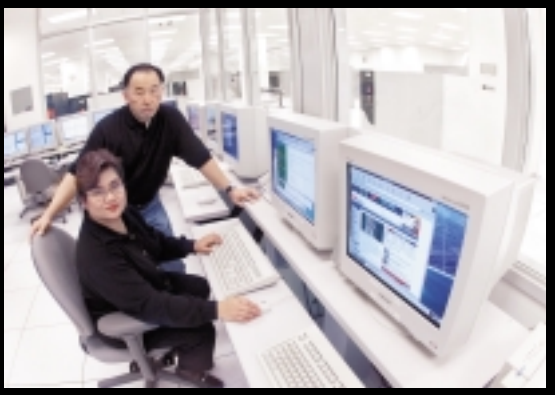
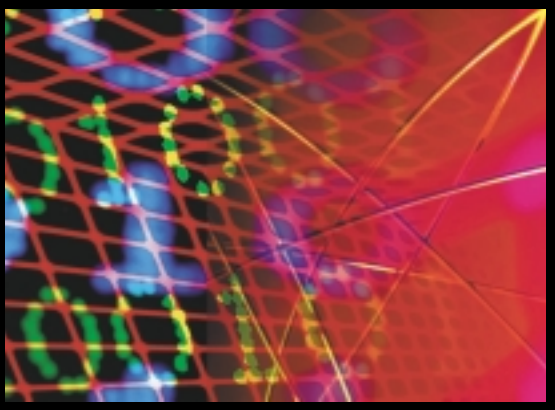
NERSC

NATIONAL ENERGY RESEARCH SCIENTIFIC COMPUTING CENTER

2000 Annual Report





Berkeley Lab
Ernest Orlando Lawrence
Berkeley National Laboratory

NATIONAL ENERGY RESEARCH SCIENTIFIC COMPUTING CENTER

2000 Annual Report

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

This work was supported by the Director, Office of Science,
Office of Advanced Scientific Computing Research
of the U.S. Department of Energy
under Contract No. DE-AC 03-76SF00098.

LBNL/PUB-47197, January 2001

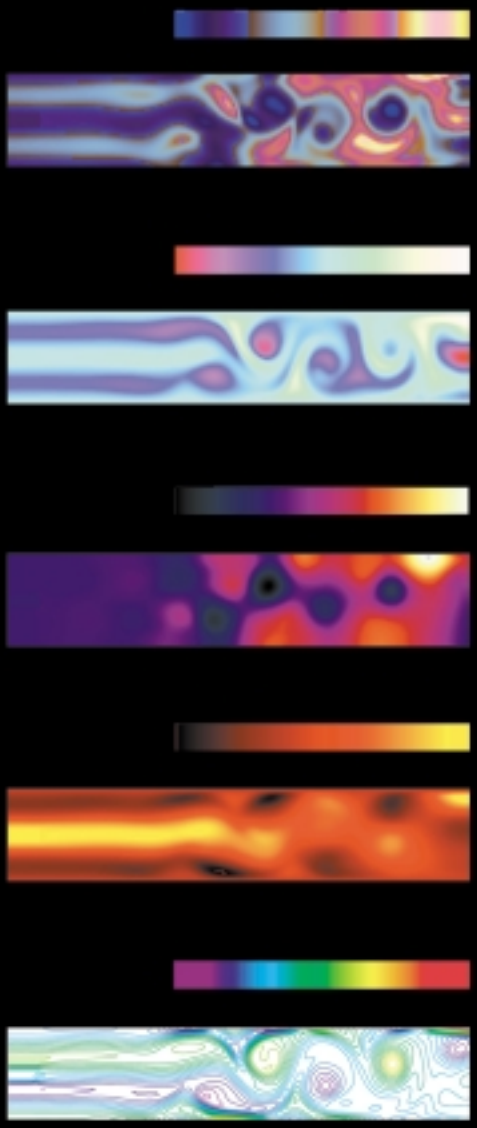
NERSC aspires to be a world leader in accelerating scientific discovery through computation. Our vision is to provide high-performance computing tools to tackle science's biggest and most challenging problems, and to play a major role in advancing large-scale computational science and computing technology. The result will be a rate of scientific progress previously unknown.

NERSC's mission is to accelerate the pace of scientific discovery in the Department of Energy Office of Science (SC) community by providing high-performance computing, information, and communications services. NERSC has a threefold strategy for increasing researchers' productivity:

- Providing leading-edge platforms and services that make NERSC the foremost resource for large-scale computation within DOE.
- Introducing the best new computer science tools to SC researchers.
- Providing intellectual services (for example, development of innovative algorithms, simulations, and visualization techniques) that make these complicated technologies useful for computational science.

Contents

Director's Perspective



Snapshots of two-dimensional jet flow in a tokamak divertor simulation, showing density, temperature, pressure, velocity, and vorticity contours. See page 64 for details.

There was not much to see five years ago when I arrived at the newly relocated NERSC in Berkeley — only a handful of new employees; anxious Berkeley Lab and DOE staff worrying whether their boldness would really pay off; major construction on the first floor of Building 50, which as yet bore no resemblance to a computer room; and trenches filled with rainwater in the parking lot. It seems remarkable how the pioneers of the new NERSC were undaunted by those challenges, and how, with neither staff nor computers in place, we boldly proclaimed that we would reinvent the high performance computing center. And, by all accounts, we did it.

By 1998 it became clear that the space and power requirements of the next few generations of high performance computing platforms would require a much larger computer room. Our Berkeley Lab colleagues in Operations and Facilities located and remodeled a cost-effective new facility in record time. In the fall of 2000, we moved most of our high-performance computing platforms to the new Oakland Scientific Facility, and our newest platform was installed in the first week of 2001, all on schedule.

The new platform will provide an unprecedented 3.8 Tflop/s peak performance for the DOE Office of Science computational community. While many of our colleagues at other sites brag about terascale computing, I believe that NERSC will be the first site where users can expect to see their applications routinely perform at teraflop/s level. This will be the culmination of years of work at NERSC to improve the utilization of highly parallel platforms, and to provide the tools that allow efficient execution of jobs requiring 512 processors and more. At the same time, our storage capability is nearing the petabyte level, thanks to continuing improvements over the last few years.

While integrating these new technologies, we continued to maintain the highest standards of service, and again enabled our community of users to attain breakthrough scientific results. The most notable accomplishment, among the many documented in this annual report, is a cover story in Nature, backed by data analysis carried out at NERSC.

In November, at the SC2000 conference in Dallas, Berkeley Lab and NERSC released three software CDs: Berkeley Lab AMR, Akenti, and Berkeley Lab VIA Software (M-VIA and MVICH). Highlights of these projects and many of our other R&D efforts are presented in this annual report, demonstrating the benefits of combining a computing facility with research and development in one organization.

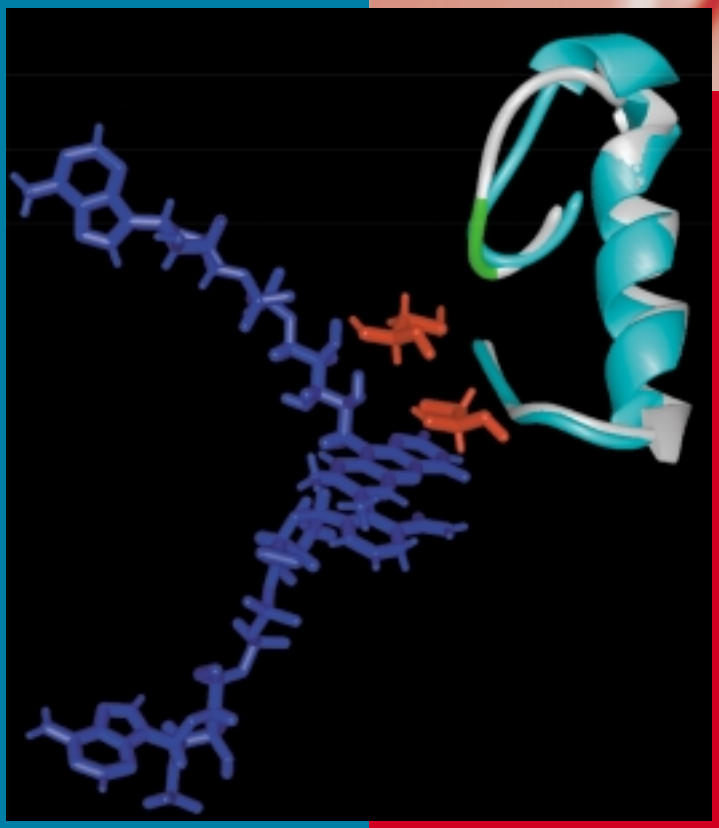
Horst D. Simon,
Division Director of NERSC

After several years of planning, computational science in the DOE Office of Science finally received a big boost through the funding of the Scientific Discovery through Advanced Computing program (SciDAC). For NERSC's clients, SciDAC offers a once-in-a-decade opportunity to demonstrate the applicability of past research and to engage in bold new projects. NERSC itself has the opportunity to further develop and then deploy the results of the computer science research of the last few years, enabling a whole new generation of computational science.

I am pleased that the Office of Advanced Scientific Computing Research is re-examining its portfolio of research activities, and that NERSC will have the opportunity to develop a new five-year plan. Given our accomplishments of the last five years, I have no doubt that we will develop a first-rate strategy for NERSC and computational science in DOE, and that we will reinvent the high performance computing center yet again.

With these exciting times ahead of us, I am grateful to our DOE Office of Science sponsors for their continued endorsement of our ambitious plans. It continues to be a pleasure to collaborate with the NERSC Users Group and its executive board members. I would like to thank them for their continued support, especially for their effort to produce the next "Greenbook" documenting the computational requirements of the Office of Science community. My special thanks and congratulations, as always, go to the NERSC staff for their skill, dedication, and tireless efforts to make NERSC the best scientific computing resource in the world.

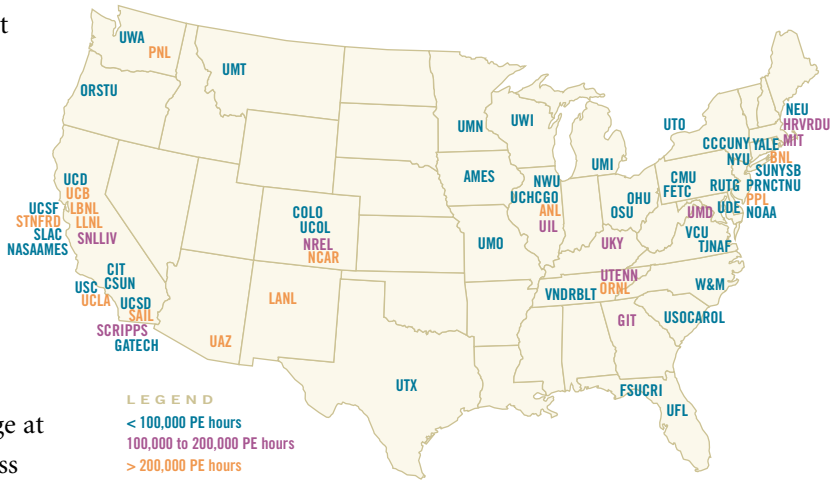
Computational Science



Molecular dynamics simulation of protein folding. see page 48 for details.

As a national facility for scientific research funded by the Department of Energy, Office of Science (DOE SC), NERSC annually serves about 2,400 scientists throughout the United States (Figure 1). These researchers work in DOE laboratories, universities, industry, and other Federal agencies (Figure 2). Computational science conducted at NERSC covers the entire range of scientific disciplines, but is focused on research that supports the DOE's mission and scientific goals (Figure 3).

Figure 1. NERSC FY00 massively parallel processing (MPP) allocations by site.

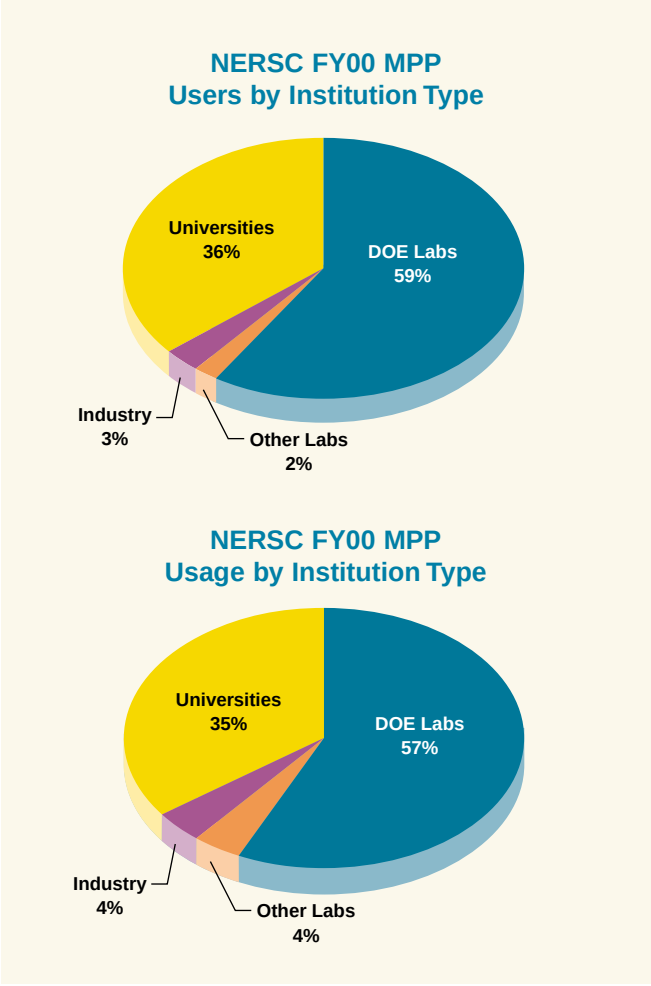


Allocations of computer time and archival storage at NERSC are awarded to research groups, regardless of their source of funding, based on an annual review of proposals. As proposals are submitted, they are subjected to peer review to evaluate the quality of science, how well the proposed research is aligned with the mission of DOE SC, and the readiness of the specific application and applicant to fully utilize the computing resources being requested.

The NERSC Program Advisory Committee (see Appendix B) is responsible for the scientific peer review process used to allocate 40 percent of NERSC's computing resources. The peer review and resource allocation process for the remaining 60 percent is managed directly by the DOE SC programs, reflecting their mission priorities.

Two other groups provide general oversight: the NERSC Policy Board (Appendix A) advises the Berkeley Lab Director on the policies that determine the impact and performance of the NERSC Center, and the NERSC Users Group (Appendix C) advises the NERSC management and provides feedback from the user community. This section of the Annual Report gives an overview of the research supported by NERSC and points out some of the year's achievements, which are described further in the Science Highlights section.

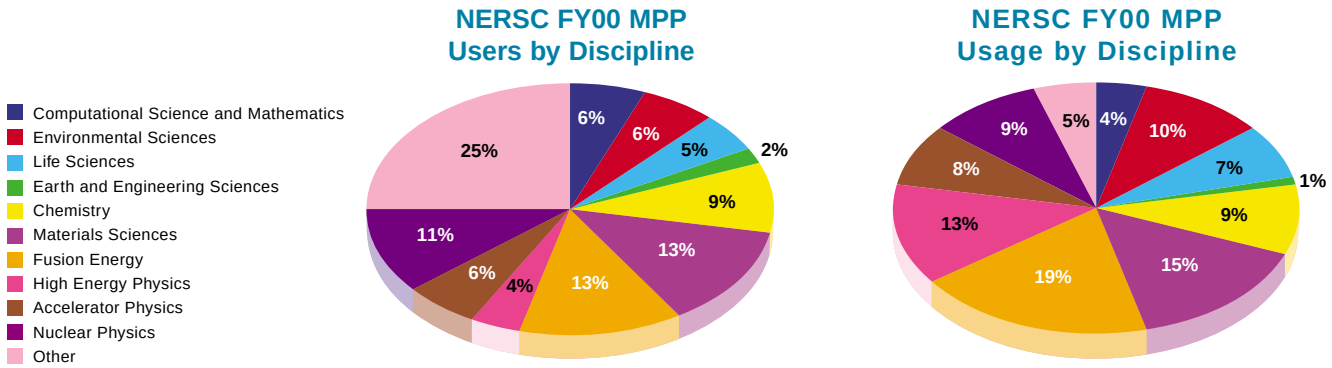
Figure 2. Percentage of NERSC MPP users and usage (computing time) by institution type.



Advanced Scientific Computing Research

DOE's Office of Advanced Scientific Computing Research (OASCR) supports a number of projects in computer science and applied mathematics. NERSC staff take the lead or participate in several of these projects, including

Figure 3. NERSC MPP users and usage by scientific discipline.



research on turbulent reacting flow by our Center for Computational Sciences and Engineering; development of software packages that enable ordinary C or Fortran computer programs to perform arithmetic with 32 or 64 decimal digit accuracy (this software is being used to explore an unresolved question regarding the regularity of vortices); developing computational tools for linear algebra problems that are ubiquitous in computational science and engineering; and researching subspace-based techniques for information retrieval, such as latent semantic indexing.

Other OASCR-funded research includes testing the scalability of parallel discrete event simulations, which are used in a wide variety of fields, from switching of cellular communications networks to studies of material failure. Another group is studying instabilities in turbulent mixing, an important issue for fluid dynamics that impacts such questions as the rate of heat transfer by the Gulf Stream, resistance of pipes to fluid flow, combustion rates in automotive engines, and the late time evolution of a supernova. Research jointly funded by OASCR and other DOE offices includes developing global optimization approaches to protein fold refinement, and developing a new generation of electron-atom and electron-molecule scattering codes that are capable of treating all details of electron impact ionization.

Basic Energy Sciences

NERSC provides computational support for a large number of materials sciences, chemical sciences, geosciences, and engineering research projects sponsored by DOE's Office of Basic Energy Sciences.

Research in the chemical sciences will have important impacts on energy efficiency, pollution prevention, and environmental restoration. This year researchers showed that molecular-based simulations of complex fluids can be used to predict properties such as the viscosity index and pressure-viscosity coefficient of lubricants. The ability to predict these properties via simulation will lead to the molecular design of improved lubricants, which will result in better energy efficiency.

A new area of research at NERSC this year was numerical simulation of combustion in homogeneous charge compression ignition (HCCI) engines. HCCI engines are an attractive alternative to diesel engines, offering the potential for diesel-like efficiencies, while producing extremely low emissions without expensive aftertreatment. Computational simulations will provide crucial direction to HCCI design efforts.

Electronic structure theory has emerged as a valuable counterpart to direct experiments for the study of reactive species that may not be characterized easily (if at all) in the laboratory — for example, large polycyclic aromatic hydrocarbon (PAH) cations, which arise in combustion processes as intermediates to the formation of soot particles, and which are also believed to play a significant role in interstellar carbon chemistry. Electronic structure research on the chemical and conformational transformations of biomolecules is beginning to yield a novel microscopic picture of biochemical dynamics. Understanding biological chemical processes at the atomic level with this and other methods will have a major impact on the drug and biotechnology industries.

In the materials sciences, Georgia Tech physicist and NERSC user Uzi Landman won this year's Feynman Prize in Nanotechnology (Theoretical) for his pioneering work in computational materials science for nanostructures (see page 39). Such computer modeling provides deep insights into the nature and properties of matter at the nanoscale, and is essential in predicting what could be built at the molecular level, reducing time spent on expensive laboratory experiments.

New modeling tools developed by Landman and other researchers are also making it possible to begin bridging the gaps between scales, so that the effects of atomic and microscopic phenomena can be seen at the macroscopic scale. As these modeling tools mature, they will make important contributions in both technological and environmental areas, including carbon sequestration, the development of high-temperature superconductors, miniaturization of electronic and mechanical devices, development of lasers and sensors, design of novel logic gates and information storage strategies, control of friction under extreme conditions, and the design of electric motors with reduced weight and improved performance.

Biological and Environmental Research

DOE's Office of Biological and Environmental Research is a major supporter of global climate studies as well as computational medical and biological research using NERSC resources. In addition to providing computational support, NERSC is also the repository for the archive of the Program for Climate Model Diagnosis and Intercomparison (PCMDI). Complete data sets from PCMDI are publicly available to researchers through NERSC (see http://www-pcmdi.llnl.gov/modeldata/ PCM_Data/pcgdahome.html).

Ongoing improvement in the precision of climate modeling codes is resulting in a growing understanding of the factors contributing to global climate change as well as better modeling of regional changes. DOE, NASA, and the National Center for Atmospheric Research are jointly developing a next-generation Community Climate System Model (CCSM), which will incorporate a higher degree of physical consistency than current models and enable longer and more detailed simulations. Development and testing of CCSM components are under way. The Parallel Climate Model (PCM) has also been improved with better sea ice and ocean components, a river transport component, and a higher resolution atmosphere component that has better definition of the continent-ocean boundaries and an improved treatment of mountains.

Steady progress is being made in determining the level of natural climate variability and distinguishing it from anthropogenic changes. Contributing to this effort is a large ensemble set of PCM simulations, showing the global climate changes due to increased greenhouse gases and changes in sulfate aerosols, for the years 1870–2100. And basic research is under way to understand the potential effectiveness and the environmental impacts of carbon sequestration in the oceans.

One of the most significant findings of the past year challenges widespread speculation that particulate pollution may offset the effects of greenhouse gases by increasing the cloud cover and reflecting solar energy back into space. Field observations of the dark haze that covered the Indian Ocean in February and March 1999, and subsequent computational analysis of the data, showed that the haze absorbed solar heat, and an atmospheric temperature increase of only 1° per day at noon was enough to significantly reduce the cloud cover (see page 52).

In the medical and life sciences, much of the research done using NERSC computers is focused on developing the computational methods needed to interpret and make use of genomic data — for example, identifying and classifying protein folds in complete genomes, predicting protein structures on the basis of genomic data, and elucidating the mechanisms of protein folding through simulations. Another major project explores the details of the chemical mechanisms employed by enzymes to serve as catalysts of biochemical reactions — events that are too fast to be measured by experiment.

Fusion Energy Sciences

The Office of Fusion Energy Sciences has historically been in the forefront of promoting computational science. Recent progress in fusion research has been accelerated by a strong coupling between theory,

Reprinted by permission from Nature 404:27 April (2000), copyright 2000 Macmillan Magazines Ltd.

Boomerang Data, Analyzed at NERSC, Reveals Flat Universe

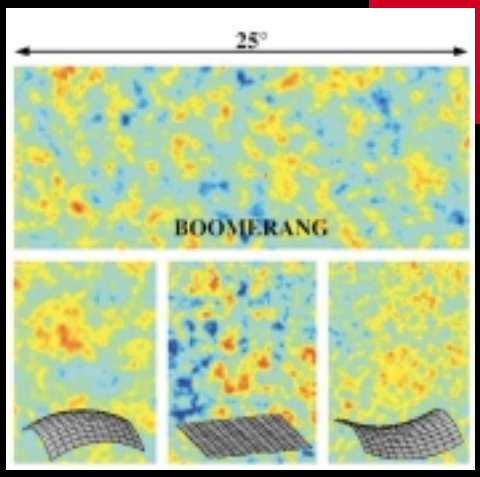
April 26, 2000, was a historic day in scientific cosmology. In an online-video news conference and a cover story in Nature, the international BOOMERANG consortium, led by Andrew Lange of the California Institute of Technology and Paolo de Bernardis of Università di Roma La Sapienza, announced results of the most detailed measurement yet made of the cosmic microwave background radiation (CMB).

BOOMERANG, which stands for "balloon observations of millimetric extragalactic radiation and geophysics," revealed that the curvature of the Universe is not positive or negative but flat. Much of the data analysis was performed at NERSC by astrophysicist Julian Borrill, a BOOMERANG team member and NERSC staff scientist.

During BOOMERANG's 10-day flight around the South Pole in January 1999, it made close to one billion measurements of tiny variations in the temperature of the CMB across a wide swath of the sky. From this dataset, the research team was able to make the most detailed map of the CMB's temperature fluctuations ever seen.

From a map of these temperature fluctuations, the researchers derived a "power spectrum," a curve that registers the strength of the fluctuations on different angular scales, and which contains information on such characteristics of the Universe as its geometry and how much matter and energy it contains. Julian calculated the power spectrum using 50,000 hours of processor time on NERSC's 696-node Cray T3E, employing software he developed called MADCAP (Microwave Anisotropy Dataset Computational Analysis Package).

Combined with other cosmological measurements, such as studies of distant supernovae by the Supernova Cosmology Project headquartered at Berkeley Lab, the BOOMERANG results support the emerging "concordance model" of a flat Universe filled with dark energy, which may correspond to the cosmological constant first proposed by Albert Einstein in 1917. For more details, see page 67.



By observing the characteristic size of hot and cold spots in the BOOMERANG images, the geometry of space can be determined. Cosmological simulations predict that if our Universe has a flat geometry (in which standard high school geometry applies), then the BOOMERANG images will be dominated by hot and cold spots of around 1° in size (bottom center). If, on the other hand, the geometry of space is curved, then the bending of light by this curvature of space will distort the images. If the universe is closed, so that parallel lines converge, then the images will be magnified by this curvature, and structures will appear larger than 1° on the sky (bottom left). Conversely, if the universe is open and parallel lines diverge, then structures in the images will appear smaller (bottom right). Comparison with the BOOMERANG image (top) indicates that space is very nearly flat.

Julian Borrill

computation, and experiments. Three-dimensional modeling contributes to the developing understanding of plasma physics, improves the analysis of experimental results, and suggests new ways to improve magnetic and heavy ion fusion reactor designs.

The Numerical Tokamak Turbulence Project this year completed the lengthy process of benchmarking turbulence simulations from several different codes. They presented the first toroidal electromagnetic simulations of tokamak microturbulence, and also discovered that electron temperature gradient turbulence (ETG modes) can, under some conditions, cause transport comparable to that resulting from ion temperature gradient (ITG) modes. Simulations that quantified the transition from electrostatic to electromagnetic turbulence with increasing β called into question the validity of the electrostatic approximation commonly employed in turbulent transport studies. The new simulations found that microturbulence takes on an electromagnetic character even at low values of β, and that significant electromagnetic effects on turbulent transport occur.

In the work of other research groups, the importance of nonlinearly generated zonal flow for the reduction of ion thermal transport was demonstrated, as well as the role played by ion-ion collisions in the bursting behavior observed in tokamak experiments. Studies of two-stream instabilities in space-charge-dominated beams in accelerators helped to explain the beam loss observed in various machines. A new working model was developed of edge localized modes (ELMs), which have been observed but poorly understood for two decades; the model successfully described ELM behavior in the DIII-D tokamak.

Gyrokinetic growth rate calculations analyzing the drift-wave stability of a variety of tokamak plasmas found that discharges with neon injection had improved energy confinement due to the suppression of ITG-mode turbulence.

High Energy and Nuclear Physics

The DOE Office of High Energy and Nuclear Physics sponsors important theoretical studies, computational simulation and analysis of experimental data, and simulations that are helping design the next generation of experimental facilities.

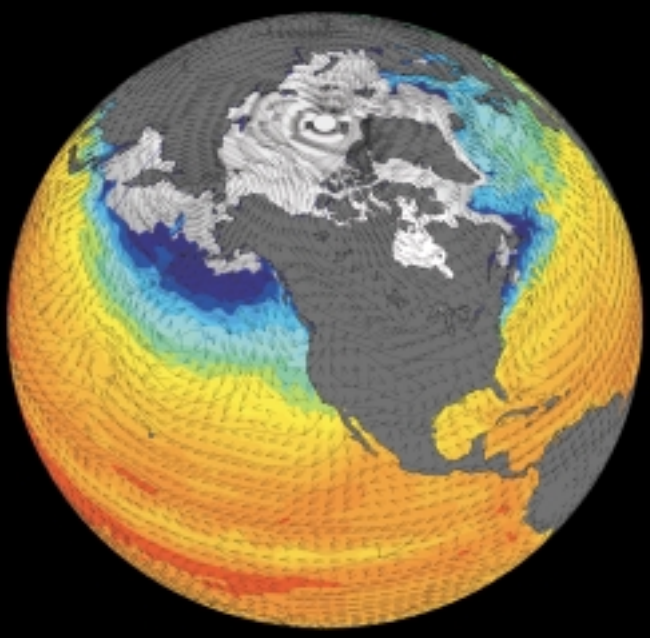
NERSC's leadership in astrophysics data analysis yielded headline-making results for the second time. Two years ago, analysis of supernova data led to the conclusion that the Universe will continue expanding forever. This year, analysis of the BOOMERANG cosmic microwave background data supported the earlier finding as well as the new conclusion that the geometry of the Universe is flat (see pages 6 and 67).

On the theoretical side, NERSC provides computational resources for several large research efforts in lattice quantum chromodynamics (QCD). Lattice QCD provides the most promising approach to understanding the behavior of quarks and gluons, the building blocks of strongly interacting particles. QCD studies are relevant to the physics of the early Universe and are crucial to interpreting the results of experimental attempts to produce a quark-gluon plasma. QCD calculations test the Standard Model and may provide clues to a new physics.

Nuclear physics researchers are working to explain the properties and reactions of nuclei in terms of interacting nucleons (protons and neutrons). A team from Argonne National Laboratory has achieved calculations of six- through ten-nucleon systems that use realistic interactions and that are accurate to 1% for the binding energies. The resulting wave functions can be used to compute properties measured at electron and hadron scattering facilities and to compute astrophysical reaction rates, many of which cannot be measured in the laboratory.

The Advanced Computing for 21st Century Accelerator Science and Technology project (the former Computational Accelerator Physics Grand Challenge) is developing a comprehensive terascale accelerator simulation environment for the U.S. particle accelerator community. The design and construction of the next generation of accelerators will involve greater complexity than ever before, and will require unprecedented precision in accelerator design and beam control. For all these accelerator systems, terascale simulation will play a key role by facilitating important design decisions, increasing safety and reliability, optimizing performance, and helping to ensure project completion within budget and on schedule. The project will add new codes and modules to an existing set that is being used for simulations at accelerator facilities around the world.

Systems and Services



High-resolution global climate models require high-capability parallel systems such as NERSC's IBM SP and Cray T3E. See page 53 for details.

A hallmark of NERSC since its founding in 1974 has been the expertise and competence of the employees staffing the facility and the high quality of services they deliver to our users. Year after year, the NERSC staff delivers critical computing resources, applications, and information to enable users to make the most of their allocations on our high performance systems. Each year the staff introduces innovative enhancements and improvements to the systems and services as well.

While NERSC employees continued to anticipate and meet users' needs in 2000, they also undertook a number of special projects aimed at maintaining NERSC's preeminence among scientific computing facilities — most notably, the installation and testing of the Phase 1 IBM SP system, and the move of our other systems to Berkeley Lab's new Oakland Scientific Facility, where there is room for the Phase 2 SP as well as future expansion. In this section we highlight some examples of both ongoing activities and special projects. For a more comprehensive view, "How Are We Doing: A Self-Assessment of the Quality of Services and Systems at NERSC" is available at http://www.nersc.gov/aboutnersc/pubs/hawd99.pdf.

Off to a Flying Start in Y2K

In 1999, the mass media had portrayed the year 2000 as a cyberspecter threatening to undermine nearly everything remotely connected to a computer. Thanks to careful and thorough preparations at NERSC, the calendar change came and went uneventfully, with no loss of service and no impact on users. Unlike other facilities that shut down over the holiday and brought in extra staff to deal with potential problems, NERSC maintained normal operations with only one extra person on duty for security on New Year's Eve.

This business-as-usual stance was possible because most of NERSC's Y2K testing was actually completed in late 1998 and early 1999. In fact, NERSC was one of the first organizations within DOE to demonstrate Y2K compliance, and was the only site to fully test IBM supercomputers, HPSS, or a Cray T3E for Y2K compliance. Hardware, operating systems, layered software, and scientific applications were all subjected to rigorous testing.

While most of NERSC's Y2K preparations were completed in 1999, R. K. Owen, Brent Draney, John McCarthy, and Greg Butler put on the finishing touches. Greg developed a new methodology for testing the IBM SP system and had it in place within 10 days of the system's installation in April. R. K. and John prepared a backup system for user account information in NERSC's Central User Bank (CUB), and Brent upgraded CUB with a number of Y2K patches. As NERSC security analyst, Brent was on the job on New Year's Eve; he quickly fixed a handful of minor problems that cropped up with the date change and also averted an attempted hacker attack.

The Y2K compliance team, led by Jim Craw, included Harsh Anand, Greg Butler, Tina Butler, Jonathon Carter, Thomas Davis, Tina Declerck, Jed Donnelley, Keith Fitzgerald, Susan Green, Frank Hale, William Harris, John Hules, Wayne Hurlbert, Nancy Johnston, Cheri Lawrence, Ken Okikawa, Bill Saphir, Jackie Scoggins, Craig Tull, David Turner, Mike Welcome, and Tammy Welcome.

And thanks to NERSC's ongoing security program, there have been no serious security incidents to date.

NERSC Delivers New Resources with Traditional Efficiency

With Phase 1 of the IBM SP system installed (see page 10), NERSC was able to double MPP allocations from 4.7 million processor hours in FY 1999 to 9.3 million in FY 2000. FY 2001 MPP allocations are up another 30%, even without factoring in the Phase 2 SP system, which will have 2,528 processors and a peak speed of 3.8 teraflops. PVP allocations increased about 50% from FY 1999 to FY 2000, while storage resource units (SRUs) increased about 40%.

Not content merely to increase our resources, we operated them with our traditional efficiency and convenience

Phase 1 of NERSC's IBM RS/6000 SP system, with 512 processors and a peak performance of 410 gigaflop/s, will be replaced by the 3 teraflop/s Phase 2 system early in 2001.

IBM SP Launched Ahead of Schedule with Million-Hour Bonus for Users

When Phase 1 of NERSC's new IBM RS/6000 SP computer system went into full service on April 5, 2000, the launch was 25 days ahead of schedule. As a result of getting the system up and running sooner than expected, NERSC was able to provide users with one million more MPP hours than anticipated for the FY 1999 allocation period.

This achievement, which capped a very complex procurement process, was the result of a five-month effort by a team of 30 employees (listed below) led by High Performance Computing Department head Bill Kramer. They installed, configured, and customized the SP system for the NERSC production environment and users, worked with early users to test and refine the system, and created training programs and documentation as well as demonstration applications. Although the team had little prior experience with SP systems and there were no permanent IBM staff on site yet, they learned quickly on their own and created innovative solutions to the problems they encountered.

In keeping with this pioneering spirit, NERSC was the first site to commit all its files to IBM's new General Parallel File System (GPFS). Our SP team helped IBM find and resolve issues that will benefit all of our users and other sites as well. NERSC's success with GPFS has already prompted other sites to adopt it. But what's most important is that NERSC users are keeping the SP constantly busy and are happy with the system's performance.

Majdi Baddourah
David Bailey
Elizabeth Bautista
Greg Butler
Andrew Canning
Nick Cardo
Jonathan Carter
Jim Craw
Tom DeBoni
Tina DeClerck
Chris Ding
Brent Draney
Therese Enright
Richard Gerber
William Harris
Terri Kaltz
Sherri Li
Gary Mack
Osni Marques
Nancy Meyer
Esmond Ng
R. K. Owen
Harsh Anand Passi
Bill Saphir
Jackie Scoggins
Martin Stouffer
Francesca Verdier
Mike Welcome
Tammy Welcome
Adrian Wong

to users. The Cray T3E utilization rate was over 90% for most of the year (Figure 4), and IBM SP utilization topped 80% continuously from the time the system was placed in production through the end of the year — an impressive achievement with a new system (Figure 5). In addition, to help users of the Cray SV1 computers test and debug their programs, we opened one of the machines for interactive use.

Move to Oakland Goes Like Clockwork

After almost two years of planning, months of construction, and an intense week of moving and installation, NERSC's Cray supercomputers, HPSS, PDSF cluster, and auxiliary systems were up and running in Berkeley Lab's new Oakland Scientific Facility on November 6 (see photos below and on the following pages). The moved systems were down for less than a week. NERSC's IBM RS/6000 SP Phase 1 computer remained in service during the move and will stay in Berkeley until the Phase 2 system is installed in Oakland. Staff from NERSC's Computer Operations and Network Support Group have relocated to the Oakland facility, and other Berkeley Lab employees will move into the facility early in 2001.

The move was the culmination of a process that began in January 1999, when NERSC, looking for space to accommodate larger computer systems, launched an extensive site selection process. A request for proposals drew eight qualified responses, from which the Oakland site emerged as the favorite. After a contract was signed

Figure 4. Cray T3E utilization for FY 2000.

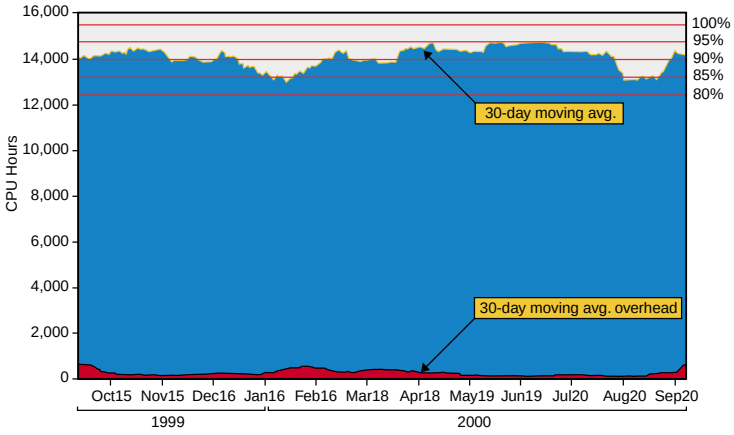
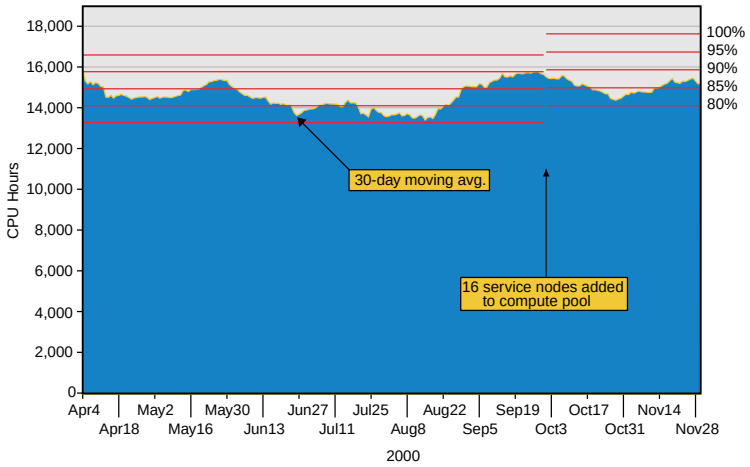


Figure 5. IBM SP utilization from April to November 2000.



in August 1999, the former bank building was stripped down to its support structure, seismically reinforced, and completely rebuilt to meet or exceed all current codes and standards. Additional improvements in the electrical supply capability and a high-volume ventilation system

Construction of Berkeley Lab's new Oakland Scientific Facility provided the floor space and electrical capacity for future expansion of NERSC's computing systems.

Through hard work and ingenuity, Steve Lowe, Wayne Hurlbert, and James Lee overcame hardware problems and a tight schedule to complete the transfer of archival storage data to an upgraded system.

were among the many special accommodations required for high performance computers. The site includes room for future expansion.

Howard Walter

Howard Walter, head of the Future Infrastructure, Networking and Security Group, managed the project for NERSC throughout the process of specification development, bid solicitation and evaluation, contract negotiation, design, permitting, construction, and relocation. A ceremony to dedicate the facility is being planned for early spring 2001.

HPSS Increases Capacity and Performance

NERSC's High Performance Storage System grew significantly in both capacity and performance during the past year thanks to the work of Nancy Meyer and the Mass Storage Group. From January to December, archival storage increased from 660 terabytes (TB) to 880 TB. The amount of data being stored at the end of the year was 145 TB. The online disk data cache grew from 1.8 TB at the beginning of the calendar year to 6 TB at the end. And the default disk speed was increased from 6 megabits per second (Mb/sec) to 32 Mb/sec. The storage environment moved from individual MicroChannel machines to IBM SP nodes, doubling the number of processors available to the storage machines and doubling the speed of each machine.

50 TB Data Transfer Overcomes Hurdles

Moving terabytes of data, even under the best conditions, can be a time-consuming chore. But when the hardware doesn't cooperate and time pressure is factored in, it's even tougher. In March 2000, Wayne Hurlbert and James Lee of the Mass Storage Group and Steve Lowe of the Computer Operations and Network Support Group were given the task of moving 50 terabytes (TB) of archival data from NERSC's IBM storage libraries to the StorageTek silos. The transfer was driven in part by a need to trade in the IBM systems by a set date and add an additional 220 TB capability to the storage system. As the data transfer began, hardware failed, causing the team to transfer the data using just 6 of 14 tape drives. Then the internal high-speed network slowed down, further impeding timely transfer of data. The team worked night and day to fix the hardware problems and came up with a workable plan to move the data. Despite the technical difficulties, they managed to move the 50 TB in four weeks at a sustained rate of 21 Mb/sec, completing the job on time.

PDSF Expands and Upgrades

Ongoing expansion and upgrades of the PDSF (Parallel Distributed Systems Facility) continued this year. The PDSF is used by several high energy and nuclear physics experiments for data analysis and simulations. Eighty-nine dual-CPU compute nodes were added for a total of 151 nodes or 277 processors, and the disk vault was expanded by 4.0 TB to a total of 7.5 TB. Tom Davis wrote a channel bonding addition to the Linux kernel that made network connectivity to the disk vault more reliable by using the two physical Ethernet connections as a single logical connection. With the move to Oakland, five disk servers were given Gigabit Ethernet fiber connections, so the PDSF now has a gigabit connection to HPSS.

New Cluster Computing Team Established

NERSC Deputy Director Bill Kramer announced in June that NERSC was establishing a team of staff from multiple groups to coordinate all NERSC Division cluster computing activities (research, development, advanced prototypes, pre-production, production, and user support). This team will assure the most effective implementations of division resources related to cluster computing.

The NERSC Cluster Computing Team is led by Tammy Welcome and is primarily composed of staff (ranging from part time to almost full time) from the Advanced Systems, Future Technologies, Computational Systems, and User Services groups. The goals of the team are:

- Continue providing high-quality support and service for production clusters such as the PDSF.
- Take maximum advantage of the NERSC production environment to improve cluster systems while maintaining appropriate levels of service.
- Investigate the feasibility and effectiveness of cluster

The initiative of Russell Huie and Elizabeth Bautista helped fellow members of the Computer Operations and Network Support Group to continue providing outstanding service to NERSC and ESnet users.

computing as a full-production, highly parallel computing platform.

- Exercise leadership in the cluster arena within DOE and the national HPC community, and collaborate with other groups.

Operations Team Takes on New Technologies

Keeping up with technological changes is a challenge in itself, but finding a way to make the information readily available can be even tougher. When Elizabeth Bautista of NERSC's Computer Operations and Network Support Group took on the task of updating documentation for the center's Operations staff, she not only had to work her way through pages of outdated printed information, she also had to scale a steep learning curve to find the best way to reorganize and update the information on the Web, so Operations staff members could more easily get the information they needed. Elizabeth not only acquired the needed expertise and completed the job on time, but she also motivated other group members to improve their troubleshooting skills as part of the process.

Russell Huie also put himself in the learning fast lane by stepping forward as the main point of contact when ESnet announced major changes in its Video Conferencing System (VCS). After taking on the responsibility, Rusty gathered all the information needed to operate, troubleshoot, and continue providing the same high level of videoconferencing services to the ESnet user community using the new Digital Collaboration Services (DCS). Once he was up to speed on the new system, Rusty provided training to other members of the Operations staff and also ensured that thorough DCS documentation was available to users.

NIM Provides Better Account Management

At the beginning of FY2001, after a year-long development project, NERSC replaced the CUB account management system with the NERSC Information Management system (NIM), a Web-based application for managing accounts. The primary reason for creating NIM was that CUB was designed for Cray vector computers and could no longer be extended to support the current high performance systems like the IBM SP. CUB also could not be ported to other architectures such as cluster or HPSS systems. Moving to a Web-based application allows NIM to be used on any platform. However, the move to a Web interface necessitated new authentication and security components, which are incorporated in NIM.

NIM makes it easier to transfer resources, either from reserve accounts to repositories, or from repo to repo. Another advantage is that NIM is based on open-source PHP (a server-side, cross-platform, HTML embedded scripting language) and uses an Oracle database, making it easier to support than the in-house produced CUB. When Version 2 is completed in the summer of 2001, NIM will provide users with richer data covering all NERSC platforms. NIM was created by members of the User Services, Computational Systems, Mass Storage, and Computer Operations and Network Support groups, with Howard Walter as project manager.

Consultation and Training Promote User Productivity

NERSC's commitment to making scientific computing more productive extends to each researcher who uses our resources. The User Services Group is the user community's primary point of contact with NERSC. This group is responsible for problem management and consulting; help with user code optimization and debugging; documentation; online, remote, and classroom training; and third-party applications and library support. In

addition, the Computer Operations and Networking Support Group is available seven days a week, 24 hours a day to help troubleshoot problems.

During the past year, User Services presented 32 training sessions, which included a three-day onsite and video-conference workshop on using the IBM SP system; two lectures on high performance linear algebra using the IBM Power 3 architecture; two days of training sessions for new and intermediate users in conjunction with the NERSC Users Group meeting at Oak Ridge National Laboratory; and six teleconference sessions with online materials. A wide range of documentation and training materials are also available at http://hpcf.nersc.gov/, with over 3,800 new training files added in just the past year.

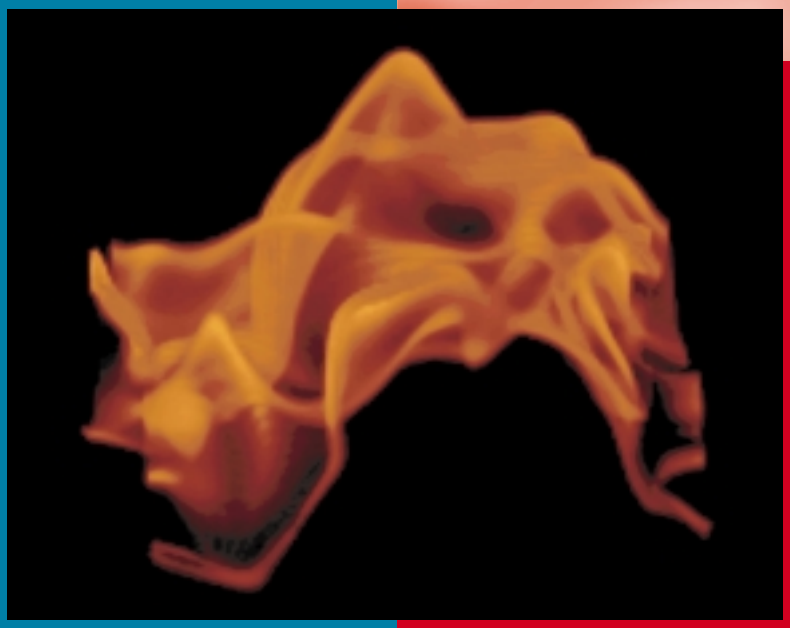
Survey Responses Lead to Improved Services, Information

Each year for the past three years, NERSC's User Services Group has asked users to complete a survey rating the center's systems and services and asking for suggestions on how NERSC can better serve its users. With each survey, the number of users participating has increased. The level of user satisfaction has also risen each year, in part because the NERSC staff implements new procedures based on comments from users. For example, based in large part on the input received in the Fiscal Year 1999 survey, NERSC made the following improvements:

- To accommodate users running large jobs on the Cray T3E, NERSC created a long-running queue (up to a maximum of 12 hours) for jobs using up to 256 PEs.
- To help users of NERSC's Cray SV1 computers test and debug their programs, the center opened one of the machines for interactive use.
- To keep users better informed of NERSC announcements and changes, User Services created new email lists, continuing changes made as a result of the 1998 survey.
- To make it easier for users to find information on the Web about running batch and interactive jobs, a new Web page with near real-time information was created at http://hpcf.nersc.gov/running_jobs/.
- To help users decipher what's being said, a NERSC Glossary and Acronym List was created and posted at http://hpcf.nersc.gov/help/glossary.html.
- To help users quickly find out whether any machine at NERSC is up or down, an automated window presenting machine status was added to the bottom of the HPCF home page, http://hpcf.nersc.gov/.

Results of the 2000 survey are available at http://hpcf.nersc.gov/about/survey/2000/.

Research and Development



Adaptive mesh refinement simulation of developing flame surface. See page 78 for details.

NERSC's concept of a scientific computing center involves much more than providing access to high performance computers and data storage — it includes providing the intellectual leadership to make computational science more productive. The NERSC Program, which provides computer access and intellectual services for the NERSC user community, is embedded in the NERSC Division at Berkeley Lab, which includes a large number of independently funded research and development efforts.

One of the key visions of the 1995 proposal that resulted in NERSC's relocation to Berkeley Lab was the mutually beneficial connection of the DOE flagship computing facility to other DOE-funded research activities in applied mathematics, computer science, and computational science — the three elements involved in developing scientific modeling and simulation codes (Figure 6). On one hand, research and development efforts in the NERSC Division directly improve the intellectual