Fig. 1. For the staggered case, in contrast with the random array, a vortex being shed from a cylinder is impeded by the ones behind it. The vortex being shed from a cylinder is impeded by the ones behind it, which is not the case for the random case. Results from 3D simulations of turbulent flow through a random arrangement of cylinders, about 296 million computational points are used. Instantaneous velocity magnitude at different elevations. Plots at (a) 0.5 %, (b) 1 %, (c) 10 %, (d) 50 %, (e) 75 %, and (f) 95 % water depth.

WHY BLUE WATERS

The study pushes the limit of the scale at which high-resolution simulations are used to study complex multi-phase flow in environmental fluid mechanics, requiring computational resources with sustained computing power at an unprecedented scale, such as Blue Waters. Simulations have been conducted for up to 296 million computational points, with the code scaling strongly up to 32,768 MPI ranks. Without access to petascale HPC like Blue Waters, completing the study within a realistic timeframe would be impossible. In addition, since visualization of a phenomenon is an effective way to understand and explain its mechanics, we will work with Blue Waters project staff to create animations of the phenomenon using data from the simulations.
QUANTUM MONTE CARLO SIMULATIONS OF MAGNETISM AND MODELS IN CONDENSED MATTER

Allocation: Illinois/202.8 Knh
P.K. Lusae K. Wagner
Collaborators: Elf Ertekine, Nurcan Alun

1University of Illinois at Urbana-Champaign

EXECUTIVE SUMMARY

This project uses quantum Monte Carlo techniques to study the behavior of electrons in materials where traditional methods of calculation are too inaccurate to be of use. These large-scale calculations will provide unprecedented detail about the correlated behavior of electrons in these materials at the subatomic scale. The simulations will then be analyzed in a systematic way to extract a compressed description of their behavior, which will help us to understand the difference at a microscopic level between a simple magnetic material and unconventional superconductors.

RESULTS & IMPACT

With Blue Waters (BW), it is now possible to achieve unprecedented accuracy in solving the Schrödinger equation using Monte Carlo techniques. Our group has used BW to show that the simulations represent a step forward in the state of the art in computing many particle quantum systems. We have been able to describe the properties of water interacting with graphene, defects in semiconductors, the properties of superconducting materials, and transitions between metallic and insulating behavior. The figure shows work performed on BW on vanadium dioxide (150°F). This transition can also be induced by strain (3) or electric field. For many years, it has been a mystery why this happens, mainly because electron correlation is important in the effect, which is difficult to access with inaccurate quantum calculations. By using BW to perform quantum Monte Carlo calculations of the electrons in this material, we were able to describe the transition from first principles, meaning that we used only the positions of the atoms to compute its properties. This new predictive capability opens the door to computer-based design of these extraordinarily sensitive materials that can be used in new devices.

WHY BLUE WATERS

The BW system provides excellent technical support and access to high-performance computing with a minimal amount of effort required from scientists, so that we can focus on the science. As one of the few National Science Foundation-supported resources that provides large allocations, it has been critical to the success of the projects listed here.

PUBLICATIONS AND DATA SETS


INTERMITTENCY, RESOLUTION EFFECTS AND HIGH SCHMIDT NUMBER MIXING IN TURBULENCE

Allocation: NSF PRAC/9,547,165 Khb
PI: P.K. Yeung
Collaborator: T. Gotoh

Georgia Institute of Technology
Nagoya Institute of Technology, Japan

EXECUTIVE SUMMARY

Substantial advances have been made in both science and computing for fluid turbulence using Blue Waters. Simulations on grids of up to 4 trillion points show clear differences between energy dissipation rates averaged locally in three versus one dimension (the latter being common in the literature). Resolution effects on extreme events are examined critically using a multi-resolution approach. We have also developed a new algorithm for turbulent mixing at low diffusivity, where scalar fluctuations arise at scales much smaller than in the velocity field. We use a dual communicator approach where different groups of MPI processes compute the velocity and scalar field at different resolutions and using different numerical schemes. Through careful use of inter-communicator communication, as well as multithreading via nested OpenMP constructs, the code scales well up to 524,288 cores at close to six percent of theoretical peak performance.

RESEARCH CHALLENGE

Turbulent flows with disorderly fluctuations over a wide range of scales are an important agent of efficient mixing in many fields of science and engineering. An enduring challenge in the theory of this subject is to understand the nature of intermittency [1] in terms of fluctuations of the energy dissipation rate (a measure of local straining) and enstrophy (a measure of local rotation) over a wide range of domain sizes. Direct numerical simulations of the scale enabled by Blue Waters are the best source of data for this purpose, but the complexity of the flow physics requires a critical examination of the accuracy and reliability of the results. Accordingly, in addition to analyzing data from a 0.5-trillion-grid-points simulation [2] we have also performed a short simulation using as many as 4 trillion grid points. Massive datasets at this level allow us to investigate the effects of differences in the local averaging procedure as well as the effects of finite resolution.

As a second focus area is the development and application of a new parallel algorithm that is uniquely suited to the study of turbulent mixing at high Schmidt number, where low molecular diffusivity leads to fluctuations at scales smaller than those in the velocity field. Our objective includes checking a scaling relation proposed in classical theory for which confirmation via either simulation or experiment has been limited by grid resolution and/or Schmidt number. Our new simulation uses a new dual-resolution parallel paradigm to make the calculation practical.

METHODS & CODES

Our technical approach is direct numerical simulation (DNS), which is based on exact physical laws and can be carried out efficiently in simplified geometries. We use Fourier pseudo-spectral methods for the velocity field, and perform local averaging both along a line (one dimension) and over a cube (three dimensions). To study resolution effects via filtering we apply successive truncations in wavenumber space before transforming to physical space coordinates. Results that differ greatly after truncation may indicate substantial errors.

RESULTS & IMPACT

For high Schmidt number mixing we retain pseudo-spectral methods for the advecting velocity field but use combined compact finite differences (CCD) on a finer grid to satisfy stringent resolution requirements for the scalar [3]. We have devised a new dual-communicator parallel algorithm [4] where distinct groups of MPI processes parallel works on the velocity and scalar fields at different resolutions. This particular setup is less communication-intensive than full pseudo-spectral codes. Rigorous efforts have been made to improve performance by overlapping communication with computation. The best performance is obtained by dedicating one OpenMP thread to communication and dividing computational work among the other threads using nested OpenMP parallelism.

WHY BLUE WATERS

The 8,192 grid resolution of our production simulations allows access to a world-class machine such as Blue Waters, and the memory capacity of Blue Waters has enabled us to obtain data at 16,384 resolution, the highest known worldwide in the turbulence community. Other important machine characteristics include support for Co-Array Fortran, topologically aware scheduling, and a large number of cores per node, which is conducive to nested OpenMP parallelism in our high-Schmidt-number algorithm.

PUBLICATIONS AND DATA SETS

170 CADENS NSF Project: Digital Literacy, Data Visualization, and the Cinematic Presentation of Science

172 Algorithms for Extreme-Scale Systems

174 Parallelization of the Multilevel Fast Multipole Algorithm (MLFMA) on Heterogeneous CPU-GPU Architectures

176 Predicting Performance Degradation and Failure of Applications through System Activity Monitoring

178 Parallel Algorithms for Solving Large Assignment Problems

180 Localizing Communication in Sparse Matrix Operations

182 Performance Evaluation of New Algebraic Algorithms and Libraries

184 Parallel Algorithms for Big Data Phylogenomics, Proteomics, and Metagenomics

186 Hardware Acceleration of Deep Learning

188 A Critical Evaluation of the OP2/OPS Parallel Meshing and Code Generation Software
CADENS NSF PROJECT: DIGITAL LITERACY, DATA VISUALIZATION, AND THE CINEMATIC PRESENTATION OF SCIENCE

Allocation: Illinois/300 Knh
PI: Donna Cox
Co-PIs: Kalina Borkiewicz, Jeff Carpenter, AJ Christensen, Stuart Levy, Robert Patterson
Collaborators: Nathan Goldhahn, Michael Norman, Leigh Ort, Brian O’Shea

National Center for Supercomputing Applications

EXECUTIVE SUMMARY

The Advanced Visualization Lab (AVL) continues work on the NSF-funded CADENS project (The Centrality of Advanced Digitally ENabled Science, ACI-1445176). The AVL co-produced and rendered visualization scenes for the recently released science documentary, “Seeing the Beginning of Time.” We have used Blue Waters for processing data as well as rendering scenes in 4K monoscopic, stereoscopic, and dome formats. A few technological byproducts of this work include a new pipeline for rendering on Blue Waters called Blurend and an open-source cinematic visualization tool called Ytini.

METHODS & CODES

To create visualizations for the science documentary “Seeing the Beginning of Time,” the AVL primarily used the commercial visual effects tool Houdini and the scientific analysis and visualization Python package yt. We have created middleware and a pipeline that uses both of these tools, called Ytini. We have run Ytini on Blue Waters to convert raw data into Houdini-compatible OpenVDB sparse volume data files. Further, we have created a Python tool called Blurend for preparing Houdini scene files for rendering on Blue Waters.

RESULTS & IMPACT

We have rendered the following sequences on Blue Waters:

• “First Light in the Renaissance Simulations” [1] in 4K stereoscopic format
• “First Light in the Renaissance Simulations” [1] in 4K dome format
• “Large Synoptic Survey Telescope Model” [6] in 4K monoscopic format

We processed data on Blue Waters for the following sequences:

• “Cosmic Bubble Bath” [2] for the 4K documentary “Seeing the Beginning of Time”
• “Formation of the Moon” [5], for our work-in-progress dome show

“Seeing the Beginning of Time” is currently available for streaming on Amazon Prime, where it is rated 4.3/5 stars. Additionally, we have created a new pipeline for preparing Houdini scene files for rendering on Blue Waters, resulting in a new software tool called Blurend. We used this tool to render several sequences for “Seeing the Beginning of Time” as well as a tornado simulation [4] in dome format.

Several of our scientific collaborators used Blue Waters to run their simulations and create large datasets residing there. It was easier for both artists and scientists to make their own cinematic scientific visualizations. Files and tutorials are distributed through www.ytini.com.

WHY BLUE WATERS

We are currently using Blue Waters to prepare data for import into the commercial visual effects tool, Houdini, where we can create cinematic treatments of the data. Converting a particular dataset from particles to a Houdini-compatible Adaptive Mesh Refinement-like volumetric format took about four days on our local cluster. It would have been incredibly difficult to iterate at this time scale. Running the same script across nodes on Blue Waters took four hours.

Additionally, we have created a new pipeline for preparing Houdini scene files for rendering on Blue Waters, resulting in a new software tool called Blurend. We used this tool to render several sequences for “Seeing the Beginning of Time” as well as a tornado simulation [4] in dome format.

Figure 1: A visualization of the Renaissance simulations [2] done by the Advanced Visualization Lab. The universe 400 million years after the Big Bang.

Figure 2: A dome rendering of a tornado simulation [4], done by the Advanced Visualization Lab.

Figure 3: A visual rendering of the Renaissance simulations [2] done by the Advanced Visualization Lab. The universe 400 million years after the Big Bang.

Figure 2: A dome rendering of a tornado simulation [4], done by the Advanced Visualization Lab.
EXECUTIVE SUMMARY

Continued increases in the performance of large-scale systems will come from greater parallelism at all levels. At the node level, we see both the increasing number of cores per processor and the use of large numbers of specialized computing elements in GPGPUs (general-purpose graphics processing units). The largest systems must network tens of thousands of nodes together to achieve the performance required for the most challenging computations.

Successfully using these systems requires new algorithms. Over the last year, we have explored two ways to reduce communication costs in large-scale systems. One is the redesign of algorithms to take into account the difference between on-node and off-node communication performance. The other is to look at the effect of performance irregularity and the use of nonblocking collectives to improve performance of algorithms that use MPI (message passing interface) collectives. We also developed an improved communication model that better matches the performance of modern parallel processors.

RESEARCH CHALLENGE

At extreme scale, even small inefficiencies can cascade to limit the overall efficiency of an application. New algorithms and programming approaches are needed to address barriers to sustained performance.

This work directly targets current barriers to effective use of extreme scale systems by applications. For example, Krylov methods such as Conjugate Gradient are used in many applications currently being run on Blue Waters (MILC is one well-known example) and other leadership-class systems. Developing and demonstrating a more scalable version of this algorithm would immediately benefit those applications. Also of importance to many computations, including Krylov methods for solving large systems of linear equations as well as methods for large-scale graph computations, are sparse-matrix vector multipliers. These involve significant communication between nodes and can lead to scalability limits by improving methods to exploit intermode and intranode communication, many applications can improve their scalability. In the longer term, the techniques that are developed in this project will provide guidance for the development of highly scalable applications.

METHODS & CODES

To address the challenges of parallelism and scale, we developed several codes that allow us to benchmark the performance of these operations, gather detailed timing results, and perform experiments with different approaches. For example, we have been developing a “noise injector” to allow us to better experiment with different amounts of performance variation in multicore nodes. We have also developed a set of benchmark codes that better measure the achievable communication performance of the communication patterns commonly used in applications.

RESULTS & IMPACT

Early results with alternative Krylov formulations have revealed several performance effects that can provide a factor of two or more improvement in performance at scale [1]. We have been using Blue Waters over the last year for an investigation into the impact of large-scale system performance variation on parallel numeric algorithms. This includes developing code for measuring and processing network performance counters, injecting network noise into nodes running another algorithm, and kernels with a variety of common communication patterns. Our initial experiments have used smaller core counts to assist in developing the code and improving the experiments in preparation for a more detailed study in the coming months involving large runs. The goal of this study is to better understand what network noise looks like on supercomputers and how we can develop parallel numeric algorithms that perform well despite noise.

We have also explored the performance models used to guide both the development of algorithms and the analysis of application performance. We discovered that the classic “postal model” is no longer effective for systems with multi-core nodes, and we developed a simple extension to the postal model that explains the observed performance of current systems [2]. This work has also informed the development of an improved approach to sparse matrix–vector multiplication that takes into account the different performance of inter- and intra-node communication [3].

WHY BLUE WATERS

Scalability research relies on the ability to run experiments at large scale, requiring tens of thousands of nodes and hundreds of thousands of processes and cores. Blue Waters provides one of the few environments available for large-scale experiments. In addition, only Blue Waters provides a highly capable I/O system, which we plan to use in developing improved approaches to extreme-scale I/O.

PUBLICATIONS AND DATA SETS


PARALLELIZATION OF THE MULTILEVEL FAST MULTIPOLAR ALGORITHM (MLFMA) ON HETEROGENEOUS CPU-GPU ARCHITECTURES

Allocation: Exploratory: 62.5 Kib
PI: lavender.Gui
Collaborators: Weng Cho Chew, Narsay A. Aluru

1University of Illinois at Urbana-Champaign

EXECUTIVE SUMMARY

The aim of this allocation was to solve unprecedentedly large scattering problems requiring large amounts of memory and to utilize a multitude of GPU’s (graphics processing units) available in Blue Waters. In the scope of this project, we solved a scattering problem with 1.5 billion unknowns on 512XE nodes; this was beyond what we could achieve before using Blue Waters. Our efficient parallelization strategy [1,2] allowed us to spread problems among large numbers of nodes to access more memory, which is required by large problems. Additionally, we achieved the largest full-wave inverse-scattering solutions in near-real-time on 256XE nodes, and it was 4.34 times faster than the same number of XE nodes [3]. This was the first GPU implementation for nonlinear inverse solutions in the literature. We implemented an efficient multi-GPU MLFMA (multilevel fast multiple algorithm) to serve as the kernel of an inverse-scattering solver.

RESEARCH CHALLENGE

We have been working on solutions of extremely large problems that are derived mainly from electromagnetics, acoustics, and optics, to name a few. An electromagnetic scattering problem may become very large as either the frequency increases or the target size gets larger. Simulating the scattering of high-frequency radar waves from a large aircraft, for instance, may require the solution of millions or even billions of unknowns. Similarly, medical imaging with microwaves requires solutions of thousands of large problems. As the problem size (i.e., the number of unknowns) grows, memory requirement increases so we need more nodes, providing access to more memory. MLFMA operates on a multilevel tree structure and is mathematically involved; therefore, it is difficult to implement the parallel algorithm on multiple CPUs. Moreover, it is not easier to implement MLFMA on GPUs because the algorithm is memory-bound due to its O(N) computational complexity. The challenge is to develop efficient parallelization strategies to spread problems evenly among large numbers of nodes to obtain larger solutions, and to exploit the hierarchical memory architecture of GPUs to obtain large computational throughput.

METHODS & CODES

The memory requirement of large problems is a challenge. Prior to Blue Waters, we reached a point where both the available total memory and the single-node memory became insufﬁcient. Therefore, we experimented with out-of-core methods to use the disk storage as additional memory, even though we had to endure a huge penalty in the time required to solve problems [4]. The huge number of computing nodes on Blue Waters makes available a much larger total memory, thus enabling the solution of such enormous problems that were impossible for us to solve before. Additionally, improving parallelization and using more nodes led to much shorter solution times. The hierarchical parallelization strategy allows us to partition the MLFMA data structures among 512XE nodes efﬁciently and evenly (i.e., in a load-balanced way). A scattering problem involving dense linear systems with 1.5 billion unknowns is solved within the 32 TB (terabytes) of memory provided by 512XE nodes.

Unfortunately, MLFMA is a memory-bound algorithm. This is mainly due to the fast nature of MLFMA, where its O(N) complexity prevents data reuse on GPUs [5]. To exploit the hierarchical memory architecture of GPUs, we reformulate the MLFMA operations as matrix-matrix multiplications; this provides good data reuse and high computational throughput. To implement matrix-matrix multiplications efﬁciently, we use a hybrid shared-memory and register tiling algorithm with thread-coarsening methods. For multi-GPU parallelization, the MLFMA tree structure is partitioned among GPUs and the operations among the branches are categorized as intra-GPU and inter-GPU. The inter-GPU data are ﬁrst transferred to CPUs (central processing units), and then MPI (message passing interface) communications are performed to swap the data among computing nodes, and ﬁnally the data are transferred back into GPUs. The MPI communications are optimized to eliminate redundancy and to minimize the communication time. To reduce idle time of CPUs and GPUs, we overlap both MPI communications and CPU–GPU transfers with GPU computations. To do that, we propose a different order than the standard one such that the data-transfer time is completely overlapped with GPU computations. This provides excellent inter-node parallelization of MLFMA [6].

All of our codes are developed in-house with Fortran, C++, and CUDA. The PETSc framework is used in the Fortran code for iterative solutions.

RESULTS & IMPACT

We have increased the number of unknowns in the problems we can solve by achieving the solution of an electromagnetic scattering problem with 1.5 billion unknowns. The ensuing 1,500,000,000 x 1,500,000,000 dense matrix equation is solved with MLFMA and parallelized on 512XE nodes of Blue Waters.

We obtained an efficient multi-GPU MLFMA implementation for volumetric problems for the first time within the scope of this project. A single GPU speedup is 55.14 and 3.97 times with respect to the sequential and 16-core execution baselines. The GPU executions are obtained on XE nodes whereas the GPU executions are obtained on XE nodes. This implies 3.97 times speedup on XE nodes over XE nodes. The speedup of 16 XE nodes is 846.41 and 15.34 times over sequential and a single XE node. This implies 96% parallelization efﬁciency among XE nodes.

Large-scale inverse-scattering solutions on GPUs are obtained on 256XE nodes, where each node is employed with an MLFMA solver. This parallelization scheme decreases an inverse solution from 11.5 hours (sequential execution) to 7 seconds on 128XE nodes. This provides images in near-real-time, allowing our fast and massively-parallel algorithms to be employed for real-life applications. We plan to employ supercomputers (e.g., Blue Waters) for the benefit of complex imaging problems.

WHY BLUE WATERS

Our team is located at the University of Illinois at Urbana-Champaign, and therefore, it is very easy for us to contact the Blue Waters staff. Occasionally, our Ph.D. students visit the NCSC building several times a day to solve their issues on compilers, libraries, and other Blue Waters utilities with the help of the staff. The vast amount of DRAM (dynamic random-access memory) and the number of GPUs available on Blue Waters enable scaling our science to much larger problems. Blue Waters (with its large number of GPU and CPU nodes) is ideally suited for our research. We can also make short-term (e.g., one-hour) reservations for a large number of nodes (e.g., 1,024 nodes) for large benchmark runs [7].

PUBLICATIONS AND DATA SETS


Hidayetoğlu, M., et al., Large-Scale Inverse Scattering Solutions with Parallel Born-Type Fast Solvers. Progress on Electromagnetics Research Symposium (PIERS), Shanghai, China, August 8–11, 2016, pp. 914–920.


PREDICTING PERFORMANCE DEGRADATION AND FAILURE OF APPLICATIONS THROUGH SYSTEM ACTIVITY MONITORING

EXECUTIVE SUMMARY

Our overarching investigation addresses complex data-driven problems associated with online system monitoring for understanding causes of application failures and performance degradations. Our contributions include the design of machine learning-based deep analytics frameworks to distinguish the cause of application and system performance degradation due to failure- and nonfailure-related issues. The tool leverages probabilistic graphical models to conduct machine learning at scale for runtime detection of congestion parameters and its effects on running applications. Our approach correctly identifies the cause of congestion in 74% of the cases (of 302 cases found in our dataset) and attributes it either to resource contention issues or to failures in the system. In addition, working with Sandia National Laboratories and the National Energy Research Scientific Computing Center, we are addressing the issue of congestion in the Cray Aries interconnect to help advance scheduling decisions of applications by porting our algorithms for online system monitoring and mitigation.

RESEARCH CHALLENGE

Extreme-scale high-performance computing (HPC) systems require a holistic approach to monitoring and coordinating many disparate subsystems (both hardware and software) to enable continued scaling and efficient execution of applications. HPC systems are typically used for executing tightly coupled simulation applications across hundreds of thousands to millions of processors. Resource contention due to failures or design issues (of applications/systems) impacts applications in two ways: (1) by degrading application performance, which causes application runtime unpredictability and limits scaling to full system; and (2) through propagation of errors and failures in application logic and code-flow, which causes application failures or invalid/error-prone outputs. As a result, effective failure/degradation mitigation response(s) in complex systems require analysis of the propagation of faults/errors and of performance issues due to interference among applications or resource exhaustion.

Our analysis approach addresses more general scenarios that result in performance degradation in which timely and appropriate response can significantly improve both application runtimes and system throughput. The eventual product of this work will be an interoperable set of capabilities for extreme-scale systems that provides monitoring, analysis, and appropriate response to both resilience and performance issues. These capabilities will support both automated and exploratory analysis, both at runtime and in post-processing. Specifically, the Blue Waters grant allocation was used to characterize performance logs and machine-generated error logs to holistically understand the propagation of performance degradations and faults affecting applications. We use Bayesian network-based machine learning methods to distinguish between failure-based resource contention issues and nonfailure-based contention issues.

METHODS & CODES

Regardless of whether performance problems such as contention for shared resources are failure-related (e.g., caused by link failures) or nonfailure-related (e.g., caused by design issues in an application or system), they can quickly propagate in systems and result in severe performance degradation, rather than application failure as discussed below.

Case Study 1: Fig. 1 (i) shows the impact of network link failure and its corresponding recovery on the Gemini 3D torus interconnection network of Blue Waters. This impact occurs because a single-link failure can cause the traffic pattern of the application to change in the underlying network. Event logs corresponding to this failure are shown in the figure. It also shows the aggregate data passing through all the Gemini routers. Basic statistics such as these indicate bandwidth utilization of the system. To understand whether there is actual performance degradation and, if so, which Gemini router may be the source of the problem, we calculate a derived metric, “average packet latency,” which captures the average time taken by a router to deliver packets.

Case Study 2: Fig. 1 (ii) shows the impact on performance of the system and application due to interplay of design issues between the application (write pattern) and system (scheduling strategy). Specifically, a 32-node job caused high congestion in the system interconnect, triggering two congestion protection events, first at 10:00 a.m. (within 10 seconds of job launch) and then at 15:20 (red lightning bolts). Congestion within the torus can adversely impact the performance of the application generating the messages and of other running applications, and is a major cause of inconsistent application runtimes. This job was assigned a linear shape in the “2” direction by the topology-aware scheduler. Hence, when it was reading large amounts of data (over RDMA, or remote direct memory access), most of the I/O calls were funneled through specific links, causing the congestion.

RESULTS & IMPACT

Our tools take a holistic approach for differentiating design-related performance issues from failure-related issues. This is done by building a Bayesian network-based diagnostic model using features obtained from systems and applications. These features are either obtained directly from measured raw metrics (yellow-filled circles in Fig. 2) or by running unsupervised clustering algorithms on raw metrics (green-filled circles). We tested our design and tools by manually verifying the cause of congestion for 302 reported cases and achieved 74% accuracy.

WHY BLUE WATERS

Blue Waters is one of the few open-science capacity systems that provides a test bed for scaling computations to tens or hundreds of thousands of cores on central processing units (CPUs) and graphics processing units (GPUs). It also enables the study of failures and degradation of applications in production petascale systems because of its unique mix of XE6 and XK7 nodes. This allows us to understand the performance-fault-tolerance continuum in HPC systems by enabling the investigation of application-level designs for mixed CPU and GPU node systems, and fault isolation in system components to mitigate failures at the application level.

PUBLICATIONS AND DATA SETS


PARALLEL ALGORITHMS FOR SOLVING LARGE ASSIGNMENT PROBLEMS

Allocation: Illinois/210 Kish
PI: Raksh Nagi
Co-PI: Ketan Date

EXECUTIVE SUMMARY

The objective of our project is to develop fast and scalable algorithms for obtaining strong lower bounds and exact solutions for large instances of the Quadratic Assignment Problem (QAP), using Graphics Processing Unit (GPU) clusters. The QAP is an NP-Hard problem, in the strong sense. To solve a linearized model of the QAP using branch-and-bound, lower bounds must be calculated using the Lagrangian dual technique, in which a large number of Linear Assignment Problems (LAPs) are solved efficiently, using our GPU-accelerated Hungarian algorithm. Additionally, in a branch-and-bound scheme, a large number of nodes must be explored in order to find a provably optimal solution. To this end, we have used Blue Waters to develop a GPU-accelerated Lagrangian dual ascent heuristic for obtaining lower bounds on the QAP, which is used in the parallel branch-and-bound scheme to solve large QAPs to optimality.

RESEARCH CHALLENGE

The Quadratic Assignment Problem (QAP) was introduced by [1] as a mathematical model to locate indivisible economical activities (such as facilities) on a set of locations so as to minimize a quadratic cost function. Typical applications of the QAP may be found in facility layout (re)design in manufacturing, distribution systems, services (retail outlets, hospital floors, etc.), and electronic circuit design. QAP may also serve as a specialization to many “harder” optimization problems, such as the Graph Association (GA), Traveling Salesman Problem (TSP), Vehicle Routing Problem (VRP), etc., in alternative formulations. Therefore, to solve these problems efficiently, we need to solve the QAP sub-problems efficiently. As a result, a fast and scalable QAP solver can be a powerful tool for researchers working on such NP-hard problems, or problems that are exceeding difficult to computationally solve on a finite resource in a finite time. The sequential QAP solver can become computationally intensive and, therefore, the algorithm can benefit from parallelization on an appropriate parallel architecture, such as Blue Waters.

METHODS & CODES

We choose to parallelize the Lagrangian dual ascent algorithm for Level-2 Refactoring Linearization Technique (RLT2) proposed by [2], in which we need to solve O(n^6) LAPs and adjust O(n^6) Lagrange multipliers to obtain a strong lower bound on the QAP. We designed a parallel Lagrangian dual ascent heuristic for solving RLT2 using hybrid MPI+CUDA architecture. The O(n^6) LAPs are split across these GPUs and solved using our GPU-accelerated Hungarian algorithm [3], while the O(n^6) Lagrange multipliers are updated by multiple CUDA Streams in parallel.

We used this GPU-accelerated dual ascent algorithm in a branch-and-bound scheme to solve QAP instances to optimality. For a node in the search tree, we fix a facility to a location and solve the corresponding RLT2 sub-problem, whose objective value provides a lower bound on the QAP. If this value is greater than the incumbent solution then the node is fathomed; otherwise, it is branched further. Each node is processed using a bank of GPUs. By using multiple such banks, we can process multiple nodes in parallel. The algorithms were tested on the problem instances from the QAPLIB [4].

RESULTS & IMPACT

Lower bounds: The results for lower bounding tests for the various problem instances are summarized in Table 1. With our architecture, we are able to obtain strong lower bounds on problem instances with up to 42 facilities, which is a tremendous achievement.

Scalability study: Although there is a minimum required number of PEs for applying accelerated RLT2 dual ascent to a QAP of specific size, the number of GPUs can be increased and the LAPs can be solved in parallel on multiple GPUs. This allows us to achieve some parallel speedup. We performed strong scalability with 1 to 32 GPUs. We obtain good speedup in the initial stages. However, as we continue to increase the number of GPUs in the system, we get diminishing returns in the execution times, due to increased MPI communication.

Parallel branch-and-bound: The results for the parallel branch-and-bound tests are shown in Table 2. We can see that the number of nodes explored and the completion times increase exponentially with the problem size. The most challenging Nug30 problem instance required more than four days to solve optimally, using 300 GPU banks with 4 GPUs each, which is a significant achievement.

Impact: With our architecture, we are able to obtain strong lower bounds on problem instances with up to 42 facilities, and optimally solve problems with up to 30 facilities, using only a modest number of GPUs. To the best of our knowledge, this is a first-of-its-kind study that has pushed the RLT2 formulation to solving such large problems. With some additional work in memory management and GPU + CPU collaboration, our proposed algorithms may be used effectively to solve truly large QAPs with over 30 facilities, which has been impossible so far.

WHY BLUE WATERS

In a typical branch-and-bound tree, we need to explore a large number of nodes in order to find an optimal solution. As the problem size grows, the number of nodes that need to be explored grows exponentially. Therefore, we need a large number of processors that can explore the solution space in parallel. Additionally, the GPU-accelerated dual ascent procedure benefits from the large number of powerful GPU-enabled processors available at the Blue Waters facility. As seen in the results, we have used over 1,200 XK compute nodes for over 110 hours to solve Nug30 problems, which would have incurred significant costs on the proprietary systems such as the AWS. We are grateful to Blue Waters and the project staff for providing this invaluable service to the scientific community.

Table 1: RLT2 dual ascent lower bounds on QAPLIB instances (selective)

<table>
<thead>
<tr>
<th>Problem</th>
<th>GPUs</th>
<th>Optimal Value</th>
<th>Lower Bound</th>
<th>% GAP</th>
<th>Itn Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nug25</td>
<td>4</td>
<td>3744*</td>
<td>3610</td>
<td>3.58</td>
<td>7.99</td>
</tr>
<tr>
<td>Nug27</td>
<td>7</td>
<td>5234*</td>
<td>5076</td>
<td>3.02</td>
<td>8.49</td>
</tr>
<tr>
<td>Nug30</td>
<td>12</td>
<td>6124*</td>
<td>5846</td>
<td>4.54</td>
<td>9.08</td>
</tr>
<tr>
<td>Tai25a</td>
<td>1</td>
<td>1,167,256*</td>
<td>1,091,480</td>
<td>6.49</td>
<td>7.75</td>
</tr>
<tr>
<td>Tai30b</td>
<td>12</td>
<td>637,117,113*</td>
<td>620,444,000</td>
<td>2.62</td>
<td>10.73</td>
</tr>
<tr>
<td>Tho40</td>
<td>71</td>
<td>240,516</td>
<td>213,372</td>
<td>11.29</td>
<td>18.41</td>
</tr>
<tr>
<td>Sko42</td>
<td>95</td>
<td>15,812</td>
<td>14,741</td>
<td>6.77</td>
<td>21.54</td>
</tr>
</tbody>
</table>

Table 2: Branch-and-bound results for QAPLIB instances (selective)

<table>
<thead>
<tr>
<th>Problem</th>
<th>Optimal Value</th>
<th>GPU Banks</th>
<th>GPUs/Bank</th>
<th>Nodes Explored</th>
<th>Time (d:h:mm:ss)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nug25</td>
<td>3744*</td>
<td>50</td>
<td>2</td>
<td>3868</td>
<td>0:02:44:24</td>
</tr>
<tr>
<td>Nug27</td>
<td>5234*</td>
<td>100</td>
<td>2</td>
<td>55761</td>
<td>1:02:28:32</td>
</tr>
<tr>
<td>Nug30</td>
<td>6124*</td>
<td>300</td>
<td>4</td>
<td>840273</td>
<td>4:14:06:21</td>
</tr>
<tr>
<td>Tai25a</td>
<td>1,167,256*</td>
<td>100</td>
<td>2</td>
<td>523005</td>
<td>3:13:53:33</td>
</tr>
<tr>
<td>Tai30b</td>
<td>637,117,113*</td>
<td>60</td>
<td>4</td>
<td>30523</td>
<td>2:09:55:17</td>
</tr>
</tbody>
</table>

PUBLICATIONS AND DATA SETS

Date, K., and R. Nagi, GPU-accelerated Hungarian algorithms for the Linear Assignment Problem. Parallel Computing, 57 (2016), pp. 52–72, DOI: 10.1016/j.parco.2016.05.012.

1University of Illinois at Urbana-Champaign
The method developed in this work reroutes communication inside of sparse matrix operations so that MPI messages are aggregated on a node before executing internode communication on the network. When multiple processors on a node of Blue Waters communicate with processors on distant nodes, communication costs are increased further. The key challenge of this project is to limit costly, internode communication through localization of the sparse matrix operations.

METHODS & CODES

Current supercomputing architectures consist of a large number of nodes, each with several multi-core processors. For example, many sparse matrix operations use 16 or 32 processors per node on Blue Waters. The standard approach to sparse matrix communication does not consider the physical location of processors on the network. Yet, the cost of communication varies greatly depending on the locations of each endpoint. For instance, communication between two processes located on the same node incurs a significantly lower cost than communicating between two different nodes.

The results highlight the value in taking advantage of processor layout and topology in irregular communication demands, such as those introduced through sparse matrix operations. As applications and data demands continue to grow in complexity and dimension, localizing communication will be critical to achieving efficiency and taking advantage of the full capacity of the network.

WHY BLUE WATERS

Blue Waters is an ideal platform for testing scalable algorithms for future machines. The method developed in this project could be extended to additional elements of the machine topology, including the socket level and also the Gemini hubs on the network. The scale and network type made Blue Waters a necessary component of the experimentation to support the algorithm development and modeling.
EXECUTIVE SUMMARY

High-performance parallel algorithms for numerical linear algebra play a crucial role in most large-scale computational science problems. We have made progress in two aspects of advancing these types of algorithms: (1) introduction and tuning of distributed-memory functional abstractions for tensor operations, and (2) development of new communication-avoiding algorithms for matrix factorizations. Blue Waters has allowed us to deploy and evaluate these new methods on a leadership-class computing platform. In particular, we have done performance studies on the use of distributed symmetric tensor contractions for the atomic-to-molecular orbital transformation (a key kernel in quantum chemistry computations), of parallel sparse matrix multiplication routines and their use in graph analytics (betweenness centrality), and are currently evaluating a novel communication-avoiding algorithm for QR factorization of rectangular matrices.

RESEARCH CHALLENGE

The biggest challenge facing the parallel scalability of methods in computational science is the overhead of moving data between processors. Our goal is to develop algorithms and libraries that minimize communication in the number of messages as well as in the amount of data moved. To do so in the most useful way, we leverage the ubiquity of numerical linear algebra in scientific computing, targeting the development of algebraic libraries and frameworks. Matrices and tensors (multidimensional matrices) provide high-level abstractions for data sets and transformations thereof. The algorithm aims to realize a QR code that achieves optimal communication and synchronization complexity in theory and is efficient in practice. While the new algorithm is not asymptotically more efficient than the state of the art, it is substantially more simple and easier to implement (no algorithm with the same communication complexity has been implemented previously). The basic idea is to use the Cholesky–QR2 algorithm and leverage communication-optimal parallel Cholesky and matrix multiplication routines. The Cholesky–QR2 algorithm is numerically stable, so long as the matrix is reasonably well-conditioned. We plan to compare the performance of our implementation to the QR routine in ScalAPACK and test its stability on very large matrices.

METHODS & CODES

Our work does not only study hypothetical algorithms but contributes directly to libraries that are available to application developers using Blue Waters or other supercomputing platforms. In particular, our research has focused on extending and tuning Cyclops Tensor Framework (CTF). This library provides distributed-memory support for sparse and dense tensors, automatically mapping contractions and other functions on these distributed data sets. CTF uses performance models to make runtime mapping decisions, using autotuning to train the model parameters. After performing tuning at scale on a training suite of CTF driver routines, we studied performance of two kernels from radically different domains: quantum chemistry and graph analysis.

RESULTS & IMPACT

We conducted a performance study of in-memory and out-of-core CTF versions of an atomic-to-molecular orbital (AO-MO) transformation, immediately showing scalability on problems that are comparable in scale to the largest previously executed. This transformation appears in many high-accuracy quantum chemistry methods and is the most expensive step in some newly proposed methods. This work served to provide preliminary results for a many-principal investigator interdisciplinary proposal focusing on catalysis in chemical reactions.

Additionally, we evaluated the performance of a betweenness centrality code—MFBC—that leverages sparse matrix multiplication functionality in CTF as well as its support for user-defined tensor element-types and functions. We were able to calculate centrality scores for some of the largest graphs publicly available, including the Friendster graph, which has 1.8 billion edges, leading to a paper in Supercomputing ‘17. Fig. 1 displays the parallel scalability of the CTF MFBC code in terms of millions of edge traversals per second.

Over the past year, we have also developed a new parallel algorithm for QR factorization as well as a parallel implementation thereof. The algorithm aims to realize a QR code that achieves optimal communication and synchronization complexity in theory and is efficient in practice. While the new algorithm is not asymptotically more efficient than the state of the art, it is substantially more simple and easier to implement (no algorithm with the same communication complexity has been implemented previously). The basic idea is to use the Cholesky–QR2 algorithm and leverage communication-optimal parallel Cholesky and matrix multiplication routines. The Cholesky–QR2 algorithm is numerically stable, so long as the matrix is reasonably well-conditioned. We plan to compare the performance of our implementation to the QR routine in ScalAPACK and test its stability on very large matrices.
RESULTS & IMPACT
One of the main outcomes of this project is HIPPI [1], a method for protein family classification. Family identification is a basic step in many bioinformatics pipelines, such as metagenomic taxon identification and abundance profiling (first steps in microbiome analysis) and is closely related to remote homology detection, which is a basic step in protein function and structure prediction. BLAST [2] is the most well-known method for this problem, but other approaches based on profile Hidden Markov Models (HMMs) have been used as well. In this work, we developed a novel machine-learning technique to detect membership in existing protein families, where we construct an ensemble of profile HMMs to represent each protein family, and then compare each sequence (which can be short reads or full-length sequences) to each HMM in each ensemble to find the best-fitting protein HMMs to represent each protein family, and then compare each sequence (which can be short reads or full-length sequences) to each HMM in each ensemble to find the best-fitting protein family. We provided an extensive study based on the PFAM [3] database of protein families and their associated profile HMMs from HMMER [4] to compare our method to the previous best methods. This study showed that the technique outperformed all the current methods (including BLAST, HMMER, and HHsearch [5]) in terms of both precision and recall, especially when analyzing short sequences (Fig. 1).

WHY BLUE WATERS
Blue Waters is necessary for at least two reasons. First, the development of these methods requires extensive testing, which is not feasible on other platforms. Second, the analysis of large biological data sets (and even of moderate-sized data sets) often requires years of CPU time (e.g., the avian phylogenomics project spent 450 CPU years to analyze approximately 50 whole genomes). Blue Waters makes this feasible and enables biological discovery.

PUBLICATIONS AND DATA SETS
HIPPI: https://github.com/emirarab/seqp, a github site maintained by Stevah Mirarab (former student).

PASTA+BAI-Phy: https://github.com/MGNute/pasta, a github site maintained by Michael Nute (current Ph.D. student).
HARDWARE ACCELERATION OF DEEP LEARNING

EXECUTIVE SUMMARY

Our project aims to use the Blue Waters platform for hardware acceleration of deep learning for big data image analytics. To achieve near real-time learning, efforts must be paid to both scaling hardware out (increasing the number of compute nodes in a cluster) and scaling up (improving the throughput of a single node by adding hardware accelerators). In this work, we evaluated the performance of scaling up using the GPU-enabled node (XK7) for training convolutional neural networks. The key observation we obtained is that implicit data synchronization across different nodes severely limits the training process. We propose a data manager that explicitly overlaps the data transfer overhead with computation. In the first step, we test the proposed strategy on a single Blue Waters XK7 node. Experimental results show that this strategy achieves a speedup of 1.6X over the implicit data transfer implementation.

RESEARCH CHALLENGE

Deep learning has been widely used in applications such as image classification, speech processing, and object recognition. The huge amount of training data required by the deep neural networks requires more computing power to keep pace of the advance in the state-of-the-art accuracy of these tasks. Mainstream deep learning facilities are CPU-based clusters, which usually consist of thousands of compute nodes. Because the major computation step in deep learning is convolution and matrix multiplication, which is suitable for Graphic Processing Units (GPUs) to compute, modern deep learning facilities are often equipped with GPUs as hardware accelerators.

However, the straightforward implementation of deep neural networks on such GPU-enabled compute nodes will lead to underutilization of compute resources, especially for multi-node systems such as Blue Waters. Therefore, there is a strong motivation to evaluate and characterize the deep learning workload on the GPU-enabled nodes.

In this work, we evaluated the performance of popular types of deep neural networks on a GPU-enabled supercomputer (the XK7 nodes on Blue Waters). From the evaluation results, we observed good scalability of neural networks. Meanwhile, among the three types of neural network layers we evaluated, that is, convolutional layers, fully connected layers, and long short-term memory (LSTM) [1] layers, the convolutional layers have the best scalability. The difference among the three types of neural network layers in terms of scalability comes from the variation of computation per byte in each layer. Given the same number of weights, the convolution layers have a one order magnitude larger number of multiply-accumulate (MAC) operations since the computation complexity of convolution is higher than matrix multiplications. Furthermore, the convolutional layers employ the weight sharing technique, which dramatically increases the computation per byte of the network. Based on these observations, we continue to explore the design space of mapping different kinds of neural networks onto the GPU-enabled supercomputer. In real-world data centers, there are numerous neural network-based applications running concurrently. Since the optimal number of nodes allocated for each type of neural networks varies, we should design a scheduling method to achieve the best efficiency. In the next generation of work, we will conduct more application characterization on the multiple-type neural network workload on Blue Waters.

METHODS & CODE

To evaluate the performance of different types of neural networks, we chose a popular neural network, AlexNet [2], for reference of the convolutional layer and fully connected layer topologies. AlexNet has one convolution layer of (224, 3, 11—the numbers are the size of input image, the number of channels and the size of filter kernels); one convolution layer of (55, 96, 5); one convolution layer of (27, 256, 3); and two convolutional layers of (13, 384, 3). The sizes of the fully connected layers in AlexNet are 4,096; 4,096; and 1,000, respectively. For the topology of LSTM RNNs, we chose a character-based language model of which all recurrent layers have 128 neurons. Since all the LSTM layers are the same, we use only one LSTM layer to run the experiment.

We implemented these neural network layers based on Deepbench, which is a performance benchmark for deep learning hardware accelerators. We modified DeepBench to change the OpenAPI originally used to the platform API of Blue Waters. All nodes allocated are XK7 GPU-enabled nodes.

RESULTS & IMPACT

Fig. 1 shows the running time of each type of neural network on different numbers of nodes. In the figure, we can see that the running time of all three types of layers reduces along with the increase of the nodes used in parallel. The Blue Waters system shows good scalability, although there is communication overhead that makes the speedup sublinear. From the figure, we observe that the speedup of different types of neural networks is different since they have different computation per byte. This observation indicates that we cannot achieve the best performance or system efficiency if we use one single resource allocation scheme for all three types of neural networks. For example, communication dominates the latency for LSTM layers and fully connected layers in the case where we allocate eight nodes, while convolutional layers are still computation-bound. Based on this, we will design new resource allocation and algorithm mapping techniques to achieve better system performance given a fixed amount of workload.

WHY BLUE WATERS

Blue Waters offers us an opportunity to do research on the optimization of deep learning on computational clusters with GPUs. Blue Waters’ XK7 nodes, which consist of one AMD eight-core CPU and one NVIDIA K20 GPU, allows studying of scaling up the computation per node through the addition of GPUs. As GPUs are more suitable than CPUs for convolution and matrix multiplications, which are the major computation in deep learning, state-of-the-art deep learning facilities widely employ GPUs as their hardware accelerators.
A CRITICAL EVALUATION OF THE OP2/OPS PARALLEL MESHING AND CODE GENERATION SOFTWARE

Allocation: Industry/8 Ksh
PI: Kevin Olson¹
Co-PIs: Reid Atcheson¹, Edward Smyth¹, Mike Dewar¹
¹Numerical Algorithms Group

EXECUTIVE SUMMARY
OPS/OP2 is a software package designed to construct algorithms that are easier to implement, maintain, and perform well on different computer architectures, including symmetric multiprocessing (SMP), clusters and clusters of graphics processing units (GPUs). This is accomplished by a programming model where the user defines stencils for performing operations, defining data structures, defining the data and how it is stored on the data structures, and constructing parallel kernels that operate on the data. Once these steps are completed, OPS and OP2 will construct underlying code for different parallel programming paradigms that include sequential, MPI, OpenMP, CUDA, MPI+OpenMP, and MPI+CUDA. Both C/C++ and Fortran programming languages are currently supported by OPS and OP2.

RESEARCH CHALLENGE
The challenge of this research was to test the OPS/OP2 code generation software to determine how easy it is to use and if it can generate code that scales well on large, parallel computer architectures.

METHODS & CODES
OPS/OP2 is software for constructing parallel computer programs [1,2]. OPS focuses on algorithms using structured meshes while OP2 is used for unstructured meshes. Using OPS/OP2 code involves writing either Fortran or C/C++ source code that embeds OPS/OP2 function calls that abstract out the parallelism in the algorithm. A code generation script is then run on the user-written code that generates source code for different parallel programming paradigms that include MPI, OpenMP, and CUDA as well as any combination of them.

RESULTS & IMPACT
OPS was used to construct a finite volume, compressible hydrodynamics code. OP2 was used to construct a Laplace equation solver using an unstructured mesh. The main result was that OPS and OP2 can produce code that runs using all the parallel programming paradigms listed above. Scaling was good in all cases, but MPI-“only” code gave the best performance. Using OPS and OP2 requires an investment in learning its syntax and use. OPS and OP2 could be very useful tools for developing parallel algorithms since it would ease the amount of maintenance work for developing parallel algorithms. We note that OPS and OP2 are continuing to be developed and have evolved several versions ahead of those used for this study. Whether OPS/OP2 can be used in further work at Numerical Algorithms Group is still under evaluation.

WHY BLUE WATERS
Blue Waters was chosen because Numerical Algorithms Group is one of its corporate partners and Blue Waters has the right combination of hardware available: a large number of SMP and GPU nodes that allowed us to perform the required scaling studies to evaluate OPS and OP2.
192. DNA Origami Membrane Channels

194. Molecular Mechanism of Nuclear Transports

196. Improving NWChem Scalability Using the DataSpaces Framework

198. How Function Shapes Dynamics in Protein Evolution

200. Cellulosome Structure Determination by Atomistic Simulations Combined with Experimental Assays

202. Mechanism of Temperature Sensitivity in TRPV1 Channel

204. Exploring the Structure and Dynamics of Converged Ensembles of DNA and RNA Through Molecular Dynamics Simulations

206. Structural Basis for Extreme Cold Tolerance in the Eye Lenses of Teleost Fishes

208. Predicting Protein Structures with Physical Petascale Molecular Simulations


212. Non-Born–Oppenheimer Effects Between Electrons and Protons

214. Brueckner–Goldstone Quantum Monte Carlo

216. Elucidating the Molecular Mechanism of C-type Inactivation in Potassium Channels

218. Exploring the Structure and Dynamics of Converged Ensembles of DNA and RNA Through Molecular Dynamics Simulations

220. A Hybrid Stochastic-Deterministic Simulation Method Enables Fast Simulation of Cellular Processes in Eukaryotes

222. Quantum-Classical Path Integral Simulation of Charge Transfer Reactions

224. Patient-Specific HPC Models and Simulation-Based Imaging for Cardiovascular Surgical Planning

226. Quantum Effects of Proton Transfer in Biological Systems

228. Thermodynamic Characterization of Conformational Landscape in Proton-Coupled Oligopeptide Transporters

230. Machine Learning Reveals Ligand-Directed Conformational Change of p-Opioid Receptor

232. Understanding the Protein Allostery in Kinases and GPCRs

234. Data-Driven, Biologically Constrained Computational Model of the Hippocampal Network at Full Scale

236. Stretching the Cadherin Molecular Velcro® of Cell-Cell Junctions

238. Machine Learning Reveals Ligand-Directed Conformational Change of μ Opioid Receptor

240. Elucidating the Molecular Mechanism of C-type Inactivation in Potassium Channels

242. Studying Cellular Processes through the Computational Microscope

244. Understanding the Protein Allostery in Kinases and GPCRs

246. Predicting Protein Structures with Physical Petascale Molecular Simulations

248. Large-Scale Coarse-Grained Molecular Simulations of the Viral Lifecycle of HIV-1
**DNA ORIGAMI MEMBRANE CHANNELS**

**Allocation:** Blue Waters Professor/240 Ksh  
**PI:** Aleksandr Akimontrovich  
**Collaborators:** Chen-Yu Li1, Jejoong Yoo1, Kerstin Göpfrich2, Iwona Mames3, Satya Prathyusha Bharmukher4, Maria Rizzii2, Adam Mames3, Alexander Ohmann4, Mathias Winterhalter5, Izgen Stule6, Bertalan Gyenes2, Ulrich F. Keyser2

1University of Illinois at Urbana-Champaign  
2University of Cambridge  
3University of Southampton  
4Jacobs University Bremen

**EXECUTIVE SUMMARY**

DNA nanotechnology utilizes self-assembly with nanometer precision for the high-throughput construction of sub-micron-size objects. In comparison to conventional nanofabrication approaches, the DNA origami method is relatively low cost, easy to use, and has an infinite number of possible applications. Using Blue Waters, we have performed landmark molecular dynamics (MD) simulations to characterize the structure and transport properties of two biomimetic DNA origami channels—the smallest [1] and largest [2] DNA channels ever made—working in collaboration with the experimental Keyser lab (University of Cambridge). Once the technology is perfected, the DNA channels could be used to replace biological membrane channels or to deliver drugs across cellular membranes.

**RESEARCH CHALLENGE**

Membrane protein channels are biological sensors with high selectivity and efficiency. One important avenue of medical research is building a synthetic channel that has the same functionality as a biological channel or that performs a user-defined role. Recently, researchers demonstrated that DNA origami-based channels could mimic the ionic conductance and transport properties of membrane protein channels [3–8]. Only after characterizing their structural and electrical properties can these DNA channels be applied to biosensing and drug delivery.

**METHODS & CODES**

We performed explicit-solvent all-atom MD simulations with the latest version of NAMD2 [9–10] of the smallest and largest synthetic DNA channels ever designed, complementing the experimental work of our collaborators in the Keyser Lab. The smallest DNA channel was built from a single DNA helix, and the largest was a megadalton funnel-shaped DNA origami porin (a protein channel that allows passive diffusion). Chemical tags were added to stably embed the channels in a lipid bilayer. Simulations with an applied electric field were then performed to measure the ionic conductance of each channel.

**RESULTS & IMPACT**

Through all-atom MD simulations, we have shown that a membrane-spanning single DNA helix decorated with chemical tags can provide a pathway for ions across the lipid membrane despite the lack of an internal physical channel. Lipid molecules were found to rearrange around the helix, forming a narrow water-filled passage at its circumference, allowing ions and water molecules to pass through the membrane. The average conductance calculated from simulation was in excellent agreement with experiment, and the simulations provided a microscopic explanation for the large variation in ionic conductance measured in experiment and simulation. Results of the single-DNA helix channel were published in *Nano Letters* [1].

Following the same all-atom MD approach, the conductance of the large funnel-shaped DNA channel was measured to be an order of magnitude larger than any previous man-made channel, and its cross-sectional area was similar to that of the nuclear pore complex. Consistent with the results of the single DNA helix, the ion current was found to flow through both the central pore of the channel and along the channel’s walls. Results of the large funnel-shaped DNA channel were published in *ACS Nano* [2].

This work could lead to important applications at the frontier of medical science. Researchers could use synthetic DNA channels as a syringe for specific drug molecules by modifying the channels to recognize selective tissues and to open up the membrane. Furthermore, synthetic channels could be used in artificial tissues to give neighboring cells a new way to communicate.

**WHY BLUE WATERS**

Explicit-solvent all-atom MD simulation is the only computational method that can treat DNA origami objects enhanced by nonstandard functional groups and accurately characterize their structural fluctuations and transport properties [11]. Because of the size of the DNA origami structures, such MD simulations are computationally demanding. The large number of XE nodes on Blue Waters with graphics processing unit accelerators connected by the fast Gemini interconnect make it one of the best publicly available systems for performing DNA origami simulations. Over the past several years, our group has used Blue Waters to carry out a set of landmark simulations in the area of DNA nanotechnology, bringing high-performance simulations to the forefront of this research field.

**PUBLICATIONS AND DATA SETS**

into the physical mechanism of nuclear transport, with important implications for several human diseases and the development of novel gene therapies.

**METHODS & CODES**

We performed explicit-solvent all-atom MD simulations with the latest version of NAMD2 [5–6] of a solution of nup fragments and varied the protein volume fraction within a confined volume, analogous to varying synthetic pore diameter in experiment. Simulations with an applied electric field were then performed to calculate ionic conductivity as a function of nup density. In addition, we used atomic-resolution Brownian dynamics (ARBD) and traditional all-atom MD to model the entire NPC, including the disordered central channel.

**RESULTS & IMPACT**

Through an array of all-atom MD simulations, we characterized the structural fluctuations and electrical properties of a solution of nups varying in protein density. We measured the ionic current of each system in the presence of an applied electric field and observed a transition from conducting to not-conducting ions at a critical nup density, matching a key result seen in experiment [2].

Building upon an atomic model of the NPC scaffold mapped through experiment just last year [4], coarse-grained ARBD simulations were performed to determine the architecture of the central channel. We then built and performed MD simulations of the first complete all-atom model of the NPC—140 M atoms in total—including the outer scaffolding, the disordered inner channel, the nuclear envelope, and the surrounding solvent (shown in fig. 1). This system is one of largest ever modeled with NAMD on Blue Waters to date.

This work could lead to important medical applications. A number of human diseases—including cancer, viral infections, and neurodegenerative diseases [7–9]—are caused by disturbances in nuclear pore transport, and this project may offer new insights into the molecular origin of those diseases and have implications for the development of gene therapy treatments.

**WHY BLUE WATERS**

Explicit-solvent all-atom MD simulations are needed to characterize the structural fluctuations and electrical properties of the disordered central channel of the NPC. Because of the immense size of the entire NPC—140 Million atoms in total—such MD simulations are only possible on a supercomputer with the computational power of Blue Waters. Furthermore, the large number of XK nodes on Blue Waters with graphics processing unit accelerators connected by the fast Gemini interconnect makes it one of the best publicly available systems for performing simulations of the entire NPC.
The purpose of this project is to speed up coupled cluster CCSD(T) computation in NWChem and then incorporate this highly reliable level of theory into parameter optimization procedures for the AMBER classical force field for DNA, with the potential in future of extending this approach to optimization of other popular force fields. The development approach relies on revising the most memory-demanding part of the CCSD algorithm with the possibility of offloading the largest memory arrays to a DataSpaces server. The purpose of using CCSD(T) computation in the parameter optimization procedure is to employ the wealth of available experimental data, which is typically underutilized in common parameter-refining techniques. Moving the parameter optimization engine to a supercomputing platform, and automating the entire process of data extraction and parameter refining, makes the process reproducible, extendable, and portable.

**METHODS & CODES**

The work on performance optimization of the CCSD(T) method in NWChem, which is maintained at Pacific Northwest National Laboratory, relies on offloading large memory arrays from compute nodes to a dedicated data storage employing the DataSpaces data management framework developed by the team headed by Manish Parashar at Rutgers University. While this work is in progress, the previously optimized version of the CCSD(T) method [1] in NWChem [2] has been used to generate a database of experimental data on intermolecular interactions from X-ray crystallographic data of molecular crystals of DNA bases. An automated procedure was developed to extract unique intermolecular pairs from the crystal data, to relax hydrogen atoms, to compute CCSD(T) interaction energies, and to arrange the data in a structured form. A scalable tool to optimize Lennard-Jones parameters in the AMBER (Assisted Model Building with Energy Refinement) force field to fit the parameters to intermolecular interaction energies for experimental geometry of monomers has been developed. It has been tested to run on 1,000 nodes using 32,000 processing units on Blue Waters. The optimization tool generates 200 alternative parameter sets of comparable quality determined by the fitting criterion. The number of alternative parameter sets is easily adjustable in this method. The optimized parameter sets proceed to the final validation step running in parallel to perform MD simulations, to compare the geometry of the system to that from the X-ray data, and to reproduce experimental heat of sublimation, unit cell volume, and interatomic distances.

**RESULTS & IMPACT**

This project introduces a new approach to parameter optimization for classical force fields that combines a high-level electronic structure calculation method to extract additional previously inaccessible information from experimental data, to reengineer the optimization procedure, and to tailor it to maximally utilize high-performance computational resources. The developed procedure shortens the time needed for parameter optimization by roughly a factor of 10. It resolves the issue of accessibility to only a few highly capable teams, transforms the force field optimization from an empirical to a well-structured discipline, educates, and makes the results easy to reproduce by the community.
HOW FUNCTION SHAPES DYNAMICS IN PROTEIN EVOLUTION

RESEARCH CHALLENGE
The biophysical properties of protein loops may hold answers to the discovery, prediction, and annotation of protein functions [1]. Some of these biophysical properties may be governed by yet-to-be discovered evolutionary drivers that could significantly impact synthetic biology and translational medicine [2]. Incorporation of biophysics in protein structure–function studies is becoming increasingly common. However, biophysics is rarely used in evolutionary studies [3]. The main objective of our studies is to bridge disparate disciplines of biology and physics with Molecular Dynamics (MD) simulations performed at nanosecond (ns) timescales to capture evolutionary dynamics on a scale of billions of years. Here, we explore biophysical variables of the MD simulations by studying community structures of protein loop residues that describe the molecular trajectories of the loop regions. Our goal is to dissect evolutionary relationships in these data using evolutionary timelines reconstructed from robust phylogenetic methods [4].

METHODS & CODES
We completed MD simulations that were left pending in a previous allocation, in which we analyzed 87 protein loops from aaRS structural databases on a timescale of 50–70 ns. In addition, we simulated 116 protein loops belonging to single-domain metaconsensus enzymes. The protein loops of aaRS domains were mostly associated with the Gene Ontology (GO) level-1 molecular function of “binding” followed by that of “catalytic activity” (Fig. 1.A). We constructed a dynamics space, a modified version of the dynamosome [5], by calculating the eigenvalues of the top five principal components from principal component analysis (Fig. 1.B) and centrality metrics from a network (Fig. 1.D) based on the dynamic cross-correlation matrix of the motions of protein residues (Fig. 1.E). In order to assess the presence or absence of a specific network topology, we also calculated maximum modularity scores, alpha values to test power law behavior, and Bartel’s test statistic for measuring the extent of modularity, scale-freeness, and randomness of the network (Fig. 2). We are currently in the process of performing unsupervised clustering of the trajectories using the dynamosome variables. We also plan to use methods that classify community structure patterns (Fig. 1.D) exhibited by loops and their correlation to specific function. Our goal is to reconstruct a “structure–evolution” space that would complement our dynamosome.

RESULTS & IMPACT
The aim of our investigation is to detect the presence or absence of patterns of motion in molecules. We focus on an analysis of dynamic network topologies defining a three-dimensional morphospace delimited by the conceptual axes of modularity, scale-freeness, and randomness [6]. Modularity, a feature persistently observed in biological networks [7], embodies flexibility and diversity of the molecular components that make up the whole. Scale-freeness is an indicator of heterogeneity in patterns of connectivity of the network. It is a measure of “economy” (i.e., how easy it is to traverse the network structure). Randomness entails uniform connectivity of nodes throughout the network, a property that confers network fault tolerance. Fig. 2 shows network topology tendencies for 72 of the 87 protein loops that have been annotated with GO functions. This plot encapsulates trade-offs among flexibility, economy, and robustness that result in a “noisy” 2-polytope Pareto front. Interestingly, Zooko’s triangle, a concept used in the design of internet domain-name systems [8], can be thought to illustrate the design space of community networks obtained from protein dynamics. Protein loops possessing various functions tend to cluster at the center of this “triangle” (Fig. 2). They also prefer enhancing modularity in their quest to seek temporal persistence.

WHY BLUE WATERS
The petascale competencies of Blue Waters have been of great value to our project. With the help of such a state-of-the-art system, we have been able to achieve our goal of simulating a significantly large number of proteins (each at a timescale of ~70 ns) in a short time period. NAMD scales well on the Blue Waters architecture, especially when combined with GPU (graphics processing unit) nodes. This provides a significant boost in acceleration [9]. Apart from system specifications that are well-suited to our project, the domain experts/scientists in the Blue Waters support team have helped us smooth out any technical issues that have arisen during the current and previous allocations.
Cellulosomes are multi-enzyme complexes that enable deconstruction of cellulose and hemicellulose in anaerobic cellulosome-containing bacteria. Integration of cellulosomal cohesin–dockerin interactions enables the incorporation of catalytic cellulases and hemicellulases onto the cellulosome scaffoldin, which covalently binds to other components, showing that even a single mutation can cause a large change in cellulosomal structural stability.

Methods & Codes

Combined with biochemical and single-molecule experiments, we employ molecular dynamics (MD), steered MD (SMD), and molecular dynamics simulations to complement single-molecule and biochemical experiments, we characterized cellulosomal components, showing that even a single mutation can cause a large change in cellulosomal structural stability.

Research Challenge

Symbiotic bacteria greatly influence human health and play a significant role in pathogenic disease predisposition, physical fitness, and dietary responsiveness [1]. Moreover, bacteria play a key role in the second-generation biofuel industry where their cellulosytic enzymes are used for plant cell-wall degradation [2]. Here, we investigate key processes underlying bacterial activity, namely plant fiber metabolism. Specifically, we seek to understand how cellulosomes, the highly cooperative macromolecular complex that is central for this metabolic process in some bacteria [3].

Cellulosomes are multi-enzyme complexes that enable deconstruction of cellulose and hemicellulose in anaerobic cellulosome-containing bacteria. Integration of cellulosomal components occurs via highly ordered protein-protein interactions among three major components. In cellulosome assembly, a large noncatalytic polypeptide called the scaffoldin, embedded with various cohesins, anchors dockerin-containing enzymes through cohesion-dockerin interactions (see Fig. 1). Specificity of the cohesion-dockerin interaction allows incorporation of different catalytic cellulases and hemicellulases onto the scaffoldin, which may or may not be bound to another domain tethered to the cell wall [4]. Cellulosome assembly promotes the exploitation of enzyme synergism because of spatial proximity and enzyme-substrate targeting.

Using stochastic search algorithms connected to MD tools, we are building the first comprehensive structure of a cellulosome including enzymatic domains. We expect that a complete model of a cellulosome’s structure will shed light on the mechanism that allows these enzymatic complexes to be highly efficient.

Results & Impact

Following well-established protocols for large macromolecular systems, and using one of the CipA conformations that we obtained using GSAFold, we built a first model of an entire cellulosome structure. MD simulations are now being employed to study the quaternary structure stability. Also, combining biochemical and single-molecule experiments with MD and SMD simulations, we investigated a series of cellulosomal cohesins from Actinobacteria cellulolyticas. We revealed that these cellulosomal components withstand different amounts of force depending on their position in the network. In this study, we combined one-step in vitro expression and specific covalent pulldown of protein constructs to assess the mechanical stability of highly related proteins in a parallel single molecular force spectroscopy assay. Using SMD simulations, we reproduced the experimental results and identified important amino acids. In addition, we used the simulations performed on Blue Waters to suggest mutations that were experimentally performed by site-directed mutagenesis, engineering proteins to pin down single crucial amino acids promoting force resilience. Why Blue Waters

Investigating the structure and functional processes of large enzymatic complex machineries such as the cellulosomes is only possible on petascale computing resources like Blue Waters. Structures obtained using enhanced sampling techniques such as GSA are only reliable if thousands of conformations (models) are predicted. Employing GSA for the numerous linkers of the cellulosome is a well-suited task for the large-scale parallel architecture of Blue Waters.
Figure 1: TRPV1 pore domain with hydrated (left) and empty (right) protein cavities (PCs). Water in the PCs and in the central pore is shown in blue and cyan, respectively.

TRPV1 is an ion channel crucially responsible for transduction of nociceptive stimuli into pain signals. Accordingly, inhibition of TRPV1 is one of the major strategies for designing next-generation antinociceptive drugs. The polyomodal nature of TRPV1 activation, i.e., the fact that a variety of stimuli can open the channel, suggests a complex molecular mechanism of activation whose details are still largely unknown. In particular, we do not currently have a satisfactory microscopic model to explain TRPV1 temperature sensitivity. In this project, we revealed the crucial role played by four nonpolar cavities whose presence and involvement in activation have not been described before. Free energy calculations show that dehydration of these cavities triggers activation of the channel. This observation is able to explain the puzzling response of TRPV1 to diverse environmental factors such as increased cytosolic hydrostatic pressure and osmolality.

RESEARCH CHALLENGE

The nonselective channel TRPV1 is a crucial player in the human nociceptive system. TRPV1 is responsible for the detection of several harmful stimuli such as heat, low pH, and irritating chemicals that are transduced in painful signals originating from peripheral nerves. In chronic pain syndromes, this channel is hypersensitized and the threshold for activation is so low that innocuous stimuli result in constant pain. This involvement in the pain pathways makes TRPV1 an appealing target for designing novel antinpain drugs: selective modulation of this channel would inhibit the generation of the pain signal without interfering with other physiological pathways. Drug discovery campaigns aimed at this target hold promise to deliver pain killers that are virtually devoid of side effects. However, several promising molecules have failed in late stages of clinical trials. The reason for these failures is that TRPV1 is involved in body temperature regulation, and inhibitors of this channel might cause hyperthermia in patients. To make progress in identifying molecules that aptly modulate the channel without interfering with the mechanism of temperature sensitivity, a microscopic understanding of TRPV1 is needed. This project aimed at filling this gap by using long time-scale molecular dynamics (MD) simulations together with enhanced sampling techniques to estimate free energies.

METHODS & CODES

The structure of the TRPV1 capsaicin-bound (CAP-bound) state was taken from the Protein Data Bank: the PDB code is 3j5r [1]. The structure was refined and the missing residues were modeled using Rosetta software [2]. Four capsaicin molecules were docked following the protocol described in [3]. The protein with the ligands was embedded in a hydrated 1-palmitoyl-2-oleoylphosphatidylcholine (POPC) bilayer and surrounded by 150 mM NaCl solution. The overall size of the system was ~170x170x160 Å; the total number of atoms was ~400,000. Two MD trajectories were generated with the peripheral cavities (PCs) either empty or hydrated. The CHARMM36 force field [4] was used to describe the protein and the POPC lipids. For capsaicin, we used the parameters derived in [3]. The TIP3P model was used to describe water [5]. An analogous setup was used to simulate the TRPV1 apo state (PDB code 3j5p [6]). The equilibration of the systems (three in total: the CAP-bound state with empty and hydrated PCs, and the apo state) was performed using NAMD 2.10 software [7] in several steps. Simulations were performed at constant temperature and pressure (1 atm) using the Langevin piston approach. For the Van der Waals interaction, we used a cutoff of 11 Å with a switching function between 8 and 11 Å. The long-range component of electrostatic interactions was calculated using the Particle Mesh Ewald approach [8] with a cutoff for the short-range component of 11 Å. The equations of motion were integrated using a multiple time-step algorithm, with a time step of 2 fs (femtoseconds) and long-range interactions calculated every other step.

We performed metadynamics simulations using the preliminary unbiased trajectories to estimate an upper bound for the free energy barrier and the diffusion constant along the biased collective variable. These were used to obtain an a priori estimate of the error on the reconstructed free energy profile using the expressions reported in [9], which relate the error to the width, height, and deposition rate of the hills. This estimate informed our choice of the metadynamics parameters. Metadynamics simulations were performed using the collective variable module implemented in NAMD2.10 [10] at three temperatures: 280 K, 300 K and 340 K.

RESULTS & IMPACT

We found that the lower gate is open or closed depending on the conformation of N676, an amino acid located on S6 at the edge of a π-helix segment (Fig. 1). This segment is characterized by extreme conformational flexibility [11]. Not all the backbone hydrogen bonds can be simultaneously satisfied and therefore their pattern is dynamic. The presence of the π-helix allows N676 to easily rotate in and out the central pore. This motion is, in turn, controlled by the hydration state of the adjacent protein cavity (PC). Besides extensive accessibility experiments by Salazar, et al. [12], the presence of these cavities is supported by alanine scanning mutagenesis performed on the S6 segment [13]: the residues lining the PC were shown to produce the greatest perturbation to the channel activation in response to several stimuli, including capsaicin and heat. Interestingly, our molecular mechanism does not entail any large conformational rearrangement of the TRPV1 central pore, whose radius profile in not dramatically altered by the closed-open transition. The hydration/dehydration of this compartment is, in fact, controlled by N676, which upon rotation changes the hydrophobic character of the molecular surface lining the central pore. This susceptibility to perturbations is not uncommon in the pores of ion channels. Wet-to-dry transitions have been reported several times [14] and are, arguably, the result of a precise evolutionary optimization. This microscopic picture provides the basis for rational design of precise modulators of TRPV1.

WHY BLUE WATERS

We investigated a system of approximately 400,000 atoms using MD simulations. The time scales involved in the activation process of TRPV1 dictated trajectory lengths on the microsecond time scale. This was possible thanks to a massively parallel calculation enabled by the computational capabilities of Blue Waters.
EXPLORING THE STRUCTURE AND DYNAMICS OF CONVERGED ENSEMBLES OF DNA AND RNA THROUGH MOLECULAR DYNAMICS SIMULATIONS

Over the past two years of using Blue Waters and taking full advantage of its computational power, we have made extensive progress in the realm of biomolecular simulation methodologies, specializing in nucleic acid structure, dynamics, and ligand–protein binding. The AMBER package for biomolecular simulation and its GPU (graphics processing unit) code has proven to be high-performance and reliable, taking full advantage of Blue Waters. As our group continues to develop specialized software and methodologies to analyze the vast amount of sampling information, we further increase our understanding of relevant biological DNA, RNA, and protein structures.

METHODS & CODES

Molecular Dynamics (MD) simulations have been one of the most important tools in the computational chemist’s toolbox for the last 25 years. Useful as it is, this technique has considerable limitations, mainly in two areas: force field validation and conformational sampling. These two problems are deeply related: As more sampling time is achieved, force field discrepancies are found. The force field is temporarily fixed; however, as computational power grows, more sampling time is achieved and new force field issues are found. Blue Waters has enabled our group to achieve enough sampling time to rigorously validate the models used and to guide the next steps of research.

RESULTS & IMPACT

Key results are described in detail in our publications (below).

WHY BLUE WATERS

Because Blue Waters is optimized for GPU-centered calculations, as well as massive CPU parallelization, it has been a key tool for our group and allowed us to independently run hundreds of biomolecular simulations of myriad systems of DNA and RNA. In some cases, we have reached well over millisecond timescales of combined sampling time. In addition, the highly trained Blue Waters staff have been in close collaboration to develop custom and specialized code capable of performing the analysis of terabytes of simulation data.

PUBLICATIONS AND DATA SETS

http://amber.uchicago.edu


STRUCTURAL BASIS FOR EXTREME COLD TOLERANCE IN THE EYE LENSES OF TELEOST FISHES

Allocation: Illinois/113.5 Knh
PI: Christina C.H. Cheng

1University of Illinois at Urbana-Champaign

EXECUTIVE SUMMARY

Eye lenses of endothermic mammals such as the cow develop cold cataracts at a mild 17°C. In contrast, ectothermic teleost fish lenses remain transparent down to -12°C. Cold-induced cataracts arise from a liquid–liquid phase-separation of lens proteins (crystallins) resulting in a protein-rich and a protein-poor phase. Crystallins are tightly packed at high concentrations to enable refraction of incident light, and teleost lenses are especially protein-dense to achieve a refractive index change in aquatic environments. Attractive forces would enable crystallins to tightly pack in the lens but risk increasing propensity for phase separation.

We propose that teleost crystallins are structurally more flexible than mammalian paralogs to minimize the propensity of phase separation at the high concentrations necessary to function in aquatic environments, conferring the observed tolerance to very low temperatures as a side benefit.

RESEARCH CHALLENGE

Attractive forces necessary for the tight packing of lens crystallins result in a protein-rich and a protein-poor phase. Crystallins are tightly packed at high concentrations to enable refraction of incident light, and teleost lenses are especially protein-dense to achieve a refractive index change in aquatic environments. Attractive forces would enable crystallins to tightly pack in the lens but risk increasing propensity for phase separation.

We are currently utilizing the computational power of Blue Waters to run extensive molecular dynamics simulations to address our hypotheses regarding flexibility and extreme cold tolerance. With this resource, we are able to ascertain the potential contribution of flexibility to resist cold cataracts at very low temperatures. Attractive forces, size, and shape may all be essential parameters in maintaining homogeneity in the lens at high concentrations and over a range of temperatures. Flexibility, in part, appears to explain the large standing question regarding the incredible ability of teleost fish lenses to maintain transparency at extremely cold temperatures.

whY blue Waters

Our work requires simulating three (3) trials of 49 proteins at two temperatures and over a long times course of 50 nanoseconds to detect meaningful molecular behavior. This work is at the core of a Ph.D. project in determining the extreme cold tolerance observed in teleost fish lenses. Only the petascale computational power and resources of Blue Waters could allow us to achieve this core portion of the project in a reasonable amount of time for downstream analyses to test our hypotheses. Without Blue Waters, we would not be able to finish this project in a reasonable time for a Ph.D. project.

RESULTS & IMPACT

At 0°C, it is evident that zebrafish γ-crystallins are largely more flexible than the mammalian isoforms, most notably on surface loops (Fig. 1a). Fig. 1b shows simulations of mammalian γ-crystallins at 37°C and presents flexibility profiles similar in amplitude to zebrafish isoforms at 0°C. The zebrafish γ-crystallin isoforms tested at 25°C exhibit greater flexibility profiles compared to 0°C (Fig. 2a). In accordance with our hypothesis concerning variation of flexibility profiles among the γM-crystallins, the 11 zebrafish γ-crystallins are not identical in amplitude across all isoforms (Fig. 2b). This is unlikely to be due to functionally neutral changes along the γM-crystallin evolutionary trajectory, but, rather, is likely to be due to functional diversity. Identical flexibility at regions with significant attractive forces may mitigate the effects the attractive forces have in maintaining high concentrations necessary for light refraction in teleost fish lenses. Diversity with regard to flexibility, spatial distribution of attractive forces, size, and shape may all be essential parameters in maintaining homogeneity in the lens at high concentrations and over a range of temperatures.

METHODS & CODES

We ran molecular dynamics (MD) simulations on eleven zebrafish and five mammalian isoforms at a cold temperature (0°C), and 3 zebrafish and 5 mammalian γ-crystallin isoforms at the normal body temperature (25°C and 37°C respectively).

Three replicates of each γ-crystallin isoform were simulated for 50 nanoseconds in NAMD 2.12 using CHARMM27 force field parameters. Each of the five mammalian γ-crystallin isoforms was simulated using solved structures obtained from the Protein Data Bank. Simulation of zebrafish γ-crystallin isoforms used one known structure, the γM-crystallin, and the remaining 10 γ-crystallin isoforms simulated in this study were modeled onto the γM-crystallin using Tasser.

VMD 1.9.3 was used to quantify flexibility via root mean square fluctuations (RMSFs), which measure the average distance (Angstroms) of aligned backbone Ca atoms per residue of a protein over the duration of the simulation. Average RMSF values were taken for each isoform, then formatted based on amino acid sequence alignment generated by MUSCLE 3.8.31 for comparison.
PREDICTING PROTEIN STRUCTURES WITH PHYSICAL PETASCALE MOLECULAR SIMULATIONS

Allocation: NSF PRAC/2,900 Ksh
PI: Ken A. Dill
Co-PI: Alberto Perez
Collaborators: Cong Liu, Emiliano Brini, James Robertson, Lane Votapka, Roy Nassar

Collaborators:
Ken A. Dill
PI:
NSF PRAC/2,900 Knh
Allocation:
MOLECULAR SIMULATIONS
PREDICTING PROTEIN STRUCTURES WITH PHYSICAL PETASCALE BLUE WATERS ANNUAL REPORT
1Stony Brook University

EXECUTIVE SUMMARY

Blue Waters has enabled the first use of physics-based methods to make accurate atomistic predictions of protein structure based on sequence information alone in a double-blind international competition called CASP (Critical Assessment of protein Structure Prediction). This is a step forward for the computational biophysics community and an unprecedented result in the last 24 years of CASP, where physics has not had a role in ab initio modeling.

Further, we have used this methodology with great success to complete the pipeline that would lead to drug discovery: from sequence to structure, to interactions with small drugs, peptides, and other proteins. We are developing new protocols in these areas and have performed well in matching experimental results for the MDM2/p53 system, a cancer target. We continue to make headway in improving the accuracy and speed of our unique technology (MELD, or Modeling Employing Limited Data), which requires the significant GPU (graphics processing unit) resources available on Blue Waters.

RESEARCH CHALLENGE

A grand challenge in the last 60 years is to use computers to predict how proteins fold from a polymer sequence into a 3D structure. Proteins are the workhorses of the cells, made from 20 different amino acids whose arrangement in a linear chain gives rise to the 3D structure. We want to learn the physical principles that lead to the 3D structure so we can design drugs to bind to these proteins and stop them from misbehaving, or to design new protein sequences that carry out new functions (e.g., digest oil spills).

Computationally cheap knowledge databases are often used to copy and mimic structures based on sequence similarity. Computationally expensive physics-based methods are necessary to push the boundary of our predictive ability.

METHODS & CODES

Physics models are accurate but too slow to tackle problems such as folding. MELD [1,2] is our solution. In this approach, we accelerate molecular dynamics (MD) simulations while integrating ambiguous, noisy, and sparse data through a Bayesian inference approach. MELD allows us to find the best agreement between physics and a subset of the data. On one hand, this reduces the search space of the system, accelerating the convergence time of physics-based methods. On the other hand, MELD allows us to deal with imperfect sets of information, such as data derived from a generic knowledge of the system from experiments. Our MD engine is based on the OpenMM program [3] and Amber [4] for force fields and setup.

RESULTS & IMPACT

With Blue Waters, we have taken MELD and physics to the limit. We have participated in CASP, the blind protein structure prediction competition. This worldwide event in which more than 200 groups participate has taken place every other year for the last 24 years. During the three months of competition, hundreds of targets are released and predictions have to be submitted in a timely manner. Before MELD, atomistic physics-based simulations had not been possible this quickly. Thanks to MELD and Blue Waters, not only were we able to make predictions, some of them were the best of the entire competition. This is an unprecedented result in the last 24 years of CASP and in using physics-based approaches for structure prediction.

We put MELD to work on problems where current technologies fail. MELD is particularly suited for predicting the structure of small proteins with little homology. These are at the intersection of cases where the faster database methods do not work and where the computational cost of MELD is still acceptable. We are working on microproteins (~70 amino acids) with no known experimental structure as well as a glass-coiling set of proteins for which the experimental structures are known [5]. We are obtaining encouraging results in both cases, and we have active collaborations with experimentalists to prove the quality of our prediction (using circular dichroism and nuclear magnetic resonance spectroscopy). In the second case, we have so far had a 25% success rate (see Fig. 1).

Sometimes protein folds upon binding to other proteins. This means that a disordered protein in solution obtains a specific 3D shape when it interacts with a specific receptor. Common computational tools fail to predict the binding pose due to the rearrangement of the structure during the process. With MELD, we have been able to predict binding poses and relative binding free energies for the complex P53-MDM2, which is involved in the development of cancer. The protocols we designed for this project are unique and have significantly advanced the field. Without Blue Waters, this project would have taken us about 16 times longer to simulate.

Figure 1: There are 463 nonfractionable proteins in the Protein Data Bank whose structure is known. When predicting the structure based on sequence alone, using state-of-the-art homology methods (left) the predicted structure is wrong. Using MELD and physics, we get the right answer (right: experiment in red, our prediction in blue).

Figure 2: Protein-protein conformations predicted with MELD (top left). Populations (lower left) are related to free energy; hence, we can pick up the right answer by clustering results. Our method so far has scored in 19 of 20 cases (right).

Proteins bind to each other and to other molecules. With MELD, we are able to investigate the way complexes of two proteins are arranged in space (see Fig. 2) and how small ligands (i.e., small molecules) can bind to a protein. Ultimately, physics governs both processes, and MELD allows for a tightly focused search process. This is an exciting new line of research made possible by the computational resources of Blue Waters.

WHY BLUE WATERS

Blue Waters has been the perfect resource for us. It is the only cluster in the United States where we can get enough throughput GPU (graphics processing unit) usage, especially for the time-sensitive CASP competition. The staff is also helpful and quick to respond to solve issues and help us to maximize our Blue Waters usage for efficiency. We have a 100-GPU cluster in our lab (slow 2050s and 2070s models). Every MELD calculation requires at least 30 GPUs, so at most three calculations can run in the lab, whereas on Blue Waters we can sometimes run up to 30 calculations.

PUBLICATIONS AND DATA SETS


208
MULTISCALE MODELING OF BIOFILM DYNAMICS IN DRINKING WATER DISTRIBUTION SYSTEMS: TOWARD PREDICTIVE MODELING OF PATHOGEN OUTBREAKS

Allocation: Exploratory/50 Knh
PI: Ahmed Elnashai
Co-PI: Helen Nguyen
Collaborator: Jaehyuk Kwack

EXECUTIVE SUMMARY
Biofilms are aggregates of cells and extracellular polymeric substances. They are found ubiquitously in both natural and engineered systems, such as on the surfaces of pipes in drinking water distribution systems (DWDS). Biofilms in DWDS have been reported to be capable of attracting and harboring pathogens. In addition, the biofilm matrix may prevent disinfectants from reaching the cells located deep inside the biofilm. As a result, pathogenic microorganisms have been found in DWDS biofilms and have been linked to outbreaks and severe health problems. Understanding the mechanisms of pathogen attachment to biofilms developed in DWDS is of crucial interest to ensure the quality of drinking water and is a critical public health issue. This project contributes to establishing a comprehensive multiscale framework for studying the dynamics of growth and detachment of biofilms in DWDS under different operating conditions by integrating fluid mechanics, solid mechanics, and chemistry to predict pathogenic outbreaks and related public health hazards.

RESEARCH CHALLENGE
The quality of drinking water is a critical public health issue. The formation and detachment of biofilms in DWDS have been linked to pathogenic outbreaks. Yet, an understanding of the response of biofilms to different types of disinfectants and their susceptibility to fracture under different flow rates and growth environments is still lacking. Several challenges exist, including the absence of a comprehensive constitutive model for the nonlinear and inelastic response of the soft biofilm, the extreme computational demand of the 3D coupled fluid-film simulation, and the numerical challenges associated with meshing the complex biofilm geometry. As a preprocessing step, numerical meshes are decomposed by the METIS/ParMETIS library. Each MPI process accesses its own decomposed numerical domain (i.e., a series of elements and nodes) and locally computes its residual vectors and tangent matrices. After that, global communications of matrix-vector calculations, assembly process, and residual norm calculations are implemented via the PETSc library. Krylov subspace methods from PETSc and sparse direct matrix solvers (e.g., SuperLU_dist and MUMPS) are employed to solve linear equations with the globally assembled matrices and vectors. The output files are written in VTK file formats and they are post-processed with ParaView.

METHODS & CODES
Over the past allocation period, we have been able to make fundamental progress in addressing some of these challenges. We have developed a continuum nonequilibrium statistical thermodynamics framework based on the effective temperature thermodynamics framework and have been able to make fundamental progress in addressing some of these challenges. We have developed a continuum nonequilibrium statistical thermodynamics framework based on the effective temperature approach to model the nonlinear elastic-plastic response of soft amorphous materials. We implemented this material model in a nonlinear finite deformation framework within the MOOSE platform, a finite element framework from Idaho National Lab for multiphysics simulations. MOOSE provides an advanced modular computational infrastructure including the libMesh finite element library and the PETSc solver library. The program requires input of the weak form for partial differential equations to be solved as well as the material model, but it provides a variety of shape functions, stabilization options, and locking control algorithms.

For the bio-film fluid interaction, we used a finite element code for Computational Fluid Dynamics and Fluid-Solid Interaction developed by Dr. Jaehyuk Kwack, who has worked in the Scientific and Engineering Applications Support group at NCSA.

The application code employs a variational multiscale approach for turbulence and nonlinear material modeling; as a result, it provides an optimal computational performance for multiphysics problems that will be investigated through this project. For parallel computation, the code is implemented with the Message-Passing Interface (MPI). As a preprocessing step, numerical meshes are decomposed by the METIS/ParMETIS library. Each MPI process accesses its own decomposed numerical domain (i.e., a series of elements and nodes) and locally computes its residual vectors and tangent matrices. After that, global communications of matrix-vector calculations, assembly process, and residual norm calculations are implemented via the PETSc library. Krylov subspace methods from PETSc and sparse direct matrix solvers (e.g., SuperLU_dist and MUMPS) are employed to solve linear equations with the globally assembled matrices and vectors. The output files are written in VTK file formats and they are post-processed with ParaView.

RESULTS & IMPACT
Our results thus far have been pushing the limits of the state of the art in modeling biofilm mechanics. To the best of our knowledge, prior work in the field has been limited to two-dimensional models. We have been able to run full 3D fluid-structure models of biofilm with complex surface geometry. As a result, we can realize complex 3D turbulent structures near surface irregularities as well as complex stress patterns that are not apparent from the 2D simulations. Furthermore, the 3D implementation of the amorphous material model enables us to realize—for the first time—complex strain localization patterns (which will potentially map into complex 3D fracture surfaces) that have not been reported before.

WHY BLUE WATERS
The progress accomplished so far would not have been possible without Blue Waters. Several of the problems we have tested required 10,000 or more core-hours of runtime and generated tens to hundreds of gigabytes of data per simulation. These resources are not accessible on conventional cluster scales. Furthermore, cloud computing is not suitable for the nature of our problems since any parallelizable discretization of the coupled system of equations of interest would require domain decomposition and efficient communication among processors, making cloud computing out of question.
COMPREHENSIVE IN SILICO MAPPING OF DNA-BINDING PROTEIN AFFINITY LANDSCAPES

Allocation: GLCPC/310 Keh
Co-PI: Peter Freddolino
PI: Morteza Khabiri
Collaborator: Atta Jolfaia
1University of Michigan Medical School
2University of Toronto

EXECUTIVE SUMMARY

Transcription factors (TFs) and other DNA–binding proteins shape the behavior of all cells, coordinating appropriate gene expression patterns in response to internal or external cues. For any particular transcription factor, maps of the binding affinity for different DNA sequences must be obtained through laborious and expensive experiments. We used Blue Waters as our computing resource to computationally map the TFs’ binding landscapes using these methods thus remains an area of active investigation.

Figure 1: Binding free energy landscape for ELK1. Computational values (black) are compared with human (green) and mouse (red) values from experimental data. The small box in the landscape plot indicates example of work distribution achieved from mutant → wildtype (blue) and wild type → mutant (green) transition. Data from [2].

RESEARCH CHALLENGE

Transcriptional regulation is driven in large part by the action of TFs and other DNA-binding proteins that either recruit or inhibit the recruitment of RNA polymerase. The activity of TFs that interact directly with DNA depends on their potent ability to recognize and bind to specific DNA sequences. In reality, most TFs bind to a range of related sequences; however, the one that provides enough information or has the lowest energy will define a reliable TF binding site. Thus, understanding the binding affinity landscape of each transcription factor to all possible sequences of DNA will help us to predict the behavior of regulatory networks. There have been a number of efforts to predict these binding sites quantitatively based on both experimental methods (e.g., protein–DNA binding microarrays) [1] or HT-SELEX experiments [2]. However, most of these experiments carry high costs in terms of time and money. Recent computational efforts by Segeliger and co-workers [4] predicted an accurate DNA–binding affinity landscape based on atomistic molecular dynamics simulations and reproduced experimental results to near-chemical accuracy. This proved to be much more accurate than previous computational methods. By using the massive computing resources provided by Blue Waters, we attempted to apply a similar method for free-energy calculations to that used by Segeliger and coworkers to map the complete binding free-energy landscapes of four TFs with well-characterized binding affinity landscapes.

METHODS & CODES

Building on previous results that showed accurate calculation of protein–DNA binding free energies for a small number of cases [4], we applied the Crooks-Gaussian intersection (CGI) method [6] to calculate the free-energy changes for base pair substitutions in the binding site of the transcription factors of interest. This method requires calculations of very long equilibrium simulations of the protein-DNA complex and the DNA alone for each of two sequences to be compared, followed by many short simulations morphing the system between the two sequences. We performed the free-energy calculations for all possible single nucleotide perturbations of the consensus binding site for the transcription factor of interest. Our results illustrate both very poor overall correlations with the experimental results and, frequently, unphysically large magnitudes of binding free-energy changes [5]. Excluding many sources of possible errors in our simulation setup, we realized that neither extensive control simulations using simplified systems or other free-energy calculation methods, nor careful characterization of the structural features involved in the protein–DNA interface of the simulated complexes, provided an explanation for the poor correlation between calculated and experimental binding free-energy landscapes. We are currently working to resolve this difficulty so that we can realize the promise of computational predictions of DNA-binding affinity landscapes.

WHY BLUE WATERS

The computational work described here requires the capability to efficiently bring huge numbers of nodes together to run dozens of simulations of independent trajectories using GPU-accelerated molecular dynamics software, and then, for each such trajectory, to perform more than 100 short follow-up simulations using CPU-only code for the free-energy calculation. The hybrid architecture of Blue Waters has been absolutely ideal for these applications, providing us with the most efficient possible environment for each portion of our workflow, and allowing us to make progress on huge numbers of mutational calculations simultaneously.

PUBLICATIONS & DATA SETS

Figure 1: (Left) A schematic diagram of the photocycle in BLUF photoreceptor proteins. The dark state to light state conversion is believed to involve photoinduced PCET effects between nuclei and electrons in computer simulations. The development of non-Born–Oppenheimer methods has ever predicted that a positron will bind to the atomic systems. Because of this limitation, the SVM method cannot be used for positrons and one proton or proton-positron correlation was found to be essential for even qualitative accuracy, as NEO methods that neglect explicit electron-positron correlation did not predict the positron to bond to the atomic systems. The NEO-RCXHF method has better computational scaling characteristics than SVM, and SVM calculations have never been agreed well with highly accurate stochastic variational method results. The NEO-RCXHF results agreed well with highly accurate stochastic variational method (SVM) calculations. The inclusion of explicit electron-positron correlation was found to be essential for even qualitative accuracy, as NEO methods that neglect explicit electron-positron correlation did not predict the positron to bond to the atomic systems. The NEO-RCXHF method has better computational scaling characteristics than SVM, and SVM calculations have never been performed on a system with more than five quantum particles. Because of this limitation, the SVM method cannot be used for larger molecular systems. No ab initio or first principles method has ever predicted that a proton will bond to an alkane, which has been observed experimentally. The accuracy of the NEO-RCXHF calculations combined with this method’s computational tractable-scaling characteristics demonstrate that NEO-RCXHF is a promising method for computing the binding energy of positrons to molecules such as alkanes. Our free energy simulations of the AppA BLUF domain have elucidated the active site conformations before and following photoexcitation [9]. For the dominant Trp/Trp conformation, the hydrogen-bonding pattern conducive to the proton relay from Tyr21 to the flavin chromophore is not thermodynamically favorable on the ground electronic state but becomes more favorable, corresponding to approximately half of the configurations sampled, on the locally excited state of the flavin. The calculated energy gaps between the locally excited state and the charge-transfer state associated with electron transfer from Tyr21 to the flavin suggest that electron transfer from Tyr21 to the flavin is more facile for configurations conducive to proton transfer. When the active site conformation is not conducive to PCET from Tyr21, Trp104 can compete directly with Tyr21 for proton relay from Tyr21 to the flavin chromophore. The calculated energy gaps between the locally excited state and the charge-transfer state suggest that electron transfer from Tyr21 to the flavin is more facile for configurations conducive to proton transfer. When the active site conformation is not conducive to PCET from Tyr21, Trp104 can compete directly with Tyr21 for proton relay from Tyr21 to the flavin chromophore.
**EXECUTIVE SUMMARY**

We introduced and fully developed novel scalable algorithms and software for predictively accurate \textit{ab initio} electronic-structure calculations for large molecules and solids, which are not easily subjected to fast calculations by fragmentation. We transform the usual, nonscalable sum-of-products expressions of MP2, second-order Green’s function (GF2) theory, and their CBS corrections by explicitly correlated \textit{F12} ansätze [1—Monte Carlo MP2 (MC-MP2) \cite{MP2}, Monte Carlo GF2 (MC-GF2) \cite{GF2}, Monte Carlo explicitly correlated MP2 (MC-MP2-F12) \cite{MC-MP2-F12}, and Monte Carlo explicitly correlated GF2 (MC-GF2-F12) \cite{MC-GF2-F12}]. These calculations were based on the redundant-walker algorithm, which propagates more walkers than minimally necessary and permutes them in all possible ways when being substituted into the integrand, thereby multiplying the sampling efficiency. We introduced \cite{MC-MP2-F12} a two-level parallelism in which dense matrix multiplications for many walkers are fine-grained on a GPU and a Monte Carlo integration itself is coarse-grained across multiple CPU-GPUs. In this way, not only did we observe a speedup by a factor of 31,000 on 256 GPUs relative to a serial execution, but we also found that the saturation point of the acceleration is significantly delayed to a much greater number of walkers. This is a rare instance in which the parallel architecture (GPU) and algorithm (the redundant-walker algorithm) mutually enhance each other.

**METHODS & CODES**

We mathematically transformed the usual sum-of-products expressions of MP2, second-order Green’s function (GF2) theory, and their CBS corrections by explicitly correlated (F12) ansätze into single high-dimensional integrals by a Laplace transform. These integrals are then evaluated by a Metropolis Monte Carlo method with judiciously chosen weight functions. The resulting stochastic (Brueckner–Goldstone quantum Monte Carlo) methods [1—Monte Carlo MP2 (MC-MP2) \cite{MP2}, Monte Carlo GF2 (MC-GF2) \cite{GF2}, Monte Carlo explicitly correlated MP2 (MC-MP2-F12) \cite{MC-MP2-F12}, and Monte Carlo explicitly correlated GF2 (MC-GF2-F12) \cite{MC-GF2-F12}]—can compute energy differences (correlation energies and electron detachment/attachment energies) directly without a sign problem in a scalable manner with respect to both computer size (on thousands of CPUs or hundreds of GPUs) and system size (the operation cost is linear scaling per MC step and cubic to quartic scaling to achieve given relative accuracy and the memory cost is negligible) \cite{MC-MP2-F12}. They can also calculate quasiparticle energy bands of a solid for the entire Brillouin zone as nearly continuous curves of a wave vector \cite{GF2} and have been extended to third-order MP (MP3) \cite{MP3} using an expedient interpretation of Brueckner–Goldstone diagrams as well as a convergence-acceleration scheme (redundant-walker algorithm) \cite{MC-MP2-F12}.

**RESULTS & IMPACT**

The MP-MC-F12 method enabled an exact (CBS-limit) MP2 energy calculation of tetrahydrocannabinol (472 basis functions) without a local-correlation scheme (Fig. 1). Exploiting the extraordinary flexibility of this algorithm in using virtually any explicitly correlated factor, we numerically characterized the performance of 17 such factors. We observed that highly performing factors share the same short-range behavior within the radius of 1.5 Bohr, while differing greatly in the long-range behavior. This result reveals fundamental electron-correlation physics that a correlation hole of a pair of electrons has a universal size (1.5 Bohr) and concave shape (dictated by Kato’s cusp condition) regardless of its molecular environment or energy (Fig. 2). We have completed the development of the MC-GF2-F12 method, which can compute electron detachment/attachment energies directly in the CBS limit. With this, we computed exact GF2 electron affinities with a statistical uncertainty of 0.03 eV for C_{60} and C_{70}, which play important roles in heterojunction solar cells as an electron acceptor but resist a local-correlation scheme for fast calculations.

These calculations were based on the redundant-walker algorithm, which propagates more walkers than minimally necessary and permutes them in all possible ways when being substituted into the integrand, thereby multiplying the sampling efficiency. We introduced \cite{MC-MP2-F12} a two-level parallelism in which dense matrix multiplications for many walkers are fine-grained on a GPU and a Monte Carlo integration itself is coarse-grained across multiple CPU-GPUs. In this way, not only did we observe a speedup by a factor of 31,000 on 256 GPUs relative to a serial execution, but we also found that the saturation point of the acceleration is significantly delayed to a much greater number of walkers. This is a rare instance in which the parallel architecture (GPU) and algorithm (the redundant-walker algorithm) mutually enhance each other.

**WHY BLUE WATERS**

The stability and ease of use (OS, compilers, libraries, and NCSA expertise) as well as the balanced deployment of CPUs and GPUs are all essential for rapid coding/profiling of new scalable algorithms from scratch and their capacity testing.

**PUBLICATIONS AND DATA SETS**


RESEARCH CHALLENGE

We seek to understand how influenza viral infection responds to membrane perturbation. Initial data suggest that lipid-modifying drugs may have potential as antivirals, but we do not understand their mechanism. Even the fundamental lipid biophysics of viral membrane organization is unknown, and the link from atoms to patients is a grand challenge in understanding and treating viral infection.

METHODS & CODES

We incorporate structural data from X-ray crystallography and cryo-electron microscopy along with mass spectrometry to construct initial models of influenza viral and target membranes. These membrane patches are then simulated using the molecular simulation software that we helped develop, Gromacs, and are measured under perturbations that match what we do experimentally. We use in-house machine-learning methods to analyze the simulations and connect them back to experiments. This then yields a model of the membrane response and its infectious consequences at a high level of molecular detail yet also with experimental validation.

RESULTS & IMPACT

Our results thus far have yielded a new membrane-organizing principle that explains our experimental data on influenza virus binding and cell entry. This molecular explanation helps us understand how the influenza virus responds to membrane changes and is likely applicable to cellular signaling as well. Such detailed molecular models of mesoscale phenomena would not be possible without this combination of precise biophysical experiments and large-scale computing.

WHY BLUE WATERS

Blue Waters offers a large number of GPUs that are tightly coupled by a fast Cray interconnect, permitting us to run many simultaneous simulations that each use a number of GPUs to accelerate the calculation.
A HYBRID STOCHASTIC-DETERMINISTIC SIMULATION METHOD ENABLES FAST SIMULATION OF CELLULAR PROCESSES IN EUKARYOTES

EXECUTIVE SUMMARY

Stochasticity in transcription is an important source of noise that can have a profound effect upon the fate of a living cell. In recent years, we have seen the advent of a community of researchers interested in performing stochastic simulations of large biological systems (e.g., millions of particles). The Lattice Microbes (LM) software suite and its pyLM problem-solving environment provide a convenient way to set up simulations of complex biological systems. However, simulations of large systems performed using the exact Stochastic Simulation Algorithm (SSA) to solve the Chemical Master Equation (CME) are computationally expensive. To alleviate this issue, we have implemented a hybrid CME-ODE (ordinary differential equations) method for LM similar to previous mixed methodologies. Our robust hybrid implementation gives unbiased realizations of these Markov processes.

RESEARCH CHALLENGE

Many processes within living cells, especially gene expression, are characterized by low particle numbers and a high degree of randomness. CME and its spatially resolved analog the Reaction-Diffusion Master Equation (RDME) are descriptions of cellular processes where the system is considered to follow a Markov jump process on the state space of particle numbers in time, capturing the discreteness of the particles and the random nature of individual chemical reactions. Gillespie’s widely used SSA method [1] provides an effective technique for obtaining unbiased realizations of these Markov processes. However, this algorithm is limited because reaction events are accounted for explicitly by the SSA, making simulations of highly reactive systems, where the time between reactions is small, computationally expensive. Highly reactive systems are characterized by large reaction propensities that can arise in the case of high copy numbers, such as metabolites in millimolar concentrations, and/or large rate constants (fast reactions). A challenging and typical scenario is when species participating in slow reactions interact with species involved in fast reactions, making the dynamics of the slow reactions dependent on the fast reactions. To alleviate the issues faced by the SSA for high particle number systems, many researchers have developed hybrid multiscale stochastic approaches [2,3,4] in which the highly reactive parts of the system are described by ODE and the slow reactive parts are described stochastically. Our hybrid method along with an easy-to-use interface through LM [5] and pyLM [6] provides an effective way to study stochastic behavior in highly reactive systems.

RESULTS & IMPACT

Such a CME-ODE partitioning works well for both bacterial and eukaryotic systems where stochastic effects are important. Partitioning typically improves the speed of the numerical simulations by a factor of 50–100, making it an indispensable tool for complex cell simulations with a large number of species types, cellular components, and high concentrations of metabolites (sugars, etc.) inside and outside the cell. Simulations enabled by this type of hybrid algorithm will allow researchers to study larger and more detailed systems, capturing the effects of reactions involving high particle count species such as metabolites, which have a crucial role in systems such as the genetic switch studied here. We have already used this hybrid approach to perform a spatially resolved RDME-ODE study (Fig. 2) of the galactose switch system, experiencing similar speeds up to what is seen in the CME implementation.

WHY BLUE WATERS

Blue Waters was essential to generate over 1,000 replicate hybrid simulations over the simulation time of 750 minutes and a range of concentrations. Only then did we have sufficient data to make the results statistically reliable and to determine the optimal communication time. In the worst-case scenario, the full CME simulations take nearly two days of wall-clock time, while the hybrid CME-ODE implementation with a communication time of one second requires less than 30 minutes. The response of the switch guided the setup for much more computationally costly RDME-ODE simulations on Blue Waters, which account for the spatial heterogeneous environment (nucleus, cytoplasm, membrane, etc.) of the yeast cell.

METHODS & CODES

The galactose switch system, with its four feedback loops and millimolar galactose concentration, is separated into a regime of species whose reactions will be simulated stochastically and another whose reactions will be simulated deterministically (Fig. 1). At the beginning of each timestep, the differential equation solver (DES) is updated with the species counts obtained from the stochastic regime (transcription, translation) simulated via the SSA. The DES then takes adaptive timesteps to evolve the high particle number species through time in the deterministic regime. At the conclusion of a timestep, the stochastic rates of reactions involving low particle number species interacting with high particle number species are updated with the species counts found by the ODE solver. At this time, the hybrid algorithm also communicates updated species counts generated from reactions in the CME regime to the ODE regime. The optimal communication times between the stochastic and deterministic descriptions, as well as the timesteps for each method, need to be assessed (Fig. 1) to verify that the hybrid description accurately describes the stochastic dynamics, which often have great impact on the cell’s behavior.

Figure 1: (top) The S. cerevisiae galactose switch system with ODE/CME separation. The reactions depicted (boxed area) are simulated continuously; those outside this region are simulated stochastically. (bottom) Distribution at steady state of the G3i species. As the communication timestep is decreased from 5 minutes to 1 second, the distributions given by the hybrid algorithm more closely match pure SSA results.

Figure 2: (top) A visualization of the cell compartmentalization that is imposed in a hybrid RDME-ODE simulation, with structural insight taken from tomograms. (bottom) The decrease in the amount of transcription factor present in the nucleus as the galactose concentration increases, matching expectations from experimental results.
QUANTUM-CLASSICAL PATH INTEGRAL SIMULATION OF CHARGE TRANSFER REACTIONS

Allocation: Illinois/100 Koch 175,000
PI: Nancy Makri
Co-PI: Peter L. Walters

EXECUTIVE SUMMARY
Quantum mechanical calculations on condensed phase or biological systems are prohibitively expensive. Mixed quantum-classical approximations are highly efficient and attractive. However, conventional quantum-classical methods introduce major assumptions in the treatment of the interaction between quantum and classical degrees of freedom. We have developed a rigorous quantum-classical formulation (QCPI) methodology that is free of assumptions and that treats all interactions in full atomistic detail. Our QCPI simulation on Blue Waters of the ultraviolet ferrocene–ferrocenium electron transfer reaction in liquid hexane, a standard in electrochemistry, yielded results of unprecedented accuracy and enabled the first quantitative demonstration of Gaussian behavior for a complex molecular solvent. Additional theoretical and algorithmic advances enabled simulation of reactions that span different time scales. In particular, simulation of this process in a series of solvents with increasing polarity exemplified the interplay between electronic and molecular time scales and energetics and its effects on the kinetics of charge transfer reactions.

RESEARCH CHALLENGE
Tunneling, quantum dispersion, and phase interference play a subtle but very important role in many chemical and biological processes. Quantum mechanical simulations of dynamical processes in the condensed phase continue to be extremely challenging because quantum mechanics is a nonlocal theory, implying an exponentially growing cost with the number of interacting particles.

For many processes of interest, quantum mechanical effects are essential only in the treatment of a small number of particles (e.g., a proton, or a small number of electronic states associated with a charge transfer reaction). The remaining particles (solvent molecules or biological medium) can be adequately described via classical dynamics. Unfortunately, the traditional Schrödinger formulation of quantum mechanics (which is based on delocalized wave functions) is incompatible with Newtontian trajectories (which are local in space). In the Schrödinger formulation, the interaction between the quantum and classical partitions cannot be treated without resorting to severe approximations.

METHODS & CODES
The focus of our work has been the development of a rigorous quantum-classical formulation based on Feynman path integral formulation of quantum mechanics. The local nature of the Feynman paths leads naturally to quantum-classical treatments that are free of ad hoc assumptions. The quantum-classical path integral (QCPI) methodology that we developed is a rigorous quantum-classical formulation. However, the QCPI expression appears impractical, as it contains an astronomical number of terms. Several advances in the understanding of interference and decoherence have recently made the QCPI methodology practical for the simulation of condensed phase reactive processes.

QCPI treats the small system of interest by full quantum mechanics, while the effects of the environment are captured via standard molecular dynamics (MD) procedures. Two widely used MD packages, NAMD and LAMMPS, are used to yield trajectories subject to forces obtained using the coordinates of the charged particle, which are specified by the given path. The dynamics captured along a classical trajectory augments the Hamiltonian of the quantum system through a time-dependent term, which leads to level fluctuations and eventually to decoherence. The exponential proliferation of trajectories with propagation time is avoided through a tensor decomposition that exploits the memory-quenching effects of condensed phase environments. The QCPI algorithm is characterized by classical molecular dynamics scaling and is fully parallelizable.

RESULTS & IMPACT
The QCPI methodology has enabled the simulation of charge transfer reactions in solution with unprecedented accuracy. The dynamics of the prototypical ferrocene–ferrocenium electron transfer pair was studied in a series of organic solvents of increasing polarity. Our first simulation of this electron transfer reaction in liquid hexane employed 1,230 atoms interacting via CHARMM force fields. More recent work has studied the same reaction in a series of solvents of increasing polarity.

The all-atom QCPI simulations shed light on the complex interplay among molecular/solvent time scales, electronic couplings, and reorganization energy, particularly on the way that these effects determine the rate of the reaction, the nature (exponential or nonexponential) of the dynamics, and the validity of Gaussian response. They also demonstrate how the interference among quantum mechanical phases leads to decoherence, and reveal substantial quantum delocalization of the otherwise classical solvent as a result of its interaction with the quantum electron transfer pair.

WHY BLUE WATERS
The QCPI formulation is well suited to a decomposition based on multi-level parallelism, and Blue Waters provides the ideal platform for its implementation. Specifically, the set of system paths is distributed across nodes; one processor within each node is assigned to the quantum mechanical calculations, while the other performs supporting trajectory computations. Moreover, because the trajectories are independent and generally relatively short, it is possible to assign a single trajectory to each core within a given processor while maintaining computational efficiency. This multi-level approach has the benefit of minimizing communication time while maximizing concurrent processing, since related classical and quantum-mechanical calculations are performed within the same node, where inter-processor communication should be much faster than if the information were more widely distributed.

PUBLICATIONS AND DATA SETS


PATIENT-SPECIFIC HPC MODELS AND SIMULATION-BASED IMAGING FOR CARDIOVASCULAR SURGICAL PLANNING

EXECUTIVE SUMMARY

Modern computational technology provides a platform to take Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) images and construct patient-specific geometric models for high-performance computing (HPC) as well as for 3D printing. Computed data obtained from numerical simulations can be converted into graphical images and used for clinical diagnostics as well as for vascular surgical planning. Geometric models can be used for developing StereoLithography (STL) files for 3D printing and generating patient-specific prototypes. In this project we focused on HPC aspects and extended and applied our methods [1,2] for blood flow modeling to patient-specific geometries.

RESEARCH CHALLENGE

The basic scientific and medical question that this research is aimed to address is to determine if computational models can provide insights into identifying regions of significant atherosclerotic plaque, luminal narrowing, and loss of wall elasticity. An important aid in the vascular access decision-making process in Transcatheter Aortic Valve Replacement (TAVR) patients is the ability of reproducing three-dimensionally the aortic, iliac, and femoral artery bifurcations from the data obtained from CT angiograms. This research helped in developing software to generate STL files for use in 3D printing. A third facet of investigation was to investigate if CT scan-based virtual models can serve as a “Virtual Patient” and provide an inside view of the arteries, which otherwise is available only through angiography.

METHODS & CODES

Automated generation of patient-specific models from CT scan images for cardiovascular models and blood flow analysis is still a bottleneck in the application of HPC to clinical applications. This research was focused on developing a framework and software that can take CT scan images to create high-fidelity patient models. Since patient-specific calculations involve uncertainty in the data, a series of numerical simulations were carried out based on statistical distribution of the data. Furthermore, blood artery interaction models are computationally expensive and therefore Blue Waters resources were needed to further explore the mathematical attributes of our non-Newtonian constitutive models as well as the coupling schemes for blood–artery interaction.

RESULTS AND IMPACT

Recent advances in computational fluid dynamics and image-based modeling permit determination of flow and pressure from CT scans, without the need for additional imaging, modification of acquisition protocols, or administration of medications. To analyze the massive amount of data in HPC one needs simulation-based images for clinical diagnostics. The software developed by Sheehan can help bring HPC and patient modeling for use by clinicians and surgeons.

WHY BLUE WATERS

Blue Waters was critical for both the development of cutting-edge software and the application of this software to perform large-scale biomechanics simulations. From a computational and algorithmic perspective, the newly developed coupled hierarchical multiscale methods tremendous benefit HPC by exploiting its large local resident memory on the processing nodes in favor of reducing the size of the global problem to be solved.

Figure 1: This is the CT Image of the branch structure created using the program developed by students, surface model of the patient, and the computational grid.

Figure 2: This is the pressure field at two instants in a typical cardiac cycle: (a) peak systole, (b) early diastole.

PUBLICATIONS AND DATA SETS


QUANTUM EFFECTS OF PROTON TRANSFER IN BIOLOGICAL SYSTEMS

Allocation: Innovation and Exploration/20 Koh
PI: James W. Mazzuca1

1Alma College

EXECUTIVE SUMMARY

Quantum effects have been shown to affect chemical reactions in a variety of condensed-phase chemical systems, including certain biological enzymatic reactions. In this study, we investigate how the quantum effects of hydrogen and its isotopes affect the rate of charge transfer in a model biological system in which a water molecule acts as an intermediary between the donor and acceptor molecule. Our calculations show an abnormally high kinetic isotope effect, which is indicative of a quantum mechanical tunneling mechanism, and we see dramatically different behavior when these quantum effects are turned off. This study makes a strong case for examining larger, more realistic biological systems using this approach in the future.

RESEARCH CHALLENGE

In most chemical dynamics calculations, it is assumed that all nuclei are classical particles and all electrons are quantum particles represented by molecular orbitals. This approach is typically quite good, especially for high-temperature chemistry and heavy nuclei, or basically anything other than hydrogen. There are certain cases, however, where the quantum mechanical effects of nuclei become important. Such cases include condensed-phase chemistry in which a hydrogen/proton transfer is the rate-limiting step. There is evidence to suggest that many biological chemical systems fit into this category, and we examine a model system in this study to determine how these quantum effects influence the mechanism of charge transfer.

These quantum effects can potentially impact both the efficiency and overall mechanism of charge transfer. The primary challenges in this field revolve around how exactly one should include the quantum effects of nuclei, but also how the electronic structure should be evaluated as the chemical system evolves from a reactant to a product state. In this particular study, we chose a model that we believe will capture the important physics of the proton transfer mechanism, but it is also simple enough that we can extract the important properties of this mechanism before examining larger, much more complicated systems.

Our calculations can give us a close look at exactly how this charge transfer process occurs, and we can use this information to improve our fundamental understanding of how nature has taken advantage of quantum mechanics in a way that has a profound influence on the chemistry of living organisms.

METHODS & CODES

We employed the ring polymer molecular dynamics (RPMD) method to include quantum effects of proton motion in our system. This method was chosen because it has been shown to be accurate for a wide range of applications, it scales linearly with respect to the system size, and the calculations can be done in parallel across many cores. Our code primarily simulates the physics of our chemical system as it evolves from a reactant state to a product state. To this end, many simulations containing many particles need to be performed in a highly parallel computing environment. Our group developed a Fortran-based code, which was highly scalable and well-suited for supercomputing environments such as Blue Waters.

RESULTS & IMPACT

Our calculations show an abnormally high kinetic isotope effect (KIE) for proton transfer in this model system. The kinetic isotope effect, in this case, is a measure of how much faster the reaction occurs when a proton, rather than a deuteron, is the charge carrier. In standard classical chemical dynamics calculations we observe a KIE of approximately 1.4 for all temperatures. In our quantum mechanical calculations we see a KIE of 35 at room temperature, and an even higher KIE at lower temperatures. This indicates that quantum tunneling of nuclei plays an important role in this charge transfer mechanism, and a physically correct picture of this process depends on the inclusion of these quantum effects.

These calculations have provided us with a very close and detailed view of a physical phenomenon that plays an important role in many biological systems and would not have been possible without highly parallel supercomputing resources. The seemingly obscure phenomenon of quantum mechanical tunneling has been shown, in this study, to profoundly impact the rate of charge transfer in a simple model biological system, and this has similar implications for more complicated but fundamentally similar systems that are found in many different natural biological systems.

WHY BLUE WATERS

The RPMD method relies on taking averages among many chemical dynamics simulations with different initial conditions. Often, we need to perform hundreds of thousands of calculations in order to see converged results. Fortunately, these jobs can be run in parallel with one another, and a supercomputing environment in which we can run massively parallel jobs over thousands of cores is ideal. Additionally, the scalability of this method typically means that if enough compute cores are available, no job should take more than 30 minutes. The efficient parallelization, short job queues, and responsiveness of the system made Blue Waters a perfect platform for our code development and job submission. Additionally, the Blue Waters Student Internship Program provided a perfect gateway into using these resources while funding an undergraduate student’s research experience on Blue Waters.

Figure 1: The image shows how potential energy changes as a proton is passed from a donor to acceptor molecule, with the help of an intervening water. The arrows represent two paths that the proton may travel to reach the product state, and quantum effects of this phenomenon were included using the ring polymer molecular dynamics (RPMD) method. Image from J. Phys. Chem. A, 121, pp. 819–826, 2017.

Figure 2: The inclusion of quantum effects in the proton transfer mechanism shows an extremely high kinetic isotope effect when the rate of proton transfer is compared to the rate of deuteron transfer, and this points to the importance of quantum effects in the proton transfer process. Image from J. Phys. Chem. A, 121, pp. 819–826, 2017.

PUBLICATIONS AND DATA SETS

EXECUTIVE SUMMARY

Proton-coupled oligopeptide transporters (POTs) use the inwardly directed proton flow to uptake small peptides and peptide-like molecules. The human POT transporters PepT1 and PepT2 provide the main route through which the body absorbs and retains dietary proteins. Human POTs also recognize several important families of peptide-like drug compounds such as β-lactam antibiotics. In order to function, POTs undergo large-scale conformational changes, whose characterization is the key in understanding the mechanism of transport by these proteins. The inward-(IF) to outward-facing (OF) structural transition of POTs, however, has remained elusive despite many experimental and computational efforts. We have employed all-atom molecular dynamics (MD) simulations along with novel enhanced sampling techniques to, for the first time, characterize the large-scale conformational changes of a bacterial POT transporter, namely GkPOT. By employing novel loosely coupled multiple-copy (LCMC) algorithms, our simulations provide a detailed description of the GkPOT conformational landscape, which sheds light on the structure–function relationship in POTs.

RESEARCH CHALLENGE

Membrane transporters provide the machinery to intimately couple active transport of molecules to various forms of cellular energy. POT transporters couple the energy from proton flow to the transport of dipeptides, tripeptides, and their analogs [1]. A key feature of POTs is their substrate promiscuity [2], which is of great interest from a biomedical perspective. Human POT transporters PepT1 and PepT2, which play a key role in absorbing and retaining dietary proteins (in the small intestine and kidney, respectively) [3], recognize several important families of peptide-like drugs such as β-lactam antibiotics [4] and can improve the uptake of poorly absorbed/restricted medications if attached to amino acids or dipeptides (prodrugs) [5]. Recent structural studies have resulted in several crystal structures of bacterial POTs [6–10], among which GkPOT, the POT transporter found in the bacterium Gambicillus kaustophilus, has the highest resolution (1.9 Å) [10]. These crystal structures, which are in the IF state, provide the basis of our understanding of POTs’ transport mechanism at the structural level. However, in order to function as active transporters, POTs are known to alternate between distinct IF and OF states. The conformation of the OF state and the transition pathway between the two functional states have remained elusive.

Previous MD simulations that relied on equilibrium simulations have failed to reliably characterize large-scale conformational changes such as those involved in POTs [10–11]. While the conventional MD can provide information on local conformational changes of a protein upon binding or unbinding of a substrate, ion, or proton, the global conformational changes observed are not often statistically significant. Functionally important conformational changes such as the IF–OF transition in membrane transporters typically occur on timescales beyond those accessible to conventional all-atom MD. The large-scale conformational changes, on the other hand, are typically studied using simplified modeling techniques such as coarse-graining, which could completely ignore or misrepresented the role of chemical events in the transport process. The main challenge in characterizing the large-scale conformational changes of proteins such as those associated with GkPOT is to reach the functionally relevant timescales without compromising the chemical details.

METHODS & CODES

We used a novel ensemble-based simulation approach [12–15] to reconstruct the entire transport cycle of GkPOT. Bias-Exchange Umbrella Sampling (BEUS) and String Method with Swarms of Trajectories (SMwST) are two independent LCMC algorithms employed. Both methods require parallel execution of hundreds of MD simulations of large protein systems with explicit representation of water and membrane, which requires hundreds of nodes. Our methodology was partly based on the techniques developed and used to investigate the thermodynamic cycle of the glyceral-3-phosphate transporter (GlpT) [14]. This is a rigorous and practical approach in characterizing large-scale conformational changes of proteins and their coupling to chemical events, which efficiently takes advantage of petascale computing. An important modification to the methodology to increase the accuracy of the results was introduced that involves a Riemannian formulation of free energy calculation and path-finding algorithms [15].

The software engine used for the simulations is NAMD, a highly scalable MD code implemented in Charm++, an object-based parallel computing system, which is used to develop massively parallelized computers of CPUs and GPUs. Multiple concurrent NAMD instances are launched with internal partitions of Charm++ and are located continuously within a single communication world. Messages between NAMD instances are passed by low-level point-to-point communication functions, which are accessible through NAMD’s Tcl-scripting interface.

RESULTS & IMPACT

The OF structure shown in Fig. 1 represents the first OF model of POT transporters, generated using our all-atom MD simulations in combination with LCMC algorithms discussed above. Our model is verifiably a stable OF structure since the subsequent equilibrium simulations show a water accessibility consistent with an OF state (see the water profiles in Fig. 1). Our simulations also suggest that the full IF–OF transition requires the binding of both proton and substrate (data not shown). The pathways generated using parallel SMwST [14] and modified using our recently introduced Riemannian formalism [15] reveal that the proton-bound GkPOT cannot transition to the OF state (data not shown). Unlike the common simulation studies, which other rely on unbiased equilibrium simulations or simple representations (e.g., coarse-graining), our approach combines the accuracy of all-atom MD with the accessibility of long timescales provided by enhanced sampling techniques. The successful employment of these cutting-edge multiple-copy algorithms using Blue Waters resources opens a new window to structural biology of membrane transporters that bypasses the limitations of computational approaches to study structure–function relationships in these proteins.

WHY BLUE WATERS

We have explicitly shown that the unbiased all-atom MD, which is routinely used in the field, could be quite misleading in deciphering mechanistic features of membrane transporters due to the great gap in the timescales associated with the conventional simulations and the function of these proteins [11]. On the other hand, we have shown that the LCMC algorithms such as BEUS and SMwST [14–15] can be used to reconstruct unknown conformational transitions of membrane transport proteins. Unlike the conventional all-atom or coarse-grained MD that can be performed on sub-petascale machines, LCMC MD simulations of membrane transporters are well suited for large petascale computational resources such as Blue Waters as they require hundreds of nodes for a single job. We note that the “weak scaling” of these algorithms (Fig. 2) makes them particularly attractive for large petascale machines, which can utilize hundreds of compute nodes with almost perfect efficiency.

PUBLICATIONS AND DATA SETS


The μ Opioid Receptor (μOR) is a G-Protein Coupled Receptor (GPCR) that mediates pain and is a key target for clinically administered analgesics (i.e., pain medicines). The current generation of prescribed opiate—drugs that bind to μOR—engender dangerous side effects such as respiratory depression and addiction due to the ligand-induced off-target conformations of the receptor. To determine both the key conformations of μOR to atomic resolution as well as the transitions between them, long timescale molecular dynamics (MD) simulations were conducted using the Blue Waters supercomputer. These simulations predict new and potentially druggable metastable states that have not been observed by crystallography. We used statistical algorithms (e.g., iICA and Transfer Entropy) to perform our analysis and discover key conformations from simulation, presenting a transferable and systematic analysis scheme. Our approach provides a complete, predictive model of the dynamics, structure of states, and structure–ligand relationships of μOR with broad applicability to GPCR biophysics and medicinal chemistry applications.

RESEARCH CHALLENGE

Because of its remarkable pain-reducing and induced-euphoria properties, opioids have been used recreationally and for medical purposes for more than 4,000 years. Unfortunately, an epidemic of opioid abuse has increasingly bedeviled the United States [1]. An efficient opioid (medicinally perfect) would be a potent pain reliever without side effects such as harmful respiratory effects or constipation, would show sustained efficacy in chronic treatments, and would not be addictive. The grand challenge questions are “How to design a perfect opioid?” and “How can we use the Blue Waters supercomputer to do that?”

FOLLOWING THE SOLUTION OF THE FIRST STRUCTURE, CRYSTALLOGRAPHY OF GPCRs HAS BOTH ILLUMINATED THE STRUCTURAL BIOLOGY AND EMPowered MEDICINAL CHEMISTRY OF THIS CLASS OF RECEPTORS [2,3]. RECENTLY, CRYSTAL STRUCTURES OF μOR ITSELF WERE SOLVED IN ITS “INACTIVE” AND “ACTIVE” CONFORMATIONS [4,5]. HOWEVER, OTHER BIOPHYSICAL AND PHARMACOLOGICAL EXPERIMENTS HAVE DEFINITIVELY DEMONSTRATED THAT μOR TRAVESSES MULTIPLE FUNCTIONALLY IMPORTANT CONFORMATIONAL STATES [6]. THESE STATES ARE IMPORTANT IN DESIGNING THE “PERFECT OPIOID” BUT ARE NOT TRACtABLE BY THE EXPERIMENTS AND CRYSTALLOGRAPHY.

USING BLUE WATERS, WE ATTEMPTED TO DISCOVER STATES OF μOR THAT HAVE DIFFERENT CONFORMATION COMPARED TO CRYSTALLOGRAPHY BUT ARE PHYSIOLOGICALLY SIGNIFICANT. IN ADDITION, WE ALSO TRIED TO UNRAVEL HOW OPIATES OF DIFFERENT SCAFFOLD CLASSES TUNE THE RECEPTOR TOWARD DISTINCT CONFORMATIONAL ENERGY LANDSCAPES.

METHODS & CODES

We performed multiple rounds of MD simulations on Blue Waters starting from the active with ligand, active APO, and inactive crystal structures. We used MDTran (the package developed in the Pande lab) to convert and assemble the trajectories. Then, we used the Conformation software package written for and applied to the

featureization of this large GPCR MD dataset. To summarize, all residue–residue pairs within 6.6 Angstroms measured by closest heavy atom distance in either crystal structure were selected. Then, for each of these approximately 2,200 residue–residue pairs, both the closest heavy atom distance and Calpha distance were computed for each frame in each trajectory, leading to 4,400 “features” for each trajectory frame. Next, the Sparse ICA algorithm was applied to determine the reaction coordinates, or slowest collective degrees of freedom (up to the 10 slowest in this case), of the protein. Finally, a Markov State Model (MSM) was constructed with a lag time of 25 nanoseconds and prior counts of 1x10. The equilibrium state probabilities from the MSMs were used individually in each condition to generate the free energy surfaces projected onto the features and ICA coordinates in Fig. 1.

RESULTS & IMPACT

Using Blue Waters, we mapped the complete free energy landscape of μOR (Fig. 1) and discovered the novel and significant conformation states not tractable by experiments (Fig. 2). Using the state-of-the-art machine-learning algorithms developed at the Pande lab (iICA and relative entropy), we reduced the high dimensionality and sheer number of data points that renders analysis so difficult. The important signal relaying residue switches we discovered (Fig. 2) shed new physical insight into the deactivation mechanisms and pathways of μOR. The newly discovered protein structures will be publically available for drug discovery projects.

We also produced a replicable framework that might enable other labs to gain actionable knowledge about their protein systems of interest through the “computational microscope” that is MD simulation. To this end, we published our simulations—both raw and factorized trajectories—as an open-source, downloadable resource, useful to both opioid researchers in particular as well as the wider structural biology and medicinal chemistry communities.

WHY BLUE WATERS

Blue Waters is an extremely powerful and versatile computational resource. In addition to powerful CPU and GPU hardware, the fast interconnect allows us to do types of calculations (rapid adaptive sampling, Markov State Model construction, force field optimization, etc.) that we could not do on other platforms such as distributed resources (e.g., Folding@home). Also, the availability of the NAMD (nanoscale molecular dynamics) simulation package on Blue Waters has particular advantages for adaptive sampling and restrained equilibrations.

PUBLICATIONS AND DATA SETS

EXECUTIVE SUMMARY

C-type inactivation of K+ channels is a molecular process of great physiological significance that affects the firing patterns of neurons in the central nervous system and the repolarization of cardiac cells in the heart. Despite years of studies, recent experimental results with semi-synthetic channels inserting unnatural amino acids in the structure suggest that a constricted filter conformation may not correspond to the C-type inactivated state. To resolve this issue, molecular dynamics simulations and free energy computations based on atomic models of the KcsA potassium channel were carried out. The computational results support the notion that the constricted conformation of the selectivity filter correspond to the functional C-type inactivated state of the KcsA channel. Our previous studies have put us on a strong path to execute this computational research [1–3].

RESEARCH CHALLENGE

Identifying the structural features associated with C-type inactivation in K+ channels is a very important goal and, in this regard, the bacterial KcsA channel has played a critical role. Although C-type inactivation has traditionally been associated with a functional behavior of voltage-gated channels such as Shaker, all the known functional markers of C-type inactivation are experimentally recapitulated by KcsA, providing a coherent view of the molecular determinants that affect inactivation. In addition, because the TTVGYGD canonical selectivity filter sequence is highly conserved, it is expected that the accessible conformational states of the selectivity filter, i.e., conductive and inactivated, ought to be fairly similar throughout most of the K+ channel family.

METHODS & CODES

We have carried out molecular dynamics simulations based on atomic models of the channels with the program NAMD using the CHARMM force field PARAM27.

RESULTS & IMPACT

Molecular dynamics simulations and free energy calculations were used to investigate the molecular mechanism of C-type inactivation. The impact of different inner-gate openings on the conformation and dynamics of the selectivity filter was systematically studied, and—for the first time—our 2D free energy calculation quantitatively characterized the conformational preferences of the selectivity filter in which a fully open gate (~23 angstroms) highly favors a constricted filter, whereas a partially open gate (~16 angstroms) notably prefers to maintain a conductive filter. The spontaneous, rapid, and consistent conductive-to-constricted transition observed in simulations with a fully open gate reveals a hydrogen-bond network controlling the cooperativity for the tetramer constriction. By contrast, the partially-to-fully open transition for the inner gate is much slower, and the rearrangement of inner helices demonstrates considerable conformational heterogeneity. Analysis based on the simulations provides comprehensive details of the long-range allosteric gating coupling, and, more importantly, reveals the molecular basis for the delayed kinetics of C-type inactivation. These recent results are described in manuscripts that are currently under review.

WHY BLUE WATERS

Blue Waters offers the ability to carry out extensive umbrella sampling computations with Hamiltonian replica-exchange molecular dynamics simulations (i.e., window swapping) using multiple copies of the system.

Figure 1: 2D PMF for the conductive-to-constricted transitions for the selectivity filter with different opening degrees of inner gate. The horizontal reaction coordinate r describes the width of the selectivity filter and is defined as the average cross-subunit pinching distance between the Cα atoms of G77. The vertical reaction coordinate z indicates the position of the external K+ ion along Z axis relative to the center of the selectivity filter. The lower panel is the one dimension PMF along horizontal reaction coordinate r, with integration of the vertical reaction coordinate z. The typical conformations for three free energy basin (left) are shown in stick for proteins and VDW representation for both water and K+ ion.
BLUE WATERS ANNUAL REPORT

STUDYING CELLULAR PROCESSES THROUGH THE COMPUTATIONAL MICROSCOPE

Allocation: NSF FACR/1354040
PI: Klaus Schulten (deceased); Emad Tajkhorshid (successor)
Co-Pi: Juan R. Perilla; James C. Phillips; John E. Stone
Collaborators: Peijun Zhang, Tatyana Polenova, Angela M. Greenenko, Christopher R. Aiken, Adam Zlotnick, Yale E. Goldman, Robert J. Woods

University of Illinois at Urbana-Champaign
University of Delaware
University of Oxford
University of Pennsylvania School of Medicine
Vanderbilt University School of Medicine
Indiana University
University of Pennsylvania School of Medicine
University of Georgia

In biology and in biomechanics, we have to realize that basically all organisms are large societies of molecules. We need a supercomputer to see that society for the first time.

—Klaus Schulten (1947–2016)

EXECUTIVE SUMMARY

Blue Waters provides a powerful platform at the interface of biology, physics, and computer science that is being leveraged to shed light on cellular processes toward direct impact on public health. Utilizing Blue Waters as a “computational microscope,” researchers are revealing key new insights into viral infection and the mechanisms by which it can be disrupted with drugs.

Virus capsids represent promising drug targets, and the world’s first microscopic molecular dynamics (MD) simulations of these remarkable molecular machines performed at all-atom resolution—a feat made achievable only by the petascale computing resources of Blue Waters—expose details underlying their biological functions and potential vulnerabilities to drug compounds.

RESEARCH CHALLENGE

Infectious viral pathogens are a major risk to public health, and millions of people die annually due to a lack of effective anti-viral treatments. The development of novel drug compounds that can target viruses depends heavily on characterizing the components of virus structure and the roles those components play in facilitating infection. One such structural component key to virus function is the capsid (Fig. 1), a protein shell that packages the viral genome and regulates its delivery to the host cell nucleus. Virus capsids are currently of great pharmaceutical interest as drug targets.

METHODS & CODES

MD simulations provide a powerful technique to investigate the dynamic structure and chemical–physical properties of virus capsids [1]. Our work has demonstrated that, when performed at the all-atom level of detail, simulations are capable of capturing even subtle effects on capsid structure and dynamics induced by bound drug molecules [2]. We employed Nanoscale Molecular Dynamics, or NAMD [3] for our simulations, a highly scalable MD code optimized specifically for Blue Waters that boasts a long and successful track record of deployment on the machine. While all-atom simulation of virus capsids comes at great computational expense, access to NAMD on Blue Waters has enabled us to reveal critical new insights into the structure and function of capsids, as well as to suggest mechanisms by which drug molecules can disrupt them. Importantly, our discoveries were accessible to state-of-the-art experimental methods, and were made possible only through access to the petascale computing power of Blue Waters.

RESULTS & IMPACT

HIV-1. Infection with human immunodeficiency virus type 1 (HIV-1) is classified as a global pandemic by the World Health Organization. Due to the extremely high mutation rate of the virus, new drug treatments must be constantly developed. We previously utilized Blue Waters to solve the all-atom structure of the mature HIV-1 capsid [1], providing an essential platform for study of its dynamical and structural properties, as well as its interaction with host-factors [3] and drug molecules [2]. At present, we have leveraged Blue Waters to perform the most monumental all-atom simulation achieved to date, characterizing the dynamical behavior of the HIV-1 capsid system (64 million atoms, Fig. 2a) over the timescale of one microsecond.

Our simulation reveals charge-specific channels in the surface of the capsid through which solvent ions translocate; these channels are likely capable of regulating translocation of DNA nucleotides, which must travel from the exterior to the interior of the capsid during reverse transcription. Further, analysis of the capsid’s collective motions reveals a belt dividing the capsid into two hemispheres suggesting a potential mechanism by which the capsid may break apart to deliver its genome. Our results, which additionally included characterization of the electrostatic and acoustic properties of the capsid, indicate new avenues for the development of drugs that seek to disrupt the capsid by altering its complex biological properties.

HBV. Hepatitis B virus (HBV) is a leading cause of liver disease worldwide, including cancer, and the World Health Organization estimates that 240 million people suffer from chronic infection. The capsid of HBV is icosahedral, and most prior structural studies imposed assumptions of icosahedral symmetry to enhance experimental resolution and reduce computational complexity. We previously utilized Blue Waters to perform the first simulation of the HBV capsid undertaken without symmetry bias and characterized drug-induced structural changes likely related to the drug’s mechanism of capsid disruption [2]. At present, we have leveraged Blue Waters to perform the most extensive unbiased simulations achieved for an icosahedral virus capsid to date, characterizing the dynamical behavior of the HBV capsid system (6 million atoms, Fig. 2b) in the presence and absence of three distinct drug compounds over timescales of one microsecond.

Our simulations reveal remarkable asymmetry in capsid motions, supporting hypotheses that the capsid can distort asymmetrically to accommodate unevenly-distributed internal strain resulting from conversion of pgRNA (pre-genomic ribonucleic acid) to DNA during reverse transcription. Further, the localization of ions during simulation provides a structural explanation for experimentally observed enhancement of capsid assembly under high salt concentrations. Finally, our results capture a variety of morphological and allosteric changes induced by bound drug molecules, providing insight into their complex mechanism of action.

Viral host factors. Integrating the results of additional simulations with experiments performed by our collaborators, we also reveal key details underlying the structures and mechanisms of large host-cell molecules that are implicated in viral infection, including the myxovirus resistance protein B (HIV-1) and cytoplasmic dynein (HIV-1 and HBV).

WHY BLUE WATERS

Due to their formidable computational expense, microscopic simulations of virus capsids are only possible on a petascale machine like Blue Waters. Capsid systems encompass millions of atoms, and computing the interactions among such large numbers of particles over such long timescales can take months, even on tens of thousands of processors. Further, analysis of the colossal data sets generated by our capsid simulations was made feasible only through access to the massively parallel computing power and high-performance Lustre filesystem provided by Blue Waters. The exciting discoveries revealed by our research underscore an essential role for petascale resources like Blue Waters in the development of anti-viral treatments, and demonstrate that access to leadership-class computing facilities holds the potential for significant impact on overall public health.

PUBLICATIONS AND DATA SETS


Figure 1: Free energy landscape of CDK and CMGI kinase indicating the key differences in their free energy landscapes.

Methods & Codes
To investigate the allosteric effects of CDK and GPCRs, we performed extensive molecular simulations of both proteins using Blue Waters. The simulation data were clustered based on the kinematically relevant state, and a Markov state model (MSM) for each system was constructed.

Results & Impact
Our MSM reveals that kinase activation is a highly complicated dynamic process that involves multiple intra- and intermolecular switches that regulate the kinase conformational preferences. We hypothesized that in the absence of cyclin, at least one regulatory switch in CDK2 is "off" while it is "on" in CMGI. To find the suitable regulatory switches, all available crystal structures of CDK2 have been compared. Four possible molecular switches have been identified including the conserved K-E bond, alignment of the R-spine residues, the helical twist in the A-loop, and availability of DFG-loop aromatic residues for substrate interaction.

Formation of a helical region in the beginning of the activation loop is an important characteristic of inactive CDK structures, which prevents binding of the substrate protein (PDB ID: 3PXR and 3PXR). The helical turn pushes the αC-helix out, thereby acting as a molecular switch that could alter the cyclin dependence of CDK kinases. This auto-inhibitory mechanism is observed in several kinases, including CDKs, Src and Abl [2]. The crystal structural analysis shows high degrees of correlation between the presence of the helical turn and the K-E bond. The simulation results mapped into a two-dimensional conformational landscape of K-E-distance versus the root mean-squared deviation of the helical turn with respect to the inactive structure (PDB ID: 3XPR) shows a barrier of ~6 kcal/mol for unfolding of the helical turn in CDK2, whereas the barrier is not observed due to a stable intermediate state in CMGI (Fig. 1 b, respectively). The helical secondary structure moves from the beginning of the A-loop toward its end in the β intermediate state in CMGI (Fig. 1 b), which helps the A-loop to unfold with a lower barrier. These free energy landscapes support the hypothesis that the differences in the stability of the helical turn between CDK2 and CMGI are one of the factors causing differences in their activation process.

Our MSMs built from GPCR simulation data provide kinetic data unobtainable from experiments describing ion-binding pathways in addition to the crucial residues that drive the ion-mediated interaction. The sodium ion is recognized by residues at the extracellular loop in all other GPCRs except S3P and LA 1R as the ion interacts with extracellular residues involving different helices. The residues slowly diffuse to the primary ligand binding site and form stable interactions with Asp64 (Ballistero-Weinstein numbering). In LA 1R, the ion occupies the orthosteric site by interacting with Glu67 and Tyr165 in S1P. Finally, the ion reaches the allosteric site and establishes stable interaction with Asp64. The free energy profiles of sodium ion binding have provided the various degrees of ion-binding specificity across the GPCR family. M3 and A2AAR bind more specifically to sodium ions compared to other GPCRs. The ion forms a stable, extended interaction at the allosteric site and establishes stable interaction with Asp64. The free energy profiles of sodium ion binding for the first time using MD simulation. The sodium binding site can be used as a potential allosteric drug-binding site to lock the receptor in the GPCR inactive state to restrict the downstream signaling mechanisms. Our results provide molecular-level details of the specificity of ion binding to various GPCRs.

WHY BLUE WATERS
The understanding of long-range network interaction in proteins requires several hundred microseconds-long simulations. Blue Waters provides the necessary computer architecture needed to carry out these computational studies. The current CPU and GPU framework allows us to run hundreds of parallel simulations. Without Blue Waters, the current work would not be possible.
DATA-DRIVEN, BIOLOGICALLY CONSTRAINED COMPUTATIONAL MODEL OF THE HIPPOCAMPAL NETWORK AT FULL SCALE

Allocation: NSF PRAC/707.6 Koh
PI: Ivan Soltesz
Collaborator: Michael Hines

1Standard University
2Yale University

EXECUTIVE SUMMARY

We have extensively validated a first-of-its-kind 1:1 scale, strictly biological data-driven computational network model of the CA1 region of the rodent hippocampus. The model spontaneously generates theta oscillations, which occur in the brain during locomotion and spatial navigation. Theta oscillations are critical for ensuring that our computational models can be comparable to behavioral experiments. Simulations of our CA1 model of 10 seconds of physical time took 14 hours to run on 1,024 Blue Waters nodes. It would not have been practical or affordable to conduct simulations of such scale on other publicly or commercially available computational platforms.

METHODS & CODES

Our principal simulation environment is NEURON 7.4 [1]. NEURON is designed to simulate neuronal models that are described in terms of the membrane properties and geometric structure of neurons [2], and supports computationally efficient representation of connections among neurons in a network [3]. NEURON is formulated around the notion of continuous cable “sections” that can be connected together to form any kind of branched cable. A section can be assigned properties that vary continuously with position along its length. User-defined biophysical properties of membranes, such as ion channel dynamics, are described in terms of differential equations, kinetic schemes, and sets of simultaneous equations. These model descriptions are compiled to C, so that membrane voltage and gating states are computed efficiently using an implicit integration method. These model descriptions, as well as the intrinsic properties of each interneuronal type, are compiled to C, so that membrane voltage and gating states are computed efficiently using an implicit integration method. The CA1 model that we have completed during the project period is capable of spontaneous generation of theta and gamma oscillations, replicating published studies. In addition, we have analyzed the firing patterns of each neuron type in the network. Next, we have conducted extensive work on perturbing various network and neuronal parameters to determine the relative contributions of each component to the generation of biophysically relevant oscillations. The model neurons exhibited firing patterns comparable with those of the rodent hippocampus.

RESULTS & IMPACT

We have made major advances toward achieving our aims. Specifically, we have extensively validated a first-of-its-kind 1:1 scale, strictly biological data-driven computational network model of the CA1 region of the rodent hippocampus. The model spontaneously generates theta oscillations, which occur in the brain during locomotion and spatial navigation. In addition, we have studied large numbers of perturbed network configurations in order to determine the range of parameters and conditions that are necessary and sufficient for the emergence of the theta oscillations. Furthermore, we have advanced our technique for generating biologically realistic dendritic trees and applied it toward constructing a 1:1 scale model of the dentate gyrus (DG), another major hippocampal region and a key component of our modeling efforts. We have conducted a number of control simulations to validate the dynamic behavior of this model and tuned its synaptic parameters so that its spatial coding properties are consistent with recently published experimental results. These advances, which have never been achieved in computational neuroscience before, are critical for ensuring that our computational models have capabilities for processing spatial information that are comparable with those of the rodent hippocampus.

Finally, we have performed initial simulations of September 2014 and can be fully parallelized [2]. NEURON is very well optimized for branched structures [2]. NEURON is very well optimized for branched structures. A typical behavioral experiment with animals running on a linear track has a duration of tens of seconds, and therefore our simulations must be long enough to reflect the brain’s changing representation of space during locomotion. A typical behavioral experiment with animals running on a linear track has a duration of tens of seconds, and therefore our simulations must be long enough to reflect the brain’s changing representation of space during locomotion.

Figure 1: Structure of the CA1 network model. (A) The model network has dimensions identical to the in vivo hippocampus. (B) Model cells before and after in their respective layers. (C) Detailed morphology and enaque placement for each model neuron. (D) Illustration of the connectivity types. (E) Characteristics of each connection are constrained by experimental data.
STRETCHING THE CADHERIN MOLECULAR VELCRO® OF CELL-CELL JUNCTIONS

EXECUTIVE SUMMARY

The extracellular domains of classical cadherins form a Velcro®-like surface that glues cells together in the presence of calcium. This is essential for cell-cell adhesion and multicellular life, but the mechanics of multi-cadherin complexes is poorly understood. Using VMD (visual molecular dynamics) and NAMD (nanoscale molecular dynamics) on Blue Waters we were able to perform atomistic simulations of realistic adhesive systems that included up to twelve cadherin-cadherin bonds with 3.7 million atoms. These simulations revealed how cadherins respond to mechanical stimulation that mimics physiological forces, such as those experienced by tissue stretched by blood pressure, muscle movement, or impact with a foreign object. In addition, we were able to determine the collective behavior of cadherins in the absence of calcium and the relaxation of partially ruptured cadherin lattices as it may occur during wound healing. These simulations shed light on the basic molecular mechanisms that underlie tissue development, mechanics, and repair.

RESEARCH CHALLENGE

Selective and robust adhesion between cells is essential for multicellular life. Classical cadherin proteins are found on the surface of cells and act as molecular Velcro® that glues adhesive cells together [1,2]. The adhesion mediated by cadherins depends on calcium ions and is the basis for the formation of organs and the maintenance of tissue integrity in humans and in multiple species across the animal kingdom. The molecular architecture of a key cadherin protein named E-cad revealed how its extracellular domain protrudes from the cell surface as a hook that engages with another E-cad from an adjacent cell [3]. This E-cad/E-cad complex effectively forms a bond that links adjacent cells together in cell-cell junctions. E-cad also has a transmembrane helix and a cytoplasmic domain that anchors the protein to the actin cytoskeleton, thus effectively forming a bond that links adjacent cells together in cell-cell junctions. To test reversibility, we released forces and observed the recovery of curvature for individual E-cad molecules, hinting at the steps required to reestablish cell-cell adhesion after rupture. We also explored the behavior of the lattice in the absence of calcium, which resulted in disordered and floppy E-cad chains. Additional simulations will explore the elastic response of the calcium-free lattice as well as of lattices that are not anchored to the cytoskeleton or that are subjected to shearing stress. Overall, NAMD simulations of cadherin lattices are providing an unprecedented atomistic view of the mechanics of cell-cell adhesion.

METHODS & CODES

To test how an E-cad lattice responds to forces that mimic the effect of cells being stretched by blood pressure, muscle movement, or impact with a foreign object, we generated atomistic models with lattices containing one to twelve E-cad/E-cad bonds [3]. Mechanical properties of cadherin proteins are best studied in simulations where all atoms in the system, including critical calcium ions, are explicitly modeled, and where the system is hydrated with water molecules and ions that mimic the native physiological environment of cell-cell junctions. Such systems were built and simulated with VMD and NAMD [8,9], a pair of programs for molecular visualization and dynamics simulation created by the Theoretical and Computational Biophysics Group at the University of Illinois at Urbana-Champaign. This software suite can handle large atomistic systems, which in our case encompassed up to 3.7 million atoms. NAMD can efficiently use thousands of cores in Blue Waters, which allowed us to simulate these large systems and use steered molecular dynamics (SMD), a technique in which forces are applied to proteins to test their mechanical response in silico [10].

RESULTS & IMPACT

Systems including one, four, and twelve E-cad/E-cad bonds in a lattice arrangement were simulated for twenty or more nanoseconds, in equilibrium and under tension. To make the systems realistic, we also incorporated constraints that mimic cytoplasmic attachment to the cytoskeleton. To the best of our knowledge, these are the first all-atom SMD simulations revealing detailed dynamics of a complete and realistic cadherin lattice. We were able to observe how the initially curved E-cad/E-cad bonds became straight upon application of force, suggesting that cadherins at the cell-cell junction can act as molecular shock absorbers that extend without the E-cad/E-cad bond breaking at low force. As tension increased, we observed E-cad/E-cad bond rupture without any unfoiling of the individual E-cad molecules, despite lateral interactions that stabilized the lattice. To test reversibility, we released forces and observed the recovery of curvature for individual E-cad molecules, hinting at the steps required to reestablish cell-cell adhesion after rupture. We also explored the behavior of the lattice in the absence of calcium, which resulted in disordered and floppy E-cad chains. Additional simulations will explore the elastic response of the calcium-free lattice as well as of lattices that are not anchored to the cytoskeleton or that are subjected to shearing stress. Overall, NAMD simulations of cadherin lattices are providing an unprecedented atomistic view of the mechanics of cell-cell adhesion.

WHY BLUE WATERS

Molecular dynamics simulations of large atomistic systems are extremely computationally demanding and cannot be divided into smaller independent simulations to be distributed among poorly networked computational resources. Only a fast networked and massively parallel system like Blue Waters can be used to achieve the multi-nanosecond time scales relevant to study the dynamics and elasticity of cadherin complexes.
EXECUTIVE SUMMARY

There is direct evidence that commercial air travel spreads common infectious diseases including influenza, SARS, tuberculosis, and measles. This has motivated calls for restrictions on air travel, for example, during the 2014 Ebola outbreak. Such restrictions, however, carry considerable economic and human costs. Ideally, decision-makers ought to take steps to mitigate the likelihood of an epidemic without imposing the above costs. Science-based policy analysis can yield useful insight to decision-makers.

The effectiveness of a policy depends on human response to it. Given inherent uncertainties in human behavior, we simulate a variety of scenarios and identify the vulnerability of policies under the potential scenarios. Supercomputing is used to deal with the large number of scenarios and the need for a short response time in case of national emergencies. Our results identify new boarding procedures that can result in substantial reduction in the risk of the spread of Ebola and SARS.

RESEARCH CHALLENGE

Our goal is to develop models and a novel methodology that can provide insight to decision-makers on policies and procedures that will reduce the likelihood of infection spread during air travel. In addition, our research contributions promise major advances in the field of infectious disease modeling. Supercomputing is used to deal with the large number of scenarios and the need for a short response time in case of national emergencies. Our results identify new boarding procedures that can result in substantial reduction in the risk of the spread of Ebola and SARS.

METHODS & CODES

We modeled pedestrian movement during air travel as particles based on a force-field approach proposed by Helbing, et al. [1]. Both pedestrian density and speed of immediate neighbor in a pedestrian line determine pedestrian speed and trajectory [2–3]. Our modifications incorporate these aspects into the pedestrian movement model. The pedestrian trajectory information is then integrated with a discrete-time stochastic Susceptible–Infectious (SE) model for infection transmission. When (SE)Infectives come into contact with $m$ susceptibles estimated by the pedestrian movement model, the newly infected and the probability of their infection can be estimated using a binomial distribution approximated as Poisson distribution. This accounts for demographic stochasticity and variations in susceptibility of the population. This approach (see Fig. 1) provides insight into the consequences of policy choices that change a passenger’s behavior. We input this information to a global phylogeography model to assess the impact of these policies at global scale. Several airports (Phoenix, Tallahassee, and Daytona Beach) have expressed interest in collaborating with us on developing feasible policies.

Since inherent uncertainties in human behavior and insufficient data during the initial stages of an epidemic make prediction difficult, we parameterize the sources of uncertainty and evaluate vulnerability under different possible scenarios. Even if individual parameters cannot be reliably estimated due to identifiability issues, it may still be possible to precisely estimate a compound quantity of interest, such as the reproductive numbers. We use Blue Waters to deal with the computational load that arises from a large parameter-space.

We link the results of this fine-scale model with a phylogeography model, which uses genetic mutation information and geographic locations of viruses to model the spread of epidemics across large geographic scales. For the phylogeography work, we used Blue Waters to analyze 264 full genome Ebola sequences from Guíana, Liberia, Sierra Leone, Italy, the United Kingdom, and the U.S, then used the REAST software on Blue Waters to implement the phylogeography model. As part of this, we specified a generalized linear model (GLM) to study predictors of spread between discrete locations. We used five predictors: worst-case (infection) scenario, best-case scenario, airplane passenger flux between airports, and sample size from the origin and destination. We ran multiple Bayesian Markov–chain Monte Carlo simulations to produce posterior estimates, and also implemented a traditional discrete phylogeography run with a GLM that evaluated the transmission rate between locations and estimated the routes with strongest support.

RESULTS & IMPACT

In prior work, we used the above approach with Ebola, studying the impact of different procedures for boarding, disembarkation, and seat assignment on infection spread. Our results show promise for substantial impact from such changes. For example, we showed that on a 182-passenger Boeing 757 airplane, random boarding can lead to substantial reduction in infection transmission compared with current zone-wise boarding (see Fig. 2). We obtained similar results showing the potential for changes in in-plane movement, deplaning procedure, seating arrangement, and plane sizes in reducing the likelihood of infection transmission. The improvements obtained for individual flights by these policy changes can be of substantial benefit over the course of an epidemic. Based on the transportation data from 2013, if unrestricted air travel were to have occurred during the 2014 Ebola epidemic, then the probability of generating 20 infections per month from air travel could have been reduced from 67% to 40% by using better pedestrian movement strategies. This could further be reduced to 13% by exclusively using smaller 50-seat airplanes. The approach is generic and can be used for other directly transmitted diseases such as SARS.

WHO BLUE WATERS

In a new emergency, due to lack of data, one usually needs to model for a variety of scenarios. This leads to a large parameter space of uncertainties, which requires a large computational effort. In addition, the models typically need fine-tuning, which leads to an iterative process where the model is repeatedly tuned based on results from its previous validation steps. Consequently, rapid turnaround time is critical, which requires massive parallelism. Such parallelism becomes even more crucial during the course of a decision meeting, where results are typically needed in the span of a couple of minutes.

Thus cloud computing is not effective in national emergencies because of the large parameter space and need for quick runtimes during policy decision meetings. After extensive optimization (with support from the Blue Waters team) the simulation time per run is currently around 20 minutes. This can be reduced further only by parallelizing an individual simulation, which is communication-intensive. Again, a cloud environment is not useful.

Among supercomputers available to us, only Blue Waters and Stampede provide the necessary computing power. However, Stampede’s queueing policy prevents large-scale parallelism except in an emergency. An optimized scalable infrastructure cannot be set up unless there is access to large-scale parallelism ahead of the emergency.

PUBLICATIONS AND DATA SETS


EXECUTIVE SUMMARY

There is a large and growing body of evidence to support the interplay of mechanical properties of the human brain being affected by and affecting the function of the human brain. The ability to quantitatively assess the structural integrity of the brain using noninvasive techniques such as magnetic resonance imaging (MRI) provides the necessary insight to understand the status and function of the living brain. Clinical MRI systems are capable of advanced quantitative imaging, both in high resolution and unique biomarkers, but they require significant technological development, tissue modeling, and computational resources. The Blue Waters environment has allowed for more rapid progress in all three areas through analysis of the computationally intensive inversion algorithms and validation of large experimental datasets. Two studies have been completed showing correlations between memory performance and hippocampal mechanical properties. Additionally, two ongoing studies look to characterize the structural-functional properties of the normal aging human brain.

RESEARCH CHALLENGE

The human brain changes throughout its lifetime, both in subtle and drastic ways. Subtle changes to the brain microstructure occur with lifestyle choices and normal aging. More drastic changes can come from traumatic injuries and neurodegenerative diseases. Quantitative MR imaging techniques provide the opportunity to interrogate the tissue microstructure for these changes, but traditional approaches require resolution on the order of the disruption of the microstructure, which is much smaller than the achievable spatial resolution of MRI. Magnetic resonance elastography (MRE) is a quantitative imaging method that is able to accurately estimate the viscoelastic material properties of the brain tissue, from the small grey matter regions (e.g., hippocampus) to large white matter (WM) regions. MRE relies on advanced imaging methods and a nonlinear inversion (NLI) to estimate the material properties. NLI is technically and computationally challenging but has shown it is sensitive enough to detect subtle changes in human brain structures during disease and even during exercise interventions in both healthy and diseased brains. Each brain requires the computational expensive step of NLI resulting in the need for significant resources for analysis of a single brain. Blue Waters enables running full studies of 10–50 subjects to answer specific questions about the relationship between brain material properties and function.

METHODS & CODES

MRE is a noninvasive quantitative technique for measuring the mechanical properties via the response of the tissue to externally applied shear waves. The full 3D mechanical waves are captured using synchronized motion-encoding imaging, then the material properties are estimated using an iterative, finite-element-based nonlinear inversion (NLI) of the heterogeneous, steady-state Navier equation [1–2]. The size of the computational problem initially required dividing the brain into smaller regions (called "subzones") for FEM meshing and optimization of properties at a local level, then reconstituting as a global solution. Blue Waters' combination of a large number of nodes, high-speed interconnect, and large memory systems enabled us to perform a parameter study of a wide range of subzone sizes to determine the dependence of the estimated mechanical properties on these parameters.

RESULTS & IMPACT

Four separate human studies were completed, or are underway, and focus on two areas of brain health: understanding the links between tissue mechanical properties and brain performance [3–4] and effects of normal aging on the microstructure of the human brain. We have built on previous findings with MRE in studies run at the Beckman Institute exploring the relationship between tissue viscoelasticity and memory to include an analysis of fitness effects, which are known to improve brain health. The positive relationship between hippocampal structure, aerobic fitness, and memory performance has been demonstrated at both the level of the lifespan in children and older adults; evidence of this relationship among young adults, for whom the hippocampus is neither developing nor atrophying, is less consistent. Studies have typically relied on hippocampal volumetry (a gross proxy of tissue composition) to assess individual differences in hippocampal structure. Microstructural differences in hippocampal integrity may exist even among young healthy adults when volumetric differences are either absent or not diagnostic of tissue health or cognitive function.

In the current study, we investigated the relationship between hippocampal viscoelasticity from MRE and cardiovascular health, and their mutual effect on relational memory in a group of healthy young adults. We replicated our previous finding that hippocampal viscoelasticity correlates with relational memory performance. We extended this work by demonstrating better aerobic fitness, as measured by VO2max, was associated with hippocampal viscoelasticity that mediated the benefits of fitness on memory function. In contrast, hippocampal volume did not account for individual differences in memory. Therefore, these data suggest hippocampal viscoelasticity may provide a more sensitive measure of microstructural tissue organization and its consequences on brain performance among healthy young adults.

We followed up on the findings of relationships among hippocampal viscoelasticity, memory performance, and fitness by using MRE in a study of exercise training in persons with multiple sclerosis (MS). Five subjects with MS completed three-month exercise training, and we observed this intervention significantly improved their hippocampal viscoelasticity. Further, we observed the change in viscoelasticity correlated with the improvement in memory performance with training. These results highlight the sensitivity of the MRE measures of brain health as they relate to cognitive performance. By using a highly sensitive neuroimaging battery that includes MRE to monitor therapy, we can better understand the mechanisms contributing to brain health and recovery.

Finally, Blue Waters enabled a parameter study of NLI to further develop the NLI algorithms to improve the capabilities of quantitative imaging via a careful characterization of the effects of inversion parameters on the reconstructed material properties. Using the capabilities of Blue Waters' network and local memory, the size of the subzones was increased up to 15 times the previous subzone size, previously impossible on other systems. Additionally, processor speed afforded an investigation into increasing the number of conjugate gradient estimations at the subzone level to ensure proper convergence of the material property estimates. This study was the first of its kind in the breadth and depth of inversion parameter characterization.

WHY BLUE WATERS

The Blue Waters system provides a unique resource for MRI because it allows for rapid analysis of the scanners and the low-latency memory transfers necessary for highly parallel computations. This is especially critical for quantitative, high-resolution MRI that inherently involves very large datasets and extensive computation. Blue Waters has provided the refinement of these techniques and their translation in research by significantly reducing computational burden.

PUBLICATIONS AND DATA SETS


A CRYSTAL BALL OF BACTERIAL BEHAVIOR: FROM DATA TO PREDICTION USING GENOME-SCALE MODELS

EXECUTIVE SUMMARY

The democratization of mass sequencing and profiling technologies has resulted in a plethora of data that can reveal how organisms are organized and function on a cellular level. The goal of this project is to integrate the millions of data points collected for a model bacterium, Escherichia coli, so we can build predictive models of its behavior in novel, untended environments. We used Blue Waters for two major tasks. The first was to apply novel deep-learning algorithms for making sense of proteomics samples to find what proteins are present in given experimental settings. The second was to run massively parallel simulations with genome-scale, integrative models that predict the omics expression and bacterial behavior in novel environments. This work led to novel algorithms for omics data processing, multi-omics modeling, and the most integrative predictive model of bacterial behavior, with forward predictions validated experimentally. Parallelization of these processes on the Blue Waters supercomputer enabled a precise data-processing pipeline, search of optimal model architecture, and hyper-parameter optimization.

RESEARCH CHALLENGE

One of the grand challenges of systems and synthetic biology is the development of an accurate predictive model of any organism. If we had such a tool at our disposal, we could interrogate the cellular states and molecular abundances that exist in general under a specific selection pressure or environmental setting. In the realm of synthetic biology, it would allow the in silico testing of synthetic circuits and the genetic engineering of the chassis cells so that they exhibit optimal behavior. However, there are many obstacles to achieving this vision, including the large complexity of the cellular organization and machinery, the lack of cohesive datasets, and models that are capable of integrating them into one system that is more than the sum of its parts.

WHY BLUE WATERS

Our large-scale simulation for multi-scale modelling as well as the size and complexity of our datasets (30 million data points from several platforms and molecular species) necessitate the use of high performance computing for this project. Furthermore, the effectiveness of GPU capabilities in Blue Waters for our application of deep neural network models both in multi-omics- and proteomics-specific projects allowed us to train large and complex models.

RESULTS & IMPACT

The predictive performance of the multi-scale model has shown substantial improvement over prior work as well as the ability to predict not only phenotypic information (growth, traits, etc.) but also genome-wide gene, protein, and metabolite expression. This brings us a step closer to having a crystal ball for prediction of bacterial states and behavior in novel environments. Having such a tool in our arsenal will allow the fast and inexpensive testing of experimental settings, which in turn will allow us to navigate the vast experimental space in search of suitable environmental conditions for phenotypes. From biotechnology to medicine, such capability has several applications and can be transformational if applied at an industrial scale.

METHODS & CODES

The proposed protein inference method (DeepPep) in the omics data-processing pipeline is based on convolutional neural networks with deep-learning techniques to efficiently train the complex model. DeepPep is written in torch/python, and the method was validated with the six external data sets (Sigma49 [1], UPS2 [2], 18Mix [3], Yeast [4], DME [5], HumanMD [6], and HumanEKC [7]). The genome-scale model employs a recurrent neural network and a constrained regression altogether to predict genome-wide responses layer by layer.

PUBLICATIONS AND DATA SETS


https://deeppep.github.io/DeepPep/
http://www.prokaryomics.com

Figure 1: Overview of the processing pipeline and genome-scale modeling for E. coli. There are three major steps: 1) omics profiling and data processing, 2) QA/QC for processed data and multi-omics integration, and 3) predictive modeling and analysis.

Figure 2: Performance evaluation of the proposed protein inference method (DeepPep) in the omics data-processing pipeline is based on convolutional neural networks with deep-learning techniques to efficiently train the complex model. DeepPep is written in torch/python, and the method was validated with the six external data sets (Sigma49 [1], UPS2 [2], 18Mix [3], Yeast [4], DME [5], HumanMD [6], and HumanEKC [7]). The genome-scale model employs a recurrent neural network and a constrained regression altogether to predict genome-wide responses layer by layer.
EXECUTIVE SUMMARY

A key step in the lifecycle of human immunodeficiency virus type-1 (HIV-1) is the production of enveloped particles containing viral proteins and genomes from infected cells. Experimental data suggest that the HIV-1 Gag polyprotein orchestrates this process through subdomains that specifically interact with the cell membrane, viral nucleic acid, or other Gag polyproteins. Nonetheless, the molecular details of this process have remained elusive. Using large-scale coarse-grained molecular simulations, enabled by the use of Blue Waters, we investigate a network of critical interactions that regulate the early stages of HIV-1 assembly, packaging, and budding (Fig. 1).

METHODS & CODES

To investigate the self-assembly of the immature protein lattice of HIV-1, we created large-scale coarse-grained (CG) molecular models of the relevant viral components and the cell membrane based on experimental data from collaborators [2,3]. CG models are computationally efficient representations, enabling computer simulations to be performed at time- and length-scales that are otherwise impossible. We performed molecular dynamics simulations using the LAMMPS software package (Sandia National Laboratories, USA) to examine the self-assembly of viral proteins, and the influence of the cell membrane and viral genome on the “budding” of viral particles. In collaboration with experimental colleagues from around the globe, we combined our simulation results with experimental data to better understand key aspects of the early stages of the HIV-1 lifecycle.

RESULTS & IMPACT

By combining computer simulations with fluorescence localization experiments [4], we elucidate the interactions that regulate HIV-1 viral assembly dynamics. Specifically, analysis of our results reveals the influence of nuclear acids and the cell membrane in promoting the aggregation of HIV-1 Gag polyprotein via a multi-step, self-correcting nucleation process. We also illustrate the functional importance of the N-terminal, C-terminal, and spacer peptide 1 (SP1) protein domains, which are each responsible for regulating different lattice qualities. These aspects are difficult or impossible to control precisely in conventional experimental approaches but are tractable in the context of our large-scale CG models. The success of our simulations, and the methodology we use to generate CG models of complicated biomolecules, suggest a simple and robust approach to the direct incorporation of experimental data into CG model generation. We envisage that our general methodology will be transferable to the study of other biomolecular processes and open up new avenues of research into phenomena that occur at physiologically relevant time and length scales. The insight gained by large-scale CG analysis can then be used to assist in the design of new therapeutic approaches to viral infection.

WHY BLUE WATERS

Despite the relative efficiency of CG models, our simulations require large numbers of individual molecules interacting over relatively long molecular time scales. Furthermore, in order to study the influence of viral RNA and the cell membrane on protein self-assembly, we must probe the behaviors of this system under a wide range of biologically relevant conditions. It was therefore crucial to the success of our project that we had access to very large amounts of computational power, and that power needed to be combined with superior network performance: the tightly coupled parallel nature of our simulations require latency and bandwidth characteristics that are simply not available in cloud computing. Blue Waters platforms thus presented a natural choice for our work; the combination of leadership-class computing capabilities with cutting-edge network hardware allowed us to successfully investigate a system of significant biomedical interest. Our large-scale CG models are generated using techniques developed for previous work on the Blue Waters platforms [5], and existing relationships with Blue Waters technical project staff greatly assisted in the deployment of novel simulations on Blue Waters.

PUBLICATIONS AND DATA SETS

SOCIAL SCIENCE, ECONOMICS, & HUMANITIES

ECONOMICS

SOCIAL SCIENCE

232  Policy Responses to Climate Change in a Dynamic Stochastic Economy

254  Enabling Redistricting Reform: A Computational Study of Zoning Optimization
POLICY RESPONSES TO CLIMATE CHANGE IN A DYNAMIC STOCHASTIC ECONOMY

Allocation: GLCPC/250 Keh
PIs: Lien Hauon,
Co-PIs: Kenneth Judd, William Brock, Thomas Hertel
Collaborators: Simon Scholtegger, T. J. Cananai, Carlos Rangel

1University of Chicago
2The Ohio State University
3Haver Institution
4University of Wisconsin
5Purdue University
6University of Zurich
7University of Minnesota

EXECUTIVE SUMMARY

We extended our DSICE (Dynamic Stochastic Integration of Climate and the Economy) framework for evaluating policy responses to future climate change, to control the global average temperature anomaly so as not to exceed 2°C, with a stochastic economic production process [1]. We find that the social cost of carbon should be significantly greater in order to meet the 2°C target with a high probability [2].

We are extending DSICE to incorporate spatial structure of temperature and economy, sea level rise, thawing of permafrost, partial competition and collaboration, and adaptation. Our preliminary results show that rich countries in the region above 30 degrees north latitude should enact higher carbon taxes than poor countries in the tropic region [3]. Additional extension of DSICE is to incorporate dirty/clean energy sectors. We find that we should significantly increase investment in clean energy and reduce investment in dirty energy, and very soon [4].

RESULTS & IMPACT

In the past year, we published three papers based on the support of Blue Waters. In the first paper, published in Nature Climate Change, we extended DSICE to study the impact of multiple interacting tipping points. It is a large-size stochastic dynamic programming model with up to 20 continuous-state variables and five discrete-state variables. The paper showed that the five multiple interacting tipping points increases the social cost of carbon nearly eightfold.

In the second paper, published in Quantitative Economics, we used Blue Waters to develop an NLCEQ method to solve huge-dimensional dynamic stochastic problems efficiently and in parallel. NLCEQ has also been implemented to analyze the effect of climate and technological uncertainty in crop yields on the optimal path of global land use by Cai, Judd, Hertel and their collaborators [7].

In the third paper, published in Operations Research, we used Blue Waters to develop a parallel algorithm that can solve super games with states, which model strategic interactions among multiple players.

Economics, particularly when integrated with climate change science, is a field with many problems that have the size and complexity that justify the use of massively parallel computer systems. Our work using Blue Waters has allowed some economists to attack those computationally intensive economics problems.

WHY BLUE WATERS

Our parallel algorithms use the master-worker structure, and the communications among the master and workers are frequent and of small or moderate sizes, so the high latency inherent in a commodity cluster limited the ability to solve large problems in a reasonable amount of time. Moreover, the number of tasks could be huge. For example, the largest DP problem (for DSICE) we solved using Blue Waters had 372 billion of such tasks. Blue Waters allows us to use MPI (Message-Passing Interface) and solve far larger problems efficiently, as we have already shown in our previous work using Blue Waters.

PUBLICATIONS AND DATA SETS


ENABLING REDISTRICTING REFORM: A COMPUTATIONAL STUDY OF ZONING OPTIMIZATION

We have developed a scalable computational tool for redistricting that synthesizes and organizes massive amounts of computation and data to evaluate redistricting schemes. The tool allows one to create high-quality maps and tailor them to notions of “fairness” and democratic rule. It can also be used as an evaluation tool by courts, advocates, and the public to ensure non-overlapping electoral maps.

The tool allows one to create high-quality maps and tailor them to notions of “fairness” and democratic rule. It can also be used as an evaluation tool by courts, advocates, and the public to ensure non-overlapping electoral maps. Specifically, we developed a scalable, parallel, evolutionary algorithm for redistricting that includes a set of spatial evolutionary algorithm operators to handle the costly spatial configuration of redistricting maps. These maps provide the basis for additional statistical analyses.

EXECUTIVE SUMMARY

The Supreme Court has found it difficult to identify a workable standard by which we might regulate gerrymandering. We lack sufficient notions of “fairness” and democratic rule. It can also be used as a tool to provide fair representation in Congress to all communities and interest groups. Gerrymandering occurs when districts are drawn in a manner that discriminates against a partisan or racial group. Both partisan and racial gerrymanders are commonly alleged. Despite broad disdain for the practice of gerrymandering, the Supreme Court has found it difficult to identify a workable standard by which we might regulate gerrymandering. We lack sufficient tools to analyze and synthesize redistricting data, in part, because the requisite computation is massive. Without the tools to quantify the effect of electoral maps, the court is left without the ability to issue legal and consistent judgments. As a result, despite the Supreme Court’s articulated legal reasoning and mandates, we have developed a computational redistricting tool.

RESEARCH CHALLENGE

In the United States, political redistricting occurs at the national level every 10 years following the decennial census. It is intended to provide fair representation in Congress to all communities and interest groups. Gerrymandering occurs when districts are drawn in a manner that discriminates against a partisan or racial group. Both partisan and racial gerrymanders are commonly alleged. Despite broad disdain for the practice of gerrymandering, the Supreme Court has found it difficult to identify a workable standard by which we might regulate gerrymandering. We lack sufficient tools to analyze and synthesize redistricting data, in part, because the requisite computation is massive. Without the tools to quantify the effect of electoral maps, the court is left without the ability to issue legal and consistent judgments. As a result, despite the five decades since the Supreme Court declared gerrymandering to be capable of being decided by legal principles or by a court, the court has yet to identify manageable standards under which one could measure and thus declare a partisan gerrymander. The failure of the legal system in this political realm has significant ramifications for our democratic system of governance. Using the Supreme Court’s articulated legal reasoning and mandates, we have developed a computational redistricting tool.

METHODS & CODES

Our algorithm, PEAR, or Parallel Evolutionary Algorithm for Redistricting, is implemented in ANSYS C. It can be compiled on Linux, OS X, and Windows as a stand-alone makefile project. PEAR uses MPI nonblocking functions for asynchronous migration for load balancing and efficiency. It uses the C SPRING 2.0 library to provide a unique random number sequence for each MPI process, which is necessary for running a large number of evolutionary algorithms (EA) iterations.

Since the spatial configuration plays a critical role in the effectiveness and numerical efficiency of redistricting algorithms, we designed spatial EA operators that incorporate spatial characteristics to effectively search the solution space. Our parallelization of the algorithm further harnesses massive parallel computing power via the coupling of EA search processes and a highly scalable message-passing model that maximizes the overlapping of computing and communication at runtime.

RESULTS & IMPACT

Our approach is designed to identify redistricting maps that satisfy a set of user-defined criteria with a particular focus on addressing fine levels of spatial granularity. We leveraged and enhanced a scalable Parallel Genetic Algorithm (PGA) library to develop PEAR for the computationally intensive redistricting problem. By incorporating a set of spatial configuration operators and spatial EA operators to handle spatial characteristics and the associated computational challenges, and harnessing massive computing power, PEAR provides a powerful and computationally scalable redistricting tool that has never existed before. In the substantive realm, our computational model allows us to synthesize and organize massive amounts of computation and data to evaluate redistricting schemes and tailor them to notions of “fairness” and democratic rule.

The discrete optimization framework identifies large sets of quality electoral maps. This set of plans is useful at the districting drawing stage as well as for judges who are adjudicating the constitutionality of a redistricting plan. At the drawing stage, this information suitably drives an iterative bargaining process whether that process involves partisan legislators or members of an independent redistricting commission. When this bounded set of plans can be identified, the redistricting process is imbued with valuable structure that is otherwise nonexistent. That structure alone makes the redistricting process more transparent and may serve to reduce legal challenges. When a legal challenge is mounted, this large set of reasonably imperfect plans produces the relevant background and allows one to place and understand the proposed plan in context. The purpose of our optimization strategy, then, is to search, synthesize, and organize massive amounts of information that will supply a common base of knowledge to guide informed and intelligent debate and decisions.

The impact of our project is yet to be determined, but the project has thus far been well received. Our proposal for a new partisan gerrymandering standard, based on our computational tool, won the first place prize in Common Cause’s “Gerrymander Standard” writing competition that was judged by law school deans, law professors, and lawyers. The project has also received some media coverage.

WHY BLUE WATERS

Our PEAR library is designed for extreme-scale redistricting applications. From the beginning, it was intended to scale to all of the processor cores on Blue Waters through nonblocking MPI communication calls. The computational approach we implemented in our solution requires generating a very large number of electoral maps for quantitative study of redistricting phenomena. Identifying quality electoral maps requires significant computing in the combinatorial optimization process. Generating a large number of statistically independent maps is only feasible on a supercomputer at Blue Waters’ scale.

REFERENCES


Figure 1: The algorithm is able to generate billions of high-quality, stable redistricting maps that may be used by the Courts in adjudicating partisan gerrymandering cases. Image credit: Carla Wheeler/USA.
258 Resolving Plant Functional Biodiversity to Quantify Forest Drought Resistance in the Amazon
260 Modeling Nonlinear Physical-Biological Interactions: Eddies and Sargassum in the North Atlantic
262 The Role of Cosmic Rays in Isolated Disk Galaxies
264 Analyzing the Propagation of Soft Error Corruptions in HPC Applications
266 Toward Developing a Thermodynamic Model of Binding-Induced Conformational Transitions in Short, Disordered Protein Regions
268 Genomic Perspectives on the Amphibian Tree of Life
270 Multiscale Simulations of Complex Fluid Rheology
272 High Accuracy Radiative Transfer in Cloudy Atmospheres
274 High-Fidelity Blade-Resolved Wind Farm Simulations
276 Reducing the Computational Cost of Coupled Cluster Theory
278 Simulated Effects of Urban Environments on the Dynamics of a Supercell Thunderstorm
280 Monte Carlo Neutrino Closures in 3D GRMHD Simulations of Core-Collapse Supernovae and Neutron Star Mergers
282 Understanding the Role of Hydrodynamic Fluctuations in Biomacromolecular Dynamics through the Development of Hybrid Atomistic-Continuum Simulation
284 The Impacts of Hydrometeor Centrifuging on Tornado Dynamics
286 Glassy Dynamics and Identity Crises in Hard-Particle Systems
288 Magnetic Reconnection in Laser-Driven Plasmas: from Astrophysics to the Laboratory in Silico
EXECUTIVE SUMMARY
Cases of heat- and drought-induced mortality have been documented in every biome of the world, indicating that changes in global temperatures and precipitation patterns are pushing the world’s forests beyond current thresholds of stress resilience. The increased frequency and severity of droughts and their regional consequences have highlighted the potential vulnerability of the Amazon region to heat- and drought-induced stress. To adequately capture the response of tropical rainforests to water limitation, mechanistic models that incorporate three-dimensional plant morphology and traits are needed. Three-dimensional root water uptake is modeled for a forest plot in the Tapajós National Forest for the 2015–2016 El Niño drought. Results confirm the model’s ability to capture differential response of individual species to standard diurnal cycles and water limitation. Using Blue Waters, future work will model whole-forest response to water limitation and the contribution of root systems to drought resilience.

RESEARCH CHALLENGE
Cases of heat- and drought-induced mortality have been documented in every biome of the world, indicating that changes in global temperatures and precipitation patterns are pushing the world’s forests beyond current thresholds of stress resilience [1]. The Amazon Basin region, home to the world’s largest area of undisturbed tropical biomass, is critical to global energy, water, and carbon cycles. Over the past two decades, the region has been hit with multiple drought events triggered by strong shifts in sea surface temperature caused by the El Niño–Southern Oscillation (ENSO). The increased frequency and severity of droughts and their regional consequences have highlighted the potential vulnerability of the Amazon to heat- and drought-induced stress [2]. To adequately capture the response of tropical rainforests to water limitation, mechanistic models that incorporate diverse plant morphology and hydraulic function are needed.

METHODS & CODES
The computational complexity of single plant models has previously limited incorporation into hydrological models at the forest plot or ecosystem scale, but recent developments in micromodeled root hydraulic architecture has opened the door to coupled models of three-dimensional root water uptake and soil water physics. Root architectures that represent the structural and spatial distribution of roots were modeled using the open source Rootbox model [3]. Each tree system is assigned hydraulic parameterization (e.g., root hydraulic conductivity, water potential thresholds) based on statistically generated water-usage strategies. These strategies may range from risky, which favor carbon assimilation over hydraulic integrity, to conservative, which will limit carbon assimilation and, therefore, water uptake to protect hydraulic pathways from damage. Root water uptake has been coupled with the massively parallel flow and transport model, PFLOTRAN [4], using hybridization techniques from [5].

Using these tools, we are exploring how tree roots contribute to forest drought resilience in areas of the Amazon rainforest during the recent 2015–2016 El Niño drought event. To tease apart the contributions of various ecophysiological properties, ensemble modeling approaches are employed that test a multitude of risk configurations and root distributions. Each of these approaches uses spatial distributions from and is validated with data collected from our field site in the Tapajós National Forest (K67), located in the eastern Amazon River Basin.

RESULTS & IMPACT
Preliminary simulations focused on canopy-dominant species. Representing the largest individuals in the forest, canopy trees have the greatest water demand and are thus susceptible to stress or damage during water limitation. Initial results confirm the model’s capacity to model water uptake for individual trees at the forest scale. Individuals show differential response to water limitation based on both their size and physical traits such as tissue conductivity and stomatal response. Furthermore, intrinsic properties are not the only driver of an individual tree’s response. Social position, the relative location of a tree to its peers, is an important factor in determining where in the soil water column uptake occurs. As soils continue to dry, uptake will be shifted away from the center of mass and into deeper soil layers.

Future simulations will include all canopy layers—wards of 1500 trees per hectare—comprising the largest simulations of three-dimensional root water uptake ever attempted. It is expected that the additional root density and root-zone overlap will further tune individual response as competition increases in soil-water reservoirs. Individual and community drought resilience will be assessed, elucidating the contributions of root systems to forest drought resilience in the Amazon rain forest.

WHY BLUE WATERS
Blue Waters is critical to the ongoing success of this project. Simulations of this complexity and scale require the computational power of this system to make meaningful analyses. Not only are the simulation domains complex, multiple simulations are needed to account for system uncertainty. The enormous biodiversity and harsh environmental conditions of tropical forests hinder data collection needed for model parameterization. Scalable, physically based models provide a necessary tool with which to explore modes of uncertainty and help target data collection efforts.

Elizabeth Agee is a fourth-year Ph.D. student in environmental engineering at the University of Michigan. She is working under the supervision of Valeriy Ivanov, and hopes to graduate in April 2019.
MODELING NONLINEAR PHYSICAL-BIOLOGICAL INTERACTIONS: EDDIES AND SARGASSUM IN THE NORTH ATLANTIC

Maureen T. Brooks, University of Maryland Center for Environmental Science
2016-2017 Graduate Fellow

EXECUTIVE SUMMARY

The macroalga Sargassum, commonly known as “gulf weed,” inhabits the Atlantic Ocean and poses natural resource management challenges for coastal communities in its range. This study seeks to highlight interactions among Sargassum and mesoscale eddies and fronts to better predict Sargassum’s dispersal and growth. To this end, I developed a coupled model system that spans scales from individual organisms up through the basinwide circulation of the Atlantic. The resources of Blue Waters facilitated this model development and allowed for implementation at high resolution. Model results suggest that the Gulf of Mexico (GoM) and Western Tropical Atlantic play a key role in determining Sargassum distribution, and highlight the need for better understanding of the reproductive strategy of this organism. The eddy field in the GoM in particular appears to both influence Sargassum dispersal into the greater Atlantic and alter the growth conditions experienced by Sargassum in the region.

RESEARCH CHALLENGE

Pelagic Sargassum is comprised of two species: Sargassum fluitans and Sargassum natans. These are the only species of macroalgae in the world that spend their entire life cycle floating on the ocean surface. They serve as keystone species in the Sargasso Sea and throughout their range, supporting a thriving ecosystem, from invertebrates to commercial fish, in low-nutrient “ocean desert” regions [1]. Recently, increased reports of Sargassum wash-ups have highlighted its negative impacts. These events are associated with lost fishing and tourism revenue, as well as large clean-up costs for coastal communities on both sides of the Atlantic [2].

Sargassum has many air bladders which provide buoyancy, and this causes colonies to accumulate along convergent eddies and fronts. These features not only affect the regional surface flow, but also can induce vertical velocities that can alter the nutrient availability in the surface waters where Sargassum grows. Mesoscale eddies can persist at monthly time scales, long enough to impact Sargassum biomass based on measured growth rates [3]. This study examines the physical and physiological impacts of mesoscale features on pelagic Sargassum.

METHODS & CODES

This research uses a system of four coupled models to simulate Sargassum growth and interactions with ocean circulation features. A Hybrid Coordinate Ocean Model (HYCOM) [4] domain was implemented at 1/12° (~10m) resolution over a domain that encompasses the known Sargassum distribution, from 15°S to 64°N and 100°W to 15°E. This model has 28 hybrid vertical layers which capture the 3-D ocean circulation and their resolution is concentrated in the upper 200m to better simulate vertical velocities associated with surface eddies. Coupled to the HYCOM model is a biogeochemical model adapted from the work of Fennel [5], which includes nitrogen and phosphorus, as well as phytoplankton, zooplankton, and detritus to effectively capture the dynamics of biologically-mediated nutrient cycling in the upper ocean.

Sargassum colonies are modeled using a combination of Lagrangian particle and individual-based physiology models. I have modified the HYCOM Lagrangian particle code to allow for particle buoyancy to better simulate Sargassum. The particle code has also been improved to allow both forward- and backward-time integration as well as inertial effects, and interpolates the physical and biogeochemical conditions from the first two models along such particle trajectory. The Sargassum physiology model was developed for this study and is run within every individual Lagrangian particle. It uses light, temperature, and nutrient availability to determine growth rate, and because each colony is tracked it can also account for age and reproductive strategy. Vegetative propagation is simulated by initializing a new Sargassum propagule in place when a colony dies to age-related causes in regions where conditions are otherwise favorable. The Sargassum biomass distributions generated by the 4-model suite were validated against monthly satellite climatologies derived from observations over a 10-year period [6].

RESULTS & IMPACT

This multi-scale modeling effort gives an unprecedented view of basin-wide Sargassum biomass and mesoscale biological-physical interactions. Particle seeding and Lagrangian trajectory analyses examined a model domain, with an average area of 1.3×10⁶ km², and found evidence for two potential “seed” regions that exert disproportionate influence on the Sargassum seasonal cycle in the Atlantic. When Sargassum particles are seeded in the Gulf of Mexico and the Western Tropical Atlantic near the mouth of the Amazon River the seasonal distribution of biomass has a 30% reduction in RMS error as measured against the satellite observations. The inclusion of Sargassum’s vegetative reproductive strategy also improved model accuracy, reducing model bias to within 1.5% of the mean observed biomass.

Lagrangian Coherent Structure (LCS) analysis was used to accurately determine the boundaries of mesoscale eddies and fronts via the finite-time Lyapunov exponent field (Fig. 1a). Sargassum particles aggregate along attracting LCS as expected, and biomass is influenced by these structures (Fig. 1b). Because the colonies tend to stay at eddy boundaries, they are spared the low-nutrient conditions in the interior of the large, convergent eddies that pinch off from the Loop Current. Buoyant particles tend not to cross LCS lines, and the differential eddy activity in the western and eastern Gulf of Mexico also helps maintain the Sargassum population in the western GoM where the potential for growth is high and biomass can accumulate. This improved understanding should help us better predict the potential for Sargassum growth based on local oceanographic conditions.

WHY BLUE WATERS

The resources of Blue Waters have made the scale and scope of this project possible. High-resolution ocean circulation modeling alone has a high computational cost. By utilizing Blue Waters, I accomplished the high-resolution ocean circulation modeling as well as coupling it with ocean biogeochemistry, Lagrangian particles, and individual organism physiology at temporal and spatial scales that span orders of magnitude. The professionalism of the NCSA staff has also been key to the success of this project. Their responsiveness and expertise made implementing and running this code on Blue Waters as straightforward as possible.

Maureen T. Brooks is in the fourth year of her Marine-Estuarine-Environmental Sciences Ph.D. program, working under the direction of Victoria Coles at the University of Maryland Center for Environmental Science Horn Point Laboratory. She plans to complete her degree in 2018.
THE ROLE OF COSMIC RAYS IN ISOLATED DISK GALAXIES

Iryna Butsky, University of Washington
2016-2017 Graduate Fellow

EXECUTIVE SUMMARY

Cosmic ray feedback in cosmological simulations is integral to creating robust models of stellar feedback and to reproducing galactic structure that is consistent with observations. Unlike thermal gas, relativistic cosmic rays retain more pressure under adiabatic expansion and cool on significantly longer time scales. Cosmic rays stream along magnetic field lines and drive instabilities that initiate strong outflows. The goal of this project is to use simulations of isolated disk galaxies for which cosmic rays are dynamically important to study how cosmic ray-driven outflows shape the structure and kinematics of the circumgalactic medium. In order to model realistic cosmic ray behavior, it is necessary to employ a fully anisotropic, magnetohydrodynamic (MHD) treatment, which I have implemented in Enzo, a publicly available cosmological simulation code.

RESEARCH CHALLENGE

The circumgalactic medium (CGM) is a dynamically complex and diffuse multiphase gas that extends to the outer edges of galactic halos and contains the majority of a galaxy’s baryons. The interplay in the CGM between outflows from the star-forming disk and inflows from the intergalactic medium provides constraints to theories of galaxy formation and evolution. Recent observations, such as those from the COS-Halos survey on the Hubble Space Telescope (HST), show that the CGM of distant galaxies contains a substantial amount of metal-enriched, cool gas [1]. Cooling times of such a gas are very short compared to galactic timescales, and it is unclear how this material survives in such abundance. The data seem to imply an additional unknown source of nonthermal pressure—for which cosmic rays are a candidate—that supports the cool gas against evaporation. Furthermore, although metals are produced within galactic disks, galaxies retain only 20–25% of these metals in their stars and interstellar medium [2]. Data from the Sloan Digital Sky Survey suggest that metals have been lost to outflows [3], which strongly affect the mass structure and kinematics of the CGM.

Recent studies have shown that including cosmic ray stellar feedback can drive stronger and more mass-loaded galactic winds than thermal stellar feedback alone [4]. Cosmic rays are relativistic charged particles accelerated by shocks (such as supernovae). Unlike thermal gas, cosmic rays retain pressure under adiabatic expansion better and cool on significantly longer time scales. Because cosmic rays are charged particles, they do not move freely through space but propagate along magnetic field lines. This streaming drives instabilities in the magnetic field lines, which contribute to the exponential growth of the galactic magnetic field and drive strong outflows. Because magnetic fields vary greatly in strength and shape between the galactic disk and the CGM, implementing anisotropic cosmic ray transport is crucial to improved accuracy and predictive power within simulations. Therefore, the goal of this project is to employ realistic cosmic ray physics in simulations of isolated disk galaxies in order to study galactic outflows and their effects on the CGM structure and composition.

METHODS & CODES

The challenge of this project was in implementing a new cosmic ray fluid that is compatible with the Riemann solvers in Enzo, a multi-physics astrophysical simulation code. In addition to cosmic ray advection, I implemented several different modes of cosmic ray propagation along magnetic field lines, including anisotropic diffusion, streaming, and gas heating. The new cosmic ray physics will soon be publicly available to all Enzo users.

After extensive testing, I began running a suite of isolated disk galaxies following initial conditions prescribed by the AGORA [5] project. In these galaxy models, supernovae inject thermal, magnetic, and cosmic ray energy into the surrounding gas. I will systematically compare models of varying resolution and cosmic ray transport to the control model with no cosmic rays to isolate the effects of the different propagation modes.

RESULTS & IMPACT

Initial results show evidence of strong, mass-loaded outflows that enrich the CGM out to the virial radius of the halo when cosmic rays are present. Fig. 1 shows an edge-on view of the metal-enriched CGM after only 1.5 Gyr (gigayears). The contours of the magnetic field lines in black show the turbulent nature of the outflowing gas. Follow-up work will focus on comparing the simulated CGM structures to existing COS-Halos observations.

Figure 1: An edge-on view of outflows from an isolated disk galaxy after 1.5 Gyr. The color depicts the metallicity (metal enrichment relative to solar abundances) of the gas, and the structures follow the topology of the magnetic field. The dimensions of the image are 90 kpc x 150 kpc.

I will use the new Trident tool to create synthetic spectra that are tuned to the specifications of the COS-Halos instrument. Directly comparing the simulations to existing data is a state-of-the-art approach that will place better constraints on theories pertaining to the CGM and make predictions for the structure and metallicity distribution of the CGM for future COS-Halos observations.

WHY BLUE WATERS

The universal challenge faced by galaxy simulations is reconciling the great dynamic range in physical and temporal scales. Although a galaxy’s CGM scales hundreds of kiloparsecs (1 kpc = 3.09e16 km), its evolution is dictated by stellar feedback, which happens on subparsec scales. Even with adaptive mesh refinement, this is a formidable task, and galaxy simulations have yet to resolve the scale of individual stars. Furthermore, each cell in the simulation contains many different fluid properties and follows a complicated set of interaction rules in order to capture the equally complicated physics governing galaxy evolution. Therefore, galaxy simulations require the use of massively parallel, high-performance supercomputers such as Blue Waters. In addition to Blue Waters’ computational resources, I have benefited greatly from the support of its staff, who are admirably dedicated to resolving issues in a timely manner.

Iryna Butsky is in the second year of a Ph.D. program in astronomy at the University of Washington (UW). Her advisor is Tom Quinn, also of UW; she plans to graduate in June 2021.

262

263
ANALYZING THE PROPAGATION OF SOFT ERROR CORRUPTION IN HPC APPLICATIONS

Jon Calhoun, University of Illinois at Urbana-Champaign

2016-2017 Graduate Fellow

EXECUTIVE SUMMARY

Because the rate of radiation-induced soft errors impacting application data are expected to increase on future high-performance computing (HPC) systems, analyzing how corruption due to soft errors propagates inside applications becomes critical for developing efficient detection and recovery schemes. Quantifying the latency (the number of instructions or iterations) of common symptoms of soft errors—e.g., crash or detection—allows for measuring the effectiveness of soft error detectors at containing corruption. Containing corruption allows for low-cost localized recovery instead of the high-cost global roll-back recovery checkpoint-restart. To analyze corruption propagation in HPC applications, an LLVM-based compiler tool instruments an application, allowing for tracking of corruption at an instruction and application variable level. Results show that the latency of common symptoms of soft errors along with speed of propagation to other processes varies dramatically. Blue Waters provided us with a platform to run thousands of fault-injection experiments and analyze hundreds of gigabytes of data efficiently.

RESEARCH CHALLENGE

Design constraints, such as cost of procurement and power budget, may make future HPC systems more susceptible to transient radiation-induced soft errors [1]. Soft errors commonly result in bit inversion and are the reason that all HPC main memories use protection mechanisms such as error-correcting codes (ECC). On future systems, full protection from the effects of soft errors will be prohibitively expensive. Thus, HPC applications running for long durations and at large scales may experience data corruption during execution [2]. Furthermore, global rollback-recovery becomes increasingly expensive at larger scales. Analyzing the effectiveness of current software-based detection schemes at containing data corruption is central in facilitating low-cost localized recovery and ensuring correct simulation results.

METHODS & CODES

To track corruption propagation inside HPC applications, we constructed an LLVM compiler pass to create an executable file that emulates lockstep execution between a nonfaulty version of the application called Gold and a version that experiences a single bit-flip fault injection called Faulty. Gold produces a reference set of loads, stores, and correct program behavior that Faulty is judged against. The compiler pass completes the following steps when instrumenting the code: (1) it duplicates compiled code, forming two applications—Gold and Faulty, (2) it interleaves instructions from Gold and Faulty, emulating lockstep execution, (3) it instruments all loads and stores in Faulty with a function call to log deviations, (4) it instruments Faulty branches to check for control flow divergence, (5) it instruments faulty code for fault injection with FlipIt [3], and (6) it logs deviation in application-level variables at the end of iterations. This tool is used to investigate the latency (in number of LLVM instructions executed) of common symptoms of soft errors, the speed and extent of propagation between variables and parallel processes, and the impact of local problem size and compiler optimizations on propagation. The applications we test include: Jacobs, CoMD, and HPCCG. The latter two are taken from the Mantovo mini-app collection [4]. Each of these applications is modified with a lightweight soft-error detection scheme to investigate the latency of soft-error detection.

RESULTS & IMPACT

This project tracks corruption propagation at both the instruction level and the application variable level. Investigation at the instruction level shows the latency (in number of LLVM instructions) of common symptoms of soft errors: crash, control flow divergence, and soft error detection. Application crashes have the shortest latency of the symptoms considered. Results show that 90% of crashes occur within four instructions. During those four instructions, there is little chance of corruption propagation. The latency of control flow divergence and soft error detection are noticeably larger. The larger latency allows for corruption to propagate to other variables and processes. To measure the extent and magnitude of corruption propagation in variables and processes, we track propagation at the application variable level. Tracking propagation in application variables highlights dependencies between variables and processes. The application HPCCG uses the conjugate gradient linear solver to solve the sparse system of \( Ax = b \). HPCCG is run 1,000 times with a single bit-flip fault injected on process rank 3 during every run. Fig. 1 shows the average percentage of elements deviated by more than 1e-10 (color) in the solution vector, \( x \), for the application HPCCG across all MPI processes (y-axis) for subsequent iterations after injection (x-axis). As corruption in the solution vector propagates locally, the horizontal color for that row grows darker. As corruption is removed, the color lightens. Propagation between processes can be seen by looking at the color progression of columns at each iteration. Due to an MP1 ALU reduce in the algorithm, corruption is present on all processes within one iteration after injection. However, the magnitude of the corruption is below our threshold deviation tolerance of 1e-10 and does not appear in Fig. 1. As HPCCG continues to iterate, corruption propagates in all processes. Over time, corruption is removed as HPCCG refines the solution to \( Ax = b \). The largest in magnitude and most difficult to remove corruption remains on the process that experienced the fault, rank 3. Analyzing HPC applications and how they propagate corruption is useful in determining what types of soft-error detectors to use and where they are best placed. Furthermore, knowing which variables are corrupted and on which processes corruption resides when a soft error is detected enables low-cost localized recovery instead of an expensive global rollback recovery, checkpoint-restart. Going forward, as HPC design constraints increase the likelihood of soft errors that can impact HPC application data, efficient detection and recovery schemes are needed to ensure applications obtain correct results.

WHY BLUE WATERS

The Blue Waters system allows us to perform fault injection and track propagation in thousands of HPC application runs and analyze the data generated more efficiently than other available systems. The fast turnaround time in obtaining and analyzing results greatly increased the speed and quality of this project.

PUBLICATIONS AND DATA SETS


Jon Calhoun was a fifth-year Ph.D. student in computer science working under the direction of Luke N. Olson and Marc Snir at the University of Illinois at Urbana-Champaign when this work was completed. He graduated in August 2017.
TOWARD DEVELOPING A THERMODYNAMIC MODEL OF BINDING-INDUCED CONFORMATIONAL TRANSITIONS IN SHORT, DISORDERED PROTEIN REGIONS

Justin Drake, University of Texas Medical Branch
2016–2017 Graduate Fellow

EXECUTIVE SUMMARY

Intrinsically disordered regions (IDRs) in proteins are highly dynamic and often undergo conformational transitions upon binding a protein target. The thermodynamics underlying these transitions, particularly conformational entropy and its contribution to protein binding, remain elusive. In this work, we use molecular dynamics simulations to sample the conformational states of a model disordered polypeptide as a function of chain length. Backbone dihedral angles for each polypeptide were measured from the simulation trajectories and used to calculate the conformational entropy using the Quasi-Harmonic Analysis (QHA), Boltzmann Quasi-Harmonic (BQH), and two variants of the Mutual Information Expansion (MIE) method. Our results suggest that short IDRs may provide a significant source of free energy that proteins may tap through order-disorder transitions to modulate or regulate protein binding. We are currently working to develop a model that more fully describes the energetics of these conformational transitions and the extent to which they contribute to protein binding processes.

RESEARCH CHALLENGE

For much of the twentieth century, it was widely held that the three-dimensional structure of a protein dictates its cellular function. While true for a large group of proteins, at the turn of the twenty-first century, seminal papers (1, 2) opened the door to the notion that the conformational entropy (i.e., a proxy for disorder) of IDRs associated with order-disorder transitions is believed to play an important role in protein-binding processes (3). However, it is notoriously difficult to measure or approximate conformational entropy using traditional solution biophysics techniques. In this work, we use all-atom molecular dynamics simulations to sample the conformational states of successively longer oligoglycine polypeptides (Glyn, where n = 3, 4, 5, 10, and 15 residues), and use a number of methods to calculate the dihedral angle contribution to the conformational entropy. Oligoglycine is an ideal protein backbone model and is found in varying lengths in a number of examples.

METHODS & CODES

Conformational entropy (i.e., a proxy for disorder) of IDRs associated with order-disorder transitions is believed to play an important role in protein-binding processes (3). However, it is notoriously difficult to measure or approximate conformational entropy using traditional solution biophysics techniques. In this work, we use all-atom molecular dynamics simulations to sample the conformational states of successively longer oligoglycine polypeptides (Glyn, where n = 3, 4, 5, 10, and 15 residues), and use a number of methods to calculate the dihedral angle contribution to the conformational entropy. Oligoglycine is an ideal protein backbone model and is found in varying lengths in a number of IDRs. Simulations were performed with the NAMD (5) molecular dynamics (MD) package and Amber (6) force field. Backbone dihedral angles for each model were measured from the simulation trajectories and used to calculate the conformational entropy using QHA (7), BQH (8), and two variants of the MIE (9) method.

RESULTS & IMPACT

The structural and thermodynamic properties of IDRs depend on chain length. In this work, we have calculated the conformational entropy (5) of an oligoglycine model as a function of chain length using four methods that account in different ways for the reduction in entropy due to correlated dihedral motions. Figure 1 shows that conformational entropy scales remarkably linearly with chain length, with slopes ranging between 4.1 and 5.6 cal/mol/K per residue. This is consistent with the experimentally measured loss of backbone entropy (4.5 cal/mol/K per residue) upon protein folding (10). S^2^2 is a strict upper bound on the conformational entropy. S^2^2 and S^2^ do not differ in the former accounts only for dihedral angle correlations within each residue, whereas the latter considers all possible pairs of dihedral angles along the backbone, yielding a slight reduction in the slope of S^2^ (4.1 cal/mol/K/ residue) compared to S^2^ (4.7 cal/mol/K/ residue). We note that S^2^ is the simplest and most computationally efficient method of the four. Coupling of the dihedral angles within each residue appears to be the most significant source of correlations affecting the conformational entropy (Fig. 2). Preliminary data also suggest that oligoglycine conformational entropy is relatively insensitive to constraints on end-to-end distance, radius of gyration, and solvent-accessible surface area.

A at temperature of 300K, the conformational entropy of these short IDR models (e.g., 21 kcal/mol for Gly15) provides a significant free energy reservoir that proteins may tap through order-disorder transitions to modulate or regulate protein binding. We will incorporate our current results with those reported in last year’s Blue Waters Annual Report on the solvation thermodynamics of oligoglycines to create a more complete model of the thermodynamic contribution to protein binding energetics associated with these conformational transitions. We will then evaluate this current model in the context of biologically relevant systems.

WHY BLUE WATERS

Statistically converged estimates of conformational entropy require extensive sampling of the conformational states of a protein or polypeptide during molecular dynamics (MD) simulations. Access to GPU-optimized MD packages such as NAMD for use with Blue Waters XSE nodes, assistance from the project staff to develop an efficient workflow, and the high-throughput of Blue Waters was critical to performing our research in a tractable amount of time. Additionally, the Blue Waters Graduate Fellowship provided me the opportunity to successfully propose, manage, and partially fund my doctoral research, and expanded my understanding of high-performance computing. These experiences are instrumental in my pursuit of becoming an independent researcher.

PUBLICATIONS AND DATA SETS


GENOMIC PERSPECTIVES ON THE AMPHIBIAN TREE OF LIFE

Paul Hime, University of Kentucky
2016–2017 Graduate Fellow

EXECUTIVE SUMMARY

The relationships among extant amphibians (frogs, salamanders, and caecilians) has been a longstanding debate in phylogenetics. Previous studies have supported any of the three possible topologies relating these three extant orders of amphibians. This project developed a novel amphibian-specific gene-capture system to target and sequence 220 nuclear genes in a diverse set of 296 amphibian species representing all major amphibian lineages. Using an information-theoretic approach to compare the relative model support for interordinal amphibian relationships and to calculate the magnitude of support for these relationships across genes, this project has demonstrated substantial variation in how strongly different genes support any interordinal topology. A few genes supporting each of the three interordinal models resoundingly reject the alternative models (high ΔAIC values in Fig. 1), while most other genes only weakly support any model for relationships among the three amphibian models, suggesting an erosion of phylogenetic signal over these deep evolutionary timescales. Phylogenetic information content in many genes in the genome appears to have eroded in the nearly 300 million years since the three amphibian orders diverged, yet our results support a model of frogs and salamanders sharing a common ancestor despite substantial variation in phylogenetic signal across different genes. Our results also suggest a revised hypothesis for the relationships among the extant families of amphibians and support a more recent origin for many of the hyperdiverse lineages of frogs. Overall, this project highlights the power of phylogenomics and a model-based theoretical framework for testing phylogenetic hypotheses in the era of genome-scale evolutionary biology.

RESEARCH CHALLENGE

All organisms trace their ancestry back to a single common ancestor nearly four billion years ago [1]. Yet, today, life has diversified into tens of millions of species. Reconstructing these evolutionary relationships is a key aim of the field of phylogenetics, and such insights may inform nearly all aspects of modern biology. The recent advent of genome sequencing technologies has ushered in a new era of phylogenomics in which genetic data generation is no longer a rate-limiting step toward testing phylogenetic hypotheses. Yet, it is becoming increasingly clear that different regions of the genome can support conflicting phylogenetic hypotheses, and reconciling these discordant gene genealogies is a key problem facing evolutionary biologists today. Information-theoretic approaches have great promise for parsing signal from noise in large phylogenomic data sets.

METHODS & CODES

Three possible topologies exist for relationships among frogs, caecilians, and salamanders. The Procera hypothesis supports caecilians+salamanders, the Acauda hypothesis supports frogs+caecilians, and the (canonical) Batrachia hypothesis supports frogs+salamanders. To evaluate support for each of these three hypotheses across the genome, we conducted gene-by-gene tests of constrained topology, comparing the maximum likelihood estimates of gene trees for 194 nuclear genes between constraints for the three possible interordinal topologies. Gene trees were estimated in RAxML [2]. The Akaike information criterion (AIC) [3] was used to quantify the direction and magnitude of support across genes. Gene trees were then reconciled into an estimate of the species tree using Astral [4]. This species tree topology was then used in concert with a set of 25 fossil calibrations to estimate divergence times across Amphibia in MCMCTree [5].

RESULTS & IMPACT

This project has demonstrated that different genes in the amphibian genome support conflicting topologies for the relationships among the three amphibian orders. Roughly half of the genes examined support frogs and salamanders as each other’s closest relatives (sister taxa). Roughly one-quarter of the genes support frogs and caecilians as sister taxa, while another quarter support caecilians and salamanders as sister taxa. Significant numbers of genes are found to support each of these three competing topologies, suggesting that either there has been an erosion of phylogenetic signal over deep time and/or that population-level processes (large effective population sizes and rapid diversification) have led to incomplete lineage sorting and gene tree, or species tree discordance. Results from our AIC-based topology testing support the Batrachia hypothesis (frogs and salamanders are each other’s closest relatives) at the level of the species tree and suggest that large numbers of genes may need to be analyzed in order to overcome stochastic phylogenetic noise at deep timescales. Our divergence time analyses also provide a revised timescale for amphibian diversification through time and indicate that a rapid increase in net rates of species diversification occurred at the Cretaceous–Tertiary boundary. Analyzing hundreds of genes for hundreds of species has traditionally been computationally intractable for empirical data sets, and this study is one of the first to use an information-theoretic framework to address not only the direction of support for phylogenetic hypotheses across the genome but also the magnitude of that support.

Figure 1: Support for competing models of interordinal amphibian relationships—194 genes are binned on the horizontal axis by support for frogs+caecilians (Acauda hypothesis in red). The upper panel depicts the hypothesis in green), frogs+salamanders (Batrachia hypothesis in blue), or caecilians+salamanders (Procera hypothesis in red). The lower panel depicts the magnitude of support against competing hypotheses measured by the AIC.

Figure 2: Estimated divergence times across Amphibia—Annonia, caecilians, salamanders, and frogs are depicted by magenta, red, blue, and green branches, respectively. The dashed red line denotes the Cretaceous–Tertiary boundary, a period that witnessed several independent increases in the net rate of diversification in multiple amphibian lineages.

WHY BLUE WATERS

The Blue Waters fellowship has provided essential support to enable me to make progress on this project and to successfully write and defend my Ph.D. dissertation. Although many of the analyses I intended to execute on Blue Waters have been delayed due to unforeseen issues with our data set, I am hopeful that with a time extension to this allocation, I will be able to leverage the massively parallel resources on Blue Waters to perform topological testing in Neobatrachian frogs at a level of gene- and species-sampling that has not been attempted in any other empirical data set.

Paul Hime—a sixth-year Ph.D. student in evolutionary biology at the University of Kentucky—successfully defended his dissertation in April 2017. His work was directed by David W. Weisrock.
MULTISCALE SIMULATIONS OF COMPLEX FLUID RHEOLOGY

Michael P. Howard, Princeton University
2016-2017 Graduate Fellow

EXECUTIVE SUMMARY

We developed the first massively parallel, open-source implementation of the multiparticle collision dynamics (MPCD) algorithm for graphics processing units (GPUs). MPCD is a mesoscale particle-based simulation method for hydrodynamics, and is particularly useful for modeling complex fluids and soft matter. Our implementation of the MPCD algorithm scales up to 1024 nodes on Blue Waters, and GPU acceleration gives a 3x speedup for an XK node compared to an XE node. The developed software will enable studies of complex fluids at length and time scales that would be otherwise inaccessible.

RESEARCH CHALLENGE

Complex fluids, readily encountered in biology, consumer products, and industrial processing, are multicomponent mixtures that exhibit a rich variety of flow behaviors. A classic example is the cornstarch-water “oobleck” mixture, which acts like a liquid when pressed slowly but can thicken to support the weight of a person when struck quickly. Such peculiar macroscopic flow properties of complex fluids are fundamentally controlled by microscopic molecular structures and interactions. Computer simulations are ideal tools for studying this nontrivial and difficult-to-predict relationship; however, performing simulations of complex fluids at physically relevant scales presents a considerable challenge.

Many complex fluids consist of a mixture of solute macromolecules (polymers) and hard spherical particles (colloids) suspended in a liquid solvent such as water. To obtain the correct dynamics, it is essential to resolve both the direct solute-solute interactions as well as the solvent-mediated interactions. The latter dominate the computational cost for a molecular model of such a mixture; however, a molecular-level description of the solvent itself is often not of interest. A multiscale approach that simplifies the solvent model while preserving its most important interactions is required to study complex fluids at relevant length and time scales.

METHODS & CODES

Multiparticle collision dynamics (MPCD) [1] is a mesoscale simulation method that combines classical molecular dynamics (MD) simulations of microscopic solutes with a coarse-grained solvent representation that faithfully reproduces long-range hydrodynamics and thermal fluctuations. In MPCD, the solvent is modeled by a set of point particles that alternate between ballistic streaming steps and cell-based, momentum-conserving, stochastic multiparticle collisions. The frequency and nature of the collisions control the transport coefficients of the solvent. Solute molecules propagate using the MD equations of motion, and are coupled to the MPCD solvent during either the streaming or collision step [2, 3]. Significantly larger length and time scales can be accessed in MPCD than with an explicit-solvent MD model because MPCD solvent particles have simple interactions with each other. However, in practice, tens of millions of MPCD particles may be required for a simulation, necessitating a parallel computational approach. We developed the first massively parallel, open-source implementation of MPCD for graphics processing units (GPUs).

We implemented the MPCD algorithm as part of the open-source simulation package HOOMD-blue [4, 5], which has optimized MD methods designed for NVIDIA GPU architectures. All MPCD data are stored independently from HOOMD-blue’s MD data to ensure high performance. Nearly all computations are performed exclusively on the GPU, which minimizes latency from host-device data transfers. We employ a spatial domain decomposition strategy to extend support to multiple GPUs [5]. One MPI process is assigned per GPU, and we perform runtime autotuning to ensure optimal CUDA kernel launch parameters. Flexible initialization and simulation setup is supported through a scriptable Python user interface.

RESULTS & IMPACT

We performed strong-scaling benchmarks of our MPCD implementation on Blue Waters. We benchmarked two cubic simulation boxes with edge lengths $L = 200a$ and $L = 400a$, where $a$ is the size of an MPCD cell. There were 10 MPCD solvent particles per cell, giving 80 million and 640 million total particles in each simulation box, respectively. We benchmarked a CPU-only implementation of our code using 16 processes per XE node and our GPU implementation using 1 process per XK node. The CPU-only code showed excellent strong scaling up to 1024 nodes. The GPU code also showed good scaling to 1024 nodes, with some performance lost at the highest node counts owing to communication latency for the MPCD cells. GPU acceleration on the XK nodes gave a roughly 3x speedup compared to the XE nodes. The complex fluid and soft matter research communities will significantly benefit from the developed MPCD software, especially the GPU implementation, which will permit studying processes at scales that would otherwise be inaccessible.

WHY BLUE WATERS

Blue Waters is the only system available to us that delivers both the CPU and GPU resources necessary to develop and optimize our software at scale. The large number of GPUs available in the XK nodes also significantly increases our overall scientific productivity and allows us to study process dynamics that would be challenging or impossible to obtain with fewer resources.

Figure 1: Strong-scaling benchmarks of MPCD software on Blue Waters for CPU-only (XE nodes) and GPU (XK nodes) implementations.

Michael P. Howard is currently completing the fourth year of a Ph.D. program in chemical engineering at Princeton University, under the direction of Athanassios Z. Panagiotopoulos. He expects to graduate in 2018.
EXECUTIVE SUMMARY

One of the most important roles clouds play in the atmosphere is in redistributing radiative energy from the sun and that which is emitted from the Earth and atmosphere. Given the ubiquity of cloud coverage, it is imperative that we get the interactions between clouds and radiation correct if we want to accurately predict and observe weather and climate. However, radiative transfer in the atmospheric sciences is generally modeled crudely because of the perceived computational expense. Evidence of a transfer in the atmospheric sciences is generally modeled crudely due to these crude assumptions has been seen in satellite-observed properties as well as modeled cloud properties.

METHODS & CODES

Unlike the direct approach to solving the radiative transfer equation, the Monte Carlo approach has the potential to be embarrassingly parallel, since the random samples are independent from one another. The figure shows weak and strong scaling tests for both solar and internal sources of photons for the monochromatic model, “IMC-emission,” in terms of throughput, or number of photons simulated per minute, and throughput efficiency. For both weak and strong scaling experiments, the solar source of photons exhibits a faster drop-off in efficiency but has an overall higher throughput than the internal source of photons. The “IMC-emission” model shows better weak scaling efficiency than strong scaling efficiency for both sources of photons over the range of processes tested.

RESULTS & IMPACT

The overarching goal of this project is to make publicly available the radiative transfer community the models, tools, data, and products developed to aid in faster and more robust progress in addressing scientific questions about the interactions of clouds and realistic radiative transfer. An existing monochromatic 3D Monte Carlo community solar radiative transfer model was further developed to include terrestrial emission in addition to solar sources of radiation. That model was then further developed to include integration over the electromagnetic spectrum to produce the broadband 3D model discussed above. In addition to the development of these two community models, several other products have resulted so far and will be made available to the community. These include databases of high spectral resolution radiative properties of earth’s gaseous atmosphere and liquid water clouds, the largest and highest resolution publicly available databases of their kind. The tools and workflow to create and subset them will also be made available. These data can be mined to update the decades-old broadband parameterizations of cloud radiative properties that are still in wide use today, for example. Each product has been thoroughly vetted for accuracy. The results of these tests will be made available for reproduction by other scientists to test these models or their own. Finally, the first few idealized experiments with long heritage in the literature have been conducted to provide the first set of benchmark simulation results that can be used to evaluate other models.

WHY BLUE WATERS

Access to debugging and profiling tools such as CrayPat and DDT allowed me to streamline the development process. Having access to a point of contact on the SEAS staff helped me think through issues and find tailored solutions for my problems that would have otherwise delayed progress for weeks. The responsiveness of the Blue Waters staff through the JIRA ticket system allowed for limited interruption in progress when small issues or questions arose. My experience as a Blue Waters graduate fellow has been invaluable to my professional development. I hope to make use of Blue Waters for the rest of its lifetime.

PUBLICATIONS AND DATA SETS

RESULTS & IMPACT
This research enabled the largest complete wind farm simulations to date using a high-fidelity blade-resolved turbine model. This demonstrates the capability for large-scale simulations of entire wind farms and serves as a milestone for high-fidelity methods in a multi-scale problem spanning 10 orders of spatial magnitude.

WHY BLUE WATERS
Blue Waters offers a unique environment not only as a computational resource but also for its expert project staff. The design of Blue Waters makes it an excellent machine geared toward scientific output rather than just its flop rate. Blue Waters allowed us to perform large-scale wind farm simulations using tens of thousands of compute cores. In addition, the project staff provided excellent insight for optimization and maximizing throughput.

EXECUTIVE SUMMARY
Wind energy is becoming an emergent renewable energy source throughout the world. Costs have dropped dramatically over the past two decades, making wind energy a desirable alternative to fossil fuels. Improvement in wind energy application simulation technologies may have a profound economic impact through improved wind plant efficiency. The goal of this work is to develop state-of-the-art aerodynamics modeling techniques using high-fidelity blade-resolved turbine models to simulate complete wind farms. The numerical methods in this research utilize multiple mesh and multiple computational fluid dynamics flow solvers coupled in an overset framework.

RESEARCH CHALLENGE
High-fidelity numerical simulation of wind energy applications is becoming a precedent for future technologies, not only for the wind energy sector but also for lower-fidelity modeling as well. The need for high-fidelity simulation using a complete geometric description of wind turbines with tower and nacelle is essential for capturing the true aerodynamic nature of the flow, which is highly turbulent and chaotic. This is especially important in the study of downstream wake effects on wind turbines that cause the primary decrease in wind plant power production efficiency. Wind farm simulation is a truly multiscale problem; the smallest turbulent structures are on the order of microns and the largest scales present from atmospheric turbulent inflow are on the order of tens of kilometers. To simulate ten orders of spatial magnitude, tens to hundreds of thousands of computing cores are required to approach this grand challenge problem.

METHODS & CODES
The numerical methods developed in this research utilize an overset grid paradigm where multiple meshes and multiple flow solvers are used in a coupled manner to enable efficient simulation of this truly multiscale problem. Nearly all software used on our framework is developed in-house at the University of Wyoming with the exception of the p4est adaptive mesh refinement framework developed by Carsten Burstedde, et al. [1]. The two flow solvers developed at the University of Wyoming are NSU3D, an unstructured 3D finite-volume solver, and dg4est, a high-order discontinuous Galerkin finite-element solver. The overset solver used is TIOGA, developed by Jay Sitaraman of Parallel Geometric Algorithms, LLC. The complete framework is known as the Wyoming Wind and Aerodynamics Applications Komputation Environment (WwAAKE3D).

Andrew Kirby is a fifth-year Ph.D. student in mechanical engineering at the University of Wyoming. He is working under Dimitri Mavriplis and expects to graduate in December 2017.
RESEARCH CHALLENGE

Retinylidene proteins are photosensitive proteins that serve a variety of biological functions and have been found in all domains of life. For example, the rhodopsin protein and the visual pigments found in human eyes enable vision. Other retinal models require sampling different conformations of the retinal chromophore. Through the use of Blue Waters, we aim to simulate the absorption spectra of different retinal models. We successfully applied our new approach to several retinal models. Our initial findings indicate that THC-EOM-CC2 performs better than other available excited state methods that are capable of treating this system size.

METHODS & CODES

In an effort to reduce the computational cost of EOM-CC2 and other quantum chemistry methods, we developed the tensor hypercontraction (THC) approximation. We applied the THC approximation to EOM-CC2 and showed that this lowered the cost of the calculation from an order of O(N^5) to O(N^4). To further improve the efficiency of THC-EOM-CC2, we redesigned the algorithms to take advantage of parallelism. This included developing a code that was parallelized at a higher level to take advantage of multiple compute nodes and at a finer level to take advantage of acceleration with GPUs. We find that for a fixed problem size, the parallel implementation of THC-EOM-CC2 scales superlinearly in the regime where we parallelize only Laplace quadrature points across compute nodes. In the regime where we distribute both quadrature points and grid point blocks across compute nodes, we see a computational speedup that levels off as the number of compute nodes increases.

RESULTS & IMPACT

Our approach, combined with the parallelization approach developed in this work, offers a new way to parallelize and improve the efficiency of quantum chemistry methods. In addition to its use in CC2 and EOM-CC2, the THC approximation has been shown to reduce the computational complexity of other electronic structure methods. While this work was focused on parallelizing THC-EOM-CC2, the general parallelization design outlined can be applied to other quantum chemistry methods using the THC approximation. We expect this new approach will be used to improve the efficiency of many other electronic structure methods in different quantum chemistry software.

The development of the MPI-enabled and GPU-accelerated THC-EOM-CC2 code and the use of Blue Waters allow us to study the absorption spectra of several model retinal chromophores. We find that for a fixed problem size, the parallel implementation of THC-EOM-CC2 scales superlinearly in the regime where we parallelize only Laplace quadrature points across compute nodes. In the regime where we distribute both quadrature points and grid point blocks across compute nodes, we see a computational speedup that levels off as the number of compute nodes increases.

We are currently comparing the results of THC-EOM-CC2 to other excited state methods. Our initial findings indicate that the THC-EOM-CC2 method more accurately describes the excited states of the retinal models compared to the results from time-dependent density functional theory. The THC-EOM-CC2 method will be a useful tool for studying the excited states of chemical systems that are otherwise challenging for the available excited state methods.

WHY BLUE WATERS

Access to Blue Waters allowed us to develop the MPI-enabled and GPU-accelerated THC-EOM-CC2 method. Blue Waters’ code development environment and tools helped us test and debug different implementations more rapidly. This work represents the first time we have tried to combine MPI and GPU acceleration for a method within TeraChem. The Blue Waters project staff offered valuable insight into code development for high-performance computing systems. Additionally, studying the absorption spectra of the retinal models requires sampling different conformations of each model retinal. Blue Waters enabled us to calculate the ground state and multiple excited states of many different configurations.

PUBLICATIONS AND DATA SETS


We currently are comparing the results of THC-EOM-CC2 to other excited state methods. Our initial findings indicate that the THC-EOM-CC2 method more accurately describes the excited states of the retinal models compared to the results from time-dependent density functional theory. The THC-EOM-CC2 method will be a useful tool for studying the excited states of chemical systems that are otherwise challenging for the available excited state methods.

Sara Kokkila Schumacher received a Ph.D. in chemistry from Stanford University in January 2017. She currently is a postdoctoral researcher in high-performance computing at IBM.
SIMULATED EFFECTS OF URBAN ENVIRONMENTS ON THE DYNAMICS OF A SUPERCELL THUNDERSTORM

Larissa Reames, The University of Oklahoma
2016–2017 Graduate Fellow

EXECUTIVE SUMMARY
This study used real-data simulations to quantify the impacts of a large Great Plains (U.S.) urban area on the evolution and strength of a supercell thunderstorm. Simulations with urban areas are compared to those without, with the aid of hierarchical clustering analysis, to form statistically similar groups of simulations. We investigated the effects of the storm with various city-relative paths as well as the storm lifecycle stage during urban interactions. These comparisons concentrate on differences in boundary layer characteristics prior to storm formation as well as changes in supercell structure, dynamics, and evolution. Additionally, we performed a factor separation analysis to determine which aspects of the urban area (i.e., roughness or thermal differences) have the most effect. The results suggest that the urban area, particularly surface roughness, can have a significant impact on storm strength, and that these modifications change based on the city-relative path of the storm.

RESEARCH CHALLENGE
Earth’s population is increasingly concentrated in urban areas, with nearly two-thirds of the world’s population expected to live in urban areas by 2050. As the number of people within cities grows, it is becoming more important to understand and to correctly predict the interactions between urban environments and the atmosphere. As such, many studies have investigated the effect of urban areas on weakly forced precipitation systems. However, interactions between urban areas and synoptically-active convection, such as supercells, remain relatively unexamined.

METHODS AND CODES
Using the Weather Research and Forecasting (WRF) [1] model—a community mesoscale numerical weather prediction model—we investigated a total of 134 simulations of a supercell thunderstorm to quantify the impacts of a large Great Plains urban area on the evolution and strength of a supercell thunderstorm. In order to properly resolve complex urban structure, all simulations were run on a 500-m horizontal grid over a 250-km x 250-km grid. In addition, to resolve the atmospheric boundary layer well, 120 vertical grid points were used, with 20 in the lowest 1.5 km above ground. In all, over 150 million points were integrated over 75,000 time steps for each simulation. Ten of the simulations contained homogeneous land use (CTRL) to serve as a comparison point for simulations with urban areas. An urban area simulated to have both increased surface roughness and thermal properties characteristic of manmade surfaces (i.e., “full physics”) was placed in 108 grid-docked locations (WestF and EastF) throughout the domain to determine effects of the city-relative path of the storm. At eight of these city locations (four near the middle (East) and four near the beginning (West) of the storm’s track), two additional simulations were performed with only either thermodynamic (WestT and EastT) or roughness (WestR or EastR) properties of the urban area simulated, providing an additional 16 simulations.

RESULTS AND IMPACT
We compared full-physics urban simulations to CTRL with the aid of hierarchical clustering analysis (HCA) to form statistically similar groups of simulations. In this analysis, we investigated the effects of the storm having various city-relative paths, as well as the storm lifecycle stage during urban interactions. These comparisons concentrate on differences in boundary-layer characteristics prior to storm formation as well as changes in supercell structure, dynamics, and evolution. The results (Fig. 1), suggest that when the storm passes to the north or directly over the city center late in its life cycle, low-and mid-level storm rotation strength increases, and the center of rotation tracks farther south. Although differences in mid-level rotation are minimal when the storm tracks south of the city, low-level rotation increases, especially late in the storm’s lifetime.

Using the 16 individual physical models (i.e., surface roughness or thermodynamic urban properties only), we undertook a factor separation approach to determine the relative importance of the roughness and thermal characteristics of urban areas on storm modification. Results (Fig. 2) generally suggest that surface roughness and its interactions among thermodynamic properties are the dominant contributors to urban-induced effects on storm strength and evolution. Additionally, the amplitude of interactions between shear and thermodynamic modifications is often similar in magnitude to either effect individually.

Most investigations of near-surface urban properties and urban-storm interactions have focused on the impact of the thermal effects of cities, but the results of this investigation argue for a greater focus on the implications of urban surface roughness. Unlike the full-physics simulations, those simulations with the city parameterized only as a roughness element, and with a long urban-storm interaction, resulted in the largest differences from nonurban simulations of any of the factor-separation analyses.

WHY BLUE WATERS
While HCA has been used previously for attribution of variations in synoptic and mesoscale fields to various factors, this is the first time it has been used to analyze storm-scale modifications. Given their large scale of motion, synoptic [O (10 km)] and mesoscale [O (100 km)] phenomena are generally more predictable than severe storms [O (10 km)], thus few simulations are required to attribute large-scale field variations to modifications in boundary conditions and parameterization options. However, to perform attribution of small-scale effects to various factors in a real-data (i.e., nonidealized) simulation, many simulations are required to ensure that the simulated changes are significant. The general hinderance to such an analysis is the large computational requirement; hence, the resources made available on Blue Waters were vital to this work. While each simulation was relatively small, the large quantity of simulations needed to produce significant results required the large computational and data storage capacities of Blue Waters.

Larissa Reames graduated from the University of Oklahoma in May 2017 with a Ph.D. in meteorology. There, her work was directed by David Stensrud at Pennsylvania State University. She currently is a postdoctoral research associate at the National Oceanic and Atmospheric Administration’s National Severe Storms Laboratory.

Figure 1: Cumulative CTRL group average (black) and accumulated difference from CTRL of N (yellow), S (red), W (blue), and E (green) groups of (a) low-level (0 to 1-km) storm strength, (b) low-level (0 to 1-km) storm strength, (c) and (f) low-level (0 to 1-km) storm strength, (d) and (g) maximum updraft speed, and (e) and (h) near-surface circulation strength.
EXECUTIVE SUMMARY

The transport of neutrinos plays a dominant role in engine-driving core-collapse supernovae, but results from simulations of neutrino transport depend sensitively on the method and implementation details of neutrino transport codes. We have enhanced the open-source Monte Carlo neutrino transport code Sedonu and compared the results against a discrete ordinates transport code to verify for the first time multidimensional Boltzmann neutrino transport algorithms. We used these results to assess the accuracy of approximate two-moment transport schemes. We then motivate the use of time-independent Monte Carlo transport calculations in future large-scale time-dependent supernova simulations. In an orthogonal study, we also analyzed how gravitational waves—the only signal other than neutrinos that directly reflects conditions deep within the explosion—could bring it closer to the exact solution. Finally, we perform three-dimensional test calculations using snapshots from simulations using the open-source package Zelma [5,6].

METHODS & CODES

We use a variety of simulation codes to approach the problem from different directions. The key component of this project is the open-source Monte Carlo neutrino transport code Sedonu, which computes with high accuracy time-independent effects (e.g., instantaneous heating/cooling and lepton gain/loss rates) of neutrinos on one, two, and three-dimensional stationary fluid snapshots. With a paucity of direct data to validate against, we compare Monte Carlo results to those from a vastly different discrete ordinates method [2] to allow us to quantify how well we understand supernova neutrinos as we work toward a working theory of neutrino-driven core-collapse supernovae. We also assess the accuracy of popular approximate two-moment transport schemes [3] used in multidimensional core-collapse simulations. In addition, we couple Sedonu to the open-source supernova code GR1D [4]. The latter code simulates stellar core collapse with spherical symmetry using an approximate neutrino transport method, while Sedonu informs the transport approximation to bring it closer to the exact solution. Finally, we perform three-dimensional test calculations using snapshots from simulations using the open-source package Zelma [5,6].

RESULTS & IMPACT

Understanding the core-collapse supernova engine requires not only ever-increasing simulation size and complexity, but an understanding of how well these methods reflect the physics they are meant to simulate. With this in mind, we developed a new version of the open-source Monte Carlo neutrino transport code Sedonu, in which we implemented the random walk approximation. This improvement, along with many other feature additions and optimizations made as part of this project, makes the code capable of calculating highly accurate steady-state neutrino fields and interaction rates through the entire system without resorting to artificial boundary conditions. We performed the first multi-dimensional comparison of full Boltzmann neutrino transport methods, using Sedonu and the discrete ordinates code of [2]. Given that core collapse simulation results depend sensitively on small details of the neutrino transport implementation, this provides a much-needed measurement of the numerical errors associated with each method, and provides the first multi-dimensional code verification tools to the community. Observations of neutrinos and gravitational waves from nearby supernovae both give direct, unobscured information about the nature of the processes occurring deep within the collapsing star that lead to explosion. In addition to the above studies of neutrinos in supernovae, we performed a parameter study to determine what gravitational wave signals from rotating core collapse can tell us about the nature of matter at super-nuclear densities. In this study, we performed over 1,800 axisymmetric (2D) core collapse simulations. We used 18 different parameterizations for nuclear matter properties and over 100 rotation profiles, and concluded that differences in the descriptions of nuclear matter allowed by current constraints are not likely discernable by current gravitational wave detectors, even for a galactic supernova. We also established a simple universal relationship between the gravitational wave frequency and properties of the collapsed core. The resulting publicly available data set is an order of magnitude larger than those in previous studies and represents the first thorough exploration of both rotation and nuclear matter parameters.

WHY BLUE WATERS

Without access to the unique Blue Waters environment, the careful method development and verification in this project would not have been possible. The fellowship allocation enabled these calculations independently from other resources, since the calculations are far too large for local clusters and are separate from other resource requests. This project is targeted at working toward large-scale three-dimensional core-collapse simulations with a Monte Carlo-informed approximate method for neutrino transport. The domain-replication parallelism used in this approach requires both a large amount of compute time and a large amount of memory on each node. Blue Waters provides a unique environment with both of these components. In addition, the support staff with both domain and system expertise significantly simplify the task of implementing and optimizing our algorithms on Blue Waters, allowing user time to be spent on science and development rather than code porting.

PUBLICATIONS AND DATA SETS


A fifth-year Ph.D. student in physics at California Institute of Technology when he completed this research, Sherwood Richers was working under the direction of Christian Ott. He successfully defended his dissertation in June 2017 and will graduate in June 2018.
UNDERSTANDING THE ROLE OF HYDRODYNAMIC FLUCTUATIONS IN BIOMACROMOLECULAR DYNAMICS THROUGH THE DEVELOPMENT OF HYBRID ATOMIC-CONTINUUM SIMULATION

Seán L. Seyler, Arizona State University 2016-2017 Graduate Fellow

EXECUTIVE SUMMARY

Biological macromolecules are nanoscale structures that largely exist in ionic aqueous conditions. Proteins, in particular, are biopolymers that fold into functional 3D structures that—under the right solvent conditions and, possibly, lipid environment—can perform mechanosensitive work, and whose dynamics span femtosecond timescales (i.e., covalent bond oscillations) to beyond the millisecond regime (e.g., glucose transport across a lipid membrane). Explicit-solvent, atomistic molecular dynamics (MD) is necessary to fully capture solute-solvent interactions, but it is currently limited to microsecond timescales—orders of magnitude shorter than the timescale range of most processes of biophysical interest, including even the fastest conformational transitions (tens of microseconds). We are developing a hybrid atomic-continuum (HAC) method—with a view toward biomolecular simulation—coupling an MD engine to a novel discontinuous-Galerkin-based fluctuating hydrodynamics (FHD) solver; our current focus is to develop an FHD model capturing transport phenomena in dense fluids (e.g., water) arising at subnanometer scales.

RESEARCH CHALLENGE

There is growing interest in hybrid multi-physics simulations, where MD is used in a restricted subdomain requiring atomistic resolution, while an efficient FHD model replaces some, or all, of the solvent [1,2]. Nearly all HAC methods employ Landau–Lifschitz Navier–Stokes (LLNS), a system of stochastic equations describing mass, momentum, and energy transport subject to thermal fluctuations in the viscous stress tensor and heat flux vector [3]. LLNS assumes linear constitutive relations Newton’s law of viscosity (stress is proportional to velocity gradients) and Fourier’s law (heat flux is proportional to the temperature gradient). Substituting the constitutive relations into the momentum and energy equations leads to second-order spatial derivatives—the LLNS system is semiparabolic—implying, for example, an infinite speed of heat conduction; physically, however, we expect a finite transport speed. Though LLNS has proven a powerful approach to modeling many nanoscale phenomena, we expect the aforementioned assumptions to break down for dense fluids when simulation grid cells approach subnanometer dimensions (i.e., a water molecule) and hydrodynamic timescales of interest are comparable to collision times.

METHODS & CODES

Our HAC method currently uses the LAMMPS MD engine and is being developed in concert with a novel FHD model that extends LLNS. Development is currently focused on our FHD implementation, called HERMESHD (Hydraulic Equations and Relaxation Model for Extended Systems of HydroDynamics), which is based on Grad’s 5-moment (G51) approximation [4]. G13 includes additional equations describing time-dependent transport of viscous stresses (tensorial, symmetric/traceless—five equations) and heat flux (vectorial—three equations). Fluctuating terms can be incorporated either as stochastic fluxes in the momentum and energy equations, or as stochastic sources in the stress and heat flux equations. HERMESHD uses a discontinuous Galerkin (DG) spatial discretization of G13 with stochastic sampling performed on a per-cell basis (i.e., not individual DG Gaussian quadrature points). For temporal integration, we leverage FGI’s (fluctuating G13) hyperbolic structure to form a split-level, semi-implicit scheme: explicit time advance is carried out with a second- or third-order strong-stability preserving (SSP) Runge–Kutta method, while a locally implicit relaxation method is used to step over the time step constraint imposed by stiff source terms [5,6].

RESULTS & IMPACT

The additional equations in G13 generate hysteresis in stress and heat flux transport, leading to viscouselastic behavior on small spatiotemporal scales; in fluctuating G13 (FG13), stress and heat flux fluctuations, modeled as white Gaussian noise, are thus subject to memory effects and give rise to colored Gaussian noise. For nanoscale FHD and HAC simulations of dense fluids like water, we expect FG13 to produce transport phenomena that are neglected by LLNS, such as finite-speed thermoacoustic waves, which may have consequences for protein dynamics, especially as recent experiments have demonstrated a connection between heat released during enzyme catalysis and enhancement of diffusion [7]. As a first test of G13-based FHD, we augmented a linearized version of G13 (L13), with appropriate stochastic terms, to construct a fluctuating L13 (FL13) system. The numerical FL13 model was examined for qualitative correctness in the large collision frequency limit (i.e., solutions should relax to ordinary LLNS); Fig. 1 compares simulations with and without the fluctuating terms of an unstable nanoscale hydrodynamic jet for viscous, compressible, isothermal flow. The jet was initialized in a periodic domain with small, random velocity perturbations throughout and simulated until instability occurred. As expected, the jet went unstable in the FHD simulation first, though we have yet to quantify differences in the manner of jet breakup or in the overall flow fields between the two models. This comparison will soon be extended to include FL13 results for moderate collision frequency, and a complete comparison with the full FG13 system is being planned.

HERMESHD has a Python-wrapped library interface to facilitate rapid prototyping, Pythonic data manipulation, and simplified communication with external codes. Our HAC method employs a Python-based driver code to couple HERMESHD to the LAMMPS MD engine, though minimizing MD/FHD code interdependence is a priority. We plan to release the HERMESHD code under the GPLv3 license on GitHub.

WHY BLUE WATERS

We must verify that our G13-based FHD models reduce to LLNS when appropriate limits are taken and also pass common computational hydrodynamic benchmarks. Blue Waters enables these benchmarks—especially high-resolution 3D simulations that would otherwise create long turnaround times—to be executed quickly without hampering the development process. FHD and HAC model validation requires examining their correspondence to numerical experiments using gold-standard atomistic MD, necessitating large microcanonical ensemble simulations in order to avoid spurious effects introduced by thermo-/baro-statting, to mitigate long-range spatial correlations across periodic boundaries, and to obtain adequate statistical sampling. Furthermore, HAC simulations require matching the continuum region to the atomistic domain through empirical constitutive relations and equations of state, which must be calculated numerically from MD simulations of bulk fluids [8]. The multifaceted nature of our project requires running custom and existing codes as well as post-processing and visualizing diverse data, making access to Blue Waters staff and resources essential.

Sean L. Seyler, a fifth-year Ph.D. student in physics at Arizona State University, is working under the direction of Oliver Beckstein. He expects to receive his Ph.D. in December 2017.

Figure 1: Comparison of rightward horizontal velocities for a simple fluid around water density at STP for: Navier–Stokes (top) and fluctuating hydrodynamics (bottom) over 20 nanoseconds. Blue (red) corresponds to small (large) velocities (0–100 nm/ns). Simulation measures 300 nm per side (302 cells) using four-element linear quadrature and second-order Runge–Kutta integration.
THE IMPACTS OF HYDROMETEOR CENTRIFUGING ON TORNADO DYNAMICS

Ronald Stenz, University of North Dakota 2016-2017 Graduate Fellow

EXECUTIVE SUMMARY

Improving the understanding of tornadoes becomes more important as the population in regions prone to these violent weather phenomena increases. This research aims to advance our understanding of tornadoes by making the simulations used to study these destructive and dangerous storms more physically realistic. For the first time, we are quantifying the impacts that centrifuging of precipitation have on the vorticity budgets of these numerically simulated tornadoes. Preliminary findings have removed an unrealistic build-up of precipitation in the vortex center, while simulated tornadoes often have a relative maximum of precipitation in the center of the vortex. With millions and sometimes billions of dollars of damage caused by tornadoes every year, along with the risk of fatalities or serious injuries, a better understanding of these destructive weather events is needed to improve forecasting, preparedness, and mitigation of their impacts. By including the centrifuging of precipitation into the model we use to learn about tornadoes, our simulations become more consistent with what is observed in nature, facilitating the improvement of our understanding. Research findings have and continue to shape forecasting methods and plans for preparedness and damage mitigation; therefore, continued improvement of our understanding of tornadoes will provide results that can be used in operational settings, ultimately benefiting those living in regions prone to tornadoes.

RESEARCH CHALLENGE

The primary research challenge being addressed is the lack of centrifuging of precipitation in numerical simulations of tornadoes. In current simulations, precipitation follows the airflow, which creates an unrealistic build-up of precipitation in the vortex center. This, in turn, creates a source of negative buoyancy that potentially limits the stretching of vorticity in these simulated tornadoes. In nature, as precipitation moves around a circulation, such as a tornado, there is no force strong enough to keep the precipitation from moving outward, or being centrifuged, away from the circulation center. Observed tornadoes have a minimum of precipitation in the vortex center, while simulated tornadoes often have a relative maximum of precipitation in the vortex center.

With millions and sometimes billions of dollars of damage caused by tornadoes every year, along with the risk of fatalities or serious injuries, a better understanding of these destructive weather events is needed to improve forecasting, preparedness, and mitigation of their impacts. By including the centrifuging of precipitation into the model we use to learn about tornadoes, our simulations become more consistent with what is observed in nature, facilitating the improvement of our understanding. Research findings have and continue to shape forecasting methods and plans for preparedness and damage mitigation; therefore, continued improvement of our understanding of tornadoes will provide results that can be used in operational settings, ultimately benefiting those living in regions prone to tornadoes.

METHODS & CODES

Cloud Model 1 (CM1), which was designed for studying small-scale atmospheric phenomena such as thunderstorms [1], can run efficiently on supercomputers such as Blue Waters and was used for our simulations in this research. To quantify the impact that the inclusion of centrifuging has on tornado dynamics, simulations were first run without centrifuging. Just prior to the formation of a tornado, a checkpoint is used, allowing the model to be run both with and without centrifuging from this point to determine what impact the centrifuging of precipitation had on the tornado dynamics. To determine the magnitude of the centrifuging occurring, a centrifuging algorithm based on [2] uses trajectories released within the simulation to calculate the curvature of the flow and ultimately how quickly precipitation will be centrifuged, or moved outward, from the tornadic circulation. To quantify these impacts over a large sample size, atmospheric profiles of temperature, moisture, and wind from atmospheric soundings that were in close proximity to observed supercells [3] were used as the environmental conditions for our simulations of storms and their resulting tornadoes. A subset of these environments known to produce simulated tornadoes in previous research has been used for this study.

RESULTS & IMPACT

Both idealized simulations and a simulation of a full-scale storm with a resulting tornado have been completed with and without centrifuging. In simulations without centrifuging, the unrealistic precipitation from moving outward, or being centrifuged, away from the circulation center. Observed tornadoes have a minimum of precipitation in the vortex center, while simulated tornadoes often have a relative maximum of precipitation in the center of the vortex. With millions and sometimes billions of dollars of damage caused by tornadoes every year, along with the risk of fatalities or serious injuries, a better understanding of these destructive weather events is needed to improve forecasting, preparedness, and mitigation of their impacts. By including the centrifuging of precipitation into the model we use to learn about tornadoes, our simulations become more consistent with what is observed in nature, facilitating the improvement of our understanding. Research findings have and continue to shape forecasting methods and plans for preparedness and damage mitigation; therefore, continued improvement of our understanding of tornadoes will provide results that can be used in operational settings, ultimately benefiting those living in regions prone to tornadoes.

A fourth-year Ph.D. student in atmospheric sciences at the University of North Dakota, Ronald Stenz is working under the supervision of Matthew Gilmore. He expects to graduate in 2018.

Figure 1: Rain is beginning to accumulate in the simulated tornado center (left). With centrifuging turned on (right), there is an eye-like feature in the simulated tornado center consistent with radar observations of actual tornadoes.
GLASSY DYNAMICS AND IDENTITY CRISES IN HARD-PARTICLE SYSTEMS
Erin Teich, University of Michigan
2016-2017 Graduate Fellow

EXECUTIVE SUMMARY
Glass formation is a well-known outstanding mystery in the physical sciences: It is a phenomenon with no canonical thermodynamic description, yet one with wide-ranging industrial applications, occurring in materials as disparate as superconductors and sand. We utilized our Blue Waters allocation to help shed light on this murky topic, by simulating and investigating the dynamics of several glass-forming soft matter systems on time scales spanning about six orders of magnitude. We used our open-source simulation toolkit HOOMD-blue [1], scalable software built to run in parallel using domain decomposition on both CPUs and GPUs. We found that all systems studied displayed dynamical signatures indicative of cooperative rearrangement and consequent dynamic heterogeneity, both hallmarks of glass-forming behavior [2,3]. We examined local structure in these systems on other machines, and found a link between glass formation and a structural identity crisis, or competition between various local structural motifs.

RESEARCH CHALLENGE
Soft matter systems, in which thermal fluctuations are strong enough to drive particle rearrangements, are capable of self-assembling into a staggering variety of simple and complex crystalline structures. Often, however, no such assembly occurs, and the system remains disordered, displaying instead dynamical signatures characteristic of glass-forming materials. To develop robust methods of self-assembly, then, an understanding of glass formation in soft matter systems is needed. Beyond the realm of soft matter, glasses have proven useful materials to humans for over four millennia, for applications ranging from dinnerware to sand. We utilized our Blue Waters allocation to help shed light on this murky topic, by simulating and investigating the dynamics of several glass-forming soft matter systems on time scales spanning about six orders of magnitude. We chose our open-source simulation toolkit HOOMD-blue [1], scalable software built to run in parallel using domain decomposition on both CPUs and GPUs. We found that all systems studied displayed dynamical signatures indicative of cooperative rearrangement and consequent dynamic heterogeneity, both hallmarks of glass-forming behavior [2,3]. We examined local structure in these systems on other machines, and found a link between glass formation and a structural identity crisis, or competition between various local structural motifs.

METHODS & CODES
To help shed light on the dynamics and structural underpinnings of glass formation in soft matter systems, we chose to simulate and analyze nonassembling monatomic systems of hard polyhedral particles, with no interactions aside from those of excluded volume. We chose particle shapes consisting of tetrahedra with varying degrees of edge and vertex truncation. These are simple systems mediated solely by entropy maximization. Entropy maximization is known to drive phase transitions in a variety of systems, from those composed of hard spheres [4] to hard rods [5] to colloids and polymer depletants [6]. Recently, it was shown that entropy maximization drives systems of hard polyhedral particles toward a vast array of ordered structures upon compression [7], and a mechanism involving emergent, directional entropic forces was proposed to explain this ordering [8]. In [7], however, a number of systems displayed no order at all, instead exhibiting the signature slow dynamics of glassiness. In these cases, directional entropic forces drove systems toward glassy disorder rather than crystallization. Our project further explores these entropic effects.

We used the hard particle Monte Carlo (HPMC) simulation method [9] included in HOOMD-blue to sample our systems in the isochoric ensemble in equilibrium. In brief, this classic simulation method proposes trial moves for particles, and accepts or rejects them based on an acceptance criterion that dictates rejection if there are any overlaps created in the system, and acceptance otherwise. We simulated systems of 4996 particles on a single GPU at a variety of densities, and subsequently measured structural and dynamical information. To gather trajectory information at a wide range of time scales for later analysis, we wrote our system trajectories to disk remarkably frequently—every 10 Monte Carlo (MC) sweeps—and collected data for the far longer period of 100 million MC sweeps. We produced trajectories as large as 3.25 terabyte (TB) MC sweeps per simulation, each containing about 10 million simulation frames.

RESULTS & IMPACT
For each trajectory, we measured several dynamical parameters known in the glass community to be indicators of glass-forming behavior, and confirmed that our hard particle systems are indeed glass-formers. These were the mean-squared displacement of all particles in the system, the van Hove autocorrelation function, the non-Gaussian parameter characterizing the particle displacement distribution, and the self-part of the intermediate scattering function [2,3]. The mean-squared displacement and intermediate scattering function for a sample system are shown in Fig. 1. Plateaus in both parameters at intermediate times and increasing densities indicate caging behavior typical of glass-formers; at these times, particles are trapped on average by a cage of their surrounding neighbors, and only escape due to cooperative structural rearrangements and consequent dynamic heterogeneity. The mean-squared displacement shows this caged behavior directly, while the plateau in the self-intermediate scattering function (a Fourier transform of the real-space van Hove autocorrelation function) shows a slowing-down of the relaxation of the system.

We have determined via simulations on other platforms that observed glass-forming tendencies in these systems are related to competition between local structural motifs, mediated by entropic forces; future work will strive to make more concrete the connection between local structure and observed heterogeneous dynamics.

WHY BLUE WATERS
HOOMD-blue HPMC’s parallel nature on both CPUs and GPUs allowed us to take unique advantage of the computational resources on Blue Waters, and, in particular, its GPU resources. We also produced tens of TB of raw data that we needed to sift through later, and we were able to accomplish this by storing it all on Blue Waters’ generous Online file system, analyzing it, and storing resultant analysis results somewhere more permanently. Without the 500 TB quota afforded to us by Blue Waters, we would find it far more difficult to explore such fine-grained dynamical phenomena.

Erin Teich is a fifth-year Ph.D. student in applied physics at the University of Michigan. She is working under the direction of Sharon C. Glotzer, and expects to graduate in June 2018.
EXECUTIVE SUMMARY

In this project, we use the state-of-the-art particle-in-cell code OSIRIS to study a fundamental plasma process known as magnetic reconnection, which plays a key role in the evolution of plasmas from astrophysics to the laboratory. Reconnection is a promising candidate for producing the energetic particle distributions associated with explosive astrophysical sources; however, the particle acceleration properties of reconnection are not fully understood. Recently, laser-driven plasma experiments have been used to study reconnection in conditions relevant for astrophysics. By modeling these experiments on Blue Waters we were able to show that for current experimental conditions, electrons can be accelerated by reconnection with sufficient quantity and energy to be detected in the laboratory, which opens the way for new experimental studies of particle acceleration from reconnection.

We are also working on developing a new simulation method called simplex-in-cell, which may improve the accuracy and reduce the expense of plasma simulations.

RESULTS & IMPACT

From the results of these simulations we were able to show clearly that for current experimental conditions, electrons can be accelerated by reconnection with sufficient quantity and energy to be detected in the laboratory. For the conditions of recent experiments, the nonthermal electrons can be accelerated to energies more than an order of magnitude larger than the initial thermal energy. The nonthermal electrons are primarily accelerated by the reconnection electric field near the X-points, which establishes a distribution of energies that resembles a power-law spectrum. After being energized, the electrons can also become trapped inside the plasmoids (magnetic islands) that form in the current layer and gain additional energy from the electric field arising from the motion of the plasmoid. By comparing simulations for finite and infinite periodic systems, we were able to demonstrate the importance of particle escape on the shape of the spectrum.

Based on our findings, we derived an analytical estimate of the maximum electron energy and a threshold condition for observing superthermal electron acceleration in terms of the initial plasma conditions, which can now be tuned in future experiments to optimize the particle acceleration. Through the use of 3D simulations (Fig. 1) we studied the angular distribution of the accelerated particles and constructed synthetic detector spectra to determine experimental signatures. These results provide new insight into the physics of reconnection and particle acceleration, and are now helping to guide several experimental programs in the United States.

Due to limitations such as noise from artificial two-body collisions and the computational expense associated with the large number of particles required to accurately capture the development of nonthermal tails in the particle distribution, multiscale PIC simulations such as those used to study laser-driven reconnection are extremely challenging. It is thus critical to work on improved methods that could reduce the computational expense of these simulations and improve their physical accuracy. To this end, we are also developing a novel method for plasma simulation, which we refer to as simplex-in-cell (SIC). The foundation of SIC is an interpretation of the simulation particles as the vertices of an unstructured mesh that traces the evolution of the plasma distribution function in phase space [3]. This enables a new discretization using deformable phase space volume elements rather than fixed-shape, localized particles. We are using the SIC interpretation of the simulation particles for data analysis and visualization of standard PIC simulations performed using OSIRIS, and have been able to show that in certain regimes SIC can reach a given noise level using one thousand times fewer simulation particles than standard methods (Fig. 2). Future work will involve implementing SIC directly in the simulations to reduce noise and unphysical artifacts.

WHY BLUE WATERS

This project required the use of large-scale 2D and 3D simulations with sufficient size and resolution to bridge the multiscale physics, from fluid dynamics to the kinetic microscopic processes. These computationally demanding simulations can require billions of simulation particles, and demand the cores, processes. These computationally demanding simulations can require billions of simulation particles, and demand the cores, memory, and communication performance available on Blue Waters. The quick support from the NCSE staff on technical issues helped me to maximize my productivity on the machine.

PUBLICATIONS AND DATA SETS


Samuel Totorica, a fifth-year Ph.D. student in physics at Stanford University, is working under the direction of Tom Abel and Federico Fiuza. He expects to graduate in 2018.
SCIENCE

TRANSFORMATIVE SCIENCE

From the beginning of the project, the primary goal of Blue Waters has been to enable transformative science and engineering research by giving research teams access to a large number of powerful compute nodes and large data capacity. The target is to have at least 40 percent of the system’s capacity dedicated to research challenges that cannot be addressed on smaller-scale systems. These "capability" jobs are defined as those that are either large, using 512 to 4,528 XE nodes (i.e., 16,384 to 144,896 core equivalents) or 64 to 845 XE nodes, or very large, using more than 4,528 XE nodes (i.e., greater than 144,896 core equivalents), or more than 845 XE nodes. Large jobs are equal to the largest computations possible on many other compute resources, and very few systems in the world other than Blue Waters can support very large jobs. The project generally surpassed this 40 percent benchmark over the past year, though the average is lower than previous years due to more data-focused, "big data" single node workloads being added to the mix of projects awarded time on Blue Waters. Blue Waters continues to prioritize the "capability" jobs that allow researchers to push the boundaries of their disciplines and to achieve results that would not be possible without Blue Waters (see Table 2). Several science teams are regularly running jobs that use 2,048; 4,096; and 8,192 nodes on Blue Waters.

Many researchers who use Blue Waters have told us that their work would simply not be possible without this Leadership-Class system: "On Blue Waters, what would be a tremendous calculation simply cannot be accomplished in a timely manner on less powerful machines. Without Blue Waters, iNEMO re-engineering, or redesign, such as implementing GPU functionality via OpenACC, or alternatives to MPI collective operations. Such work is tracked via a coordinated work plan that is developed between the PoC and the science team to clearly indicate the scope and scale of the work involved, including milestones and deliverables. Work plans are reviewed and approved by the BWPO."
These competitive fellowships include both financial support and required to respond to and resolve service requests, which include traditional “trouble tickets” as well as requests for basic issues, and training activities for undergraduate students, graduate and for improving work. PAID forges for key service metrics over the past year. Many users have mentioned the crucial role of Blue Waters support staff in enabling their work:

“One key aspect of the Blue Waters facility was the availability of the Blue Waters staff, who provided invaluable technical insight and frequent advice in order to tune the performance of our software.”

Greg Voth, University of Chicago

“The Blue Waters Fellowship has opened new doors for me by providing me first-hand experience with the high-performance computing resources necessary for cutting-edge science in oceanography. The expert staff at NCSA have been extremely responsive to my questions and needs and have helped me grow from a novice into a proficient user with plans to continue using high-performance computing to study physical–biological interactions across marine ecosystems.”

Maureen Brooks, Blue Waters Graduate Fellow, University of Maryland College Park

The Petascale Application Improvement Discovery (PAID) program is another way the Blue Waters project supports advanced application development and improvement work. PAID forged collaborations between improvement Method Enablers (IMEs) and science and engineering research teams, helping them to create and implement new technologies that improve application performance. (A detailed report on the PAID program can be found elsewhere in this volume.) Additionally, the annual Blue Waters Symposium provides an excellent opportunity for cross-pollination among research teams and across disciplines. Attendees report that the facilitation of personal networking, interaction, and discussion is a particular strength of the symposium and that the event provides useful resources and information.

The Blue Waters project also strives to develop the next generation of computational researchers by supporting education and training activities for undergraduate students, graduate students, and professional staff. Each year, an elite group of Ph.D. students from a diverse range of domains and from institutions across the country are awarded Blue Waters Graduate Fellowships. These competitive fellowships include both financial support and access to the Blue Waters system to substantially advance the fellows’ computational research. To date, 36 young scholars have participated in the program, which over three years will award more than $1.3 million in direct financial support and over 60 million core equivalent hours to support graduate research. The 36 fellows are from 24 states and 26 institutions, including eight institutions from EPSCOR (Established Program to Stimulate Competitive Research) states and one that is a minority-serving institution.

In order to ensure that the fellows are able to effectively utilize the power of the Blue Waters system, each fellow is paired with a member of the Blue Waters SEAS group; these “points of contact” assist the fellows with essential activities such as porting and tuning codes and with any support requests. In the December 2015 review of the Blue Waters project by the National Science Foundation, the review panel pointed to this in-depth PoC model as a “best practice” for fellowship and education programs.

Some of the impressive research results achieved by the 2016–2017 Blue Waters Graduate Fellows are showcased in this annual report. Many fellows highlighted the impact that Blue Waters and their fellowship experience will have on their research and their future career path:

“The Blue Waters Graduate Fellowship has not only played a pivotal role in helping me achieve my doctorate, but it has also exposed me to a variety of fields and topics related to high-performance computing that I would have otherwise not experienced as part of my degree plan. I believe this breadth of knowledge will give me a competitive edge as I continue to the next stages of my career.”

Justin Drake, University of Texas Medical Branch

Approximately 20 students are selected each year for the Blue Waters Undergraduate Internship program, which includes an intensive two-week workshop covering the fundamentals of high-performance computing (HPC) and engages undergraduate students in year-long projects in which they apply HPC to problems in science, mathematics, and engineering with guidance from faculty mentors. To date, the Blue Waters project has funded 120 interns from 64 unique institutions in 27 states, the District of Columbia, and Puerto Rico. Of the 99 interns, 52 are women and/or minorities. Thirty-five of the 69 institutions were in EPSCOR regions, and 17 were classified as minority-serving institutions. Blue Waters has directly benefited people in 49 states (Hawaii is the outlier) as well as Puerto Rico. Blue Waters activities have also benefited individuals in more than 20 countries working as part of projects led by principal investigators in the United States. External evaluation has found that both interns and their advisors find the program beneficial. Some participants have reported altering their educational and career paths or embarking on a path with greater confidence thanks to the internship experience.

Another aspect of the Blue Waters engagement with students and early career workers is the fact that from the beginning of the project, 56.4 percent of the people using Blue Waters were students and/or postdocs. This includes 881 graduate students, 125 undergraduate students, 17 high school students, and 350 postdocs.

The project also funds, in partnership with the XSEDE project, the content in HPC University (http://hpctechnology.org), which is focused on developing educational materials for and about petascale computing and analysis. This year, the HPC University had 67,531 petascale learning modules and training materials downloaded.

In partnership with the University of Buffalo, Blue Waters published a extensive study of the work that is using Blue Waters over the first three and a half years of service. This study shows many of the characteristics of how the applications use the different architectural features of Blue Waters such as the amount of memory used and the I/O characteristics of the workload. The full report is available to provide much more detail and is available at https://arxiv.org/ftp/arxiv/papers/1703/1703.00924.pdf

Because of excellent overall project management, the partners, leadership and staff of the Blue Waters project is pleased that the National Science Foundation has approved continued operations of the Blue Waters project through at least March 2019. With this one-year-plus extension of full-service operations, the project will continue to serve the research needs of the scientific community as well as undergraduates and graduate education activities. The dedicated Blue Waters team at NCSA looks forward to continuing productive collaborations with our science and engineering partners as they continue to innovate and discover.

Table 2: Metrics for service request response and resolution time show the Blue Waters support staff are meeting and often exceeding the metric expectations throughout the year.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Service requests are recorded and acknowledged in a timely manner</td>
<td>95% of partner service requests are acknowledged by a human</td>
<td>96.6% of partner service tickets had a human response within four business hours</td>
<td>Meets Expectations</td>
<td>Exceeds Expectations</td>
<td>Meets Expectations</td>
<td>Exceeds Expectations</td>
</tr>
<tr>
<td>Most problems are solved within a reasonable time</td>
<td>80% of partner service requests are addressed within three working days, either by: - resolving them to the partner’s satisfaction within three working days, or: - for problems that will take longer, by informing the partner how the problem will be handled within three working days (and providing periodic updates on the expected resolution)</td>
<td>83.6% of partner service requests were resolved within three business days</td>
<td>Meets Expectations</td>
<td>Exceeds Expectations</td>
<td>Meets Expectations</td>
<td>Exceeds Expectations</td>
</tr>
<tr>
<td>82.4% of partner service requests were resolved within three business days</td>
<td>85.1% of partner service requests were resolved within three business days</td>
<td>81.5% of partner service requests were resolved within three business days</td>
<td>Meets Expectations</td>
<td>Exceeds Expectations</td>
<td>Meets Expectations</td>
<td>Exceeds Expectations</td>
</tr>
</tbody>
</table>

Table 2: Metrics for service request response and resolution time show the Blue Waters support staff are meeting and often exceeding the metric expectations throughout the year.
The Science and Engineering Team Advisory Committee (SETAC) brings together a diverse group of scientists and engineers who represent the various Blue Waters science and engineering research teams using Blue Waters and the breadth of research powered by the Blue Waters system. The committee provides guidance and assessment to help the Blue Waters project deliver the best possible performance and services that will in turn assist research teams in achieving breakthrough results.

The SETAC makes recommendations on technical directions, strategies, and management while identifying potential challenges for petascale applications. As users themselves, the SETAC members also provide advice for solving common issues that arise from moving applications to Blue Waters and from system software at scale.

The SETAC convenes three to four times per year and is available to the Blue Waters Project to provide guidance and advice as needed throughout the year. Members are nominated and typically serve two-year terms.

SETAC Members

PETASCALE COMPUTING RESOURCE TEAMS
- Paul Woodward, Physics and Astrophysics, University of Minnesota (SETAC Chair)
- Susan Bates, Oceanography & Climate Dynamics, NCAR
- Manuela Campanelli, Physics and Astronomy, Rochester Institute of Technology
- Tom Cheatham, Chemistry, University of Utah
- David Ceperley, Physics and Material Science, University of Illinois at Urbana-Champaign
- Tiziana Di Matteo, Physics and Cosmology, Carnegie Mellon University
- Said Elghobashi, Mechanical and Aerospace Engineering, University of California, Irvine
- Tom Jordan and Philip J. Maechling, Seismology, University of Southern California
- Paul Morin, Earth Sciences, University of Minnesota
- Brian O’Shea, Physics and Astronomy, Michigan State University
- Nikolai Pogorelov, Space Science, University of Alabama, Huntsville

GREAT LAKES CONSORTIUM FOR PETASCALE COMPUTATION TEAMS
- Ben Rogers, Information Technology Research Services, University of Iowa

INDUSTRY TEAMS
- Rick Arthur, General Electric Global Research, Advanced Computing

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN TEAMS
- Athol Kembhavi, Astronomy, University of Illinois at Urbana-Champaign
- John Hawley, Image Uncertainty Quantification
- Fatemeh Khalil-Araghi, Paracelular Transport Mechanism
- Seid Kord, Improving Performance of an Implicit FEA
- William T.C. Kramer, Blue Waters Workload Analysis
- Praveen Kumar, Ecohydrologic dynamics-climate change
- Gregory Michael Laskowski, LES of a Film Cooled Turbine Vane
- Vincent Le Chenadec, Numerical simulation of two-phase flows
- Athol Kemball, Tilted Disks Around Black Holes
- Jong Lee, Numerical study of many body transition
- David Luiz, Genotyping for very large cohorts
- Mark Neubauer, Enabling Discoveries at the LHC
- Jian Peng, Deep learning for structure prediction
- Venkat Raman, Solvers for Turbulent Combustion
- Paul Ricker, AGN Feedback in Galaxy Clusters
- William Sanders, Evaluating Theft in Smart Grids
- Justin Sirignano, High-Speed Link Simulation
- Mark Neubauer, Large-scale Deep Learning
- Marshall Stageberg, Atmospheric Response to a Wildland Fire
- Victoria Stodden, Variability of halo properties in Enzo
- Matthew Turk, The Computational Microscope
- Emad Tajkhorshid, Chemistry in the Early Universe

The projects listed here had a Blue Waters allocation during this reporting period but did not submit a report for the project.

QC HathiTrust Digital Text Archive
- Scott Althaus, Spot Scanning Proton Therapy Optimization
- Chris Beltran, EpiSimdemics Simulations of Influenza
- Abhinav Bhatele, SC16 SciNet Cosmology Experiment
- Franck Cappello, Simulations of Large-N Star Clusters
- Sourav Chatterjee, HathiTrust Public Research Dataset
- J. Stephen Downie, Reservoir Simulation on Blue Waters
- Vadim Dyadechko, Mechanics of Biofilms
- Ahmed Elbanna, Co-Evolution of Galaxies and Black Holes
- Claude-Andre Faucher-Giguere, Simulation and Scaling of MultiPhysics
- Jonathan Freund, Scalable Social Simulations
- Leslie Gasser, Protein Sequence Similarity Networks
- John Gerlt, Global crop forecasting from satellite
- John Hawley, Cloud Computing Systems at Scale
- John Hawley, Global crop forecasting from satellite
- John Hawley, Cloud Computing Systems at Scale
REFERENCES

SPACE SCIENCE

Balsara, Dinshaw

Holden, Gilbert

Huerta, Eliu (1)


Huerta , Eliu (1)

Lentz, Eric

Mckenna, Ryan

Pogorelov, Nikolai

Quinn, Thomas

Royer, Vladimir

Schive, Hsi-Yu
Shapiro, Stuart


GEOSCIENCE

DiGiaroma, Larry


[9] Garcia, Marcello


Lasher-Trapp, Sonia


Liu, Lin


Ackerman, David


Akono, Ange-Therese


Allain, Jean Paul


Valoschi, Albert J.


West, Matthew


Wuebbles, Donald J.


Um, Jensikhu


Valoschi, Albert J.

Gazzola, Mattia


Huff, Kathryn


Militzer, Burkhard

Mackenzie, Paul

Rani, Sarma

Schlief, Andre


Riedl, Caroline

Advancing the Science of High Pressure Lattice Field Theory (World Steel Association, Brussels, Belgium, 2014).


Militzer, Burkhard

Mackenzie, Paul

Rani, Sarma


Yeung, Pui-kuon


Thomas, Brian

Steel Statistical Yearbook (World Steel Association, Brussels, Belgium, 2014).


Wagner, Lucas

Wagner, Lucas


Yeung, Pui-kuon


Thomas, Brian

Steel Statistical Yearbook (World Steel Association, Brussels, Belgium, 2014).


Wagner, Lucas


Gropp, Bill


Olson, Kevin


Caetano-Anolles, Gustavo


PLoS Computational Biology, 8.6 (2012), e1002572. DOI:10.1371/journal.1.

chemistry, and molecular evolution. 

[4] Caetano-Anolles, G., W. Wang, and D. Caetano-
Anolles. Structural phylogenetic analysis: the origin of the gene code and uncovers the evolu-
tional impact of protein flexibility. 

PLOS One, 7.5 (2012), e34722.


[7] Ravasz, E., et al., Hierarchical organization of mod-
ularity in metabolic networks. 

Carnevale, Vincenzo 1400–1415.

[1] Wu, M., et al., Genetic determinants of in vivo ftn-


[1] Carlsson, E., M. Liao, Y. Cheng, D. Julius, TRPV1 struc-


[1] Ribiero, I. V., et al., QuElMOD—Incorporative Molecu-

[1] Bernacki, R. C., M.C.R. Mels, and K. Scholtz, Enhanced sampling techniques in molecular dyna-


[1] Dong, H., et al., Pure waters regulate ion per-


[1] Faivre, G., M.L. Klein, and J. Hénin, Using collec-

[1] Co, E., M. Liao, Y. Cheng, D. Julius, TRPV1 struc-

[1] Skolnick, J., and H. Zhou, Why is There a Glass Ceiling for Threading Based Protein Structure Predic-

[1] Luthey-Schulten, Zaida et al., Adaptive simulation of hybrid stochastic and deterministic models for biochemi-


Richers, Sherwood


Seyler, Sean


Stenz, Ronald


Teich, Erin


Totorica, Sam


Blue Waters is supported by the National Science Foundation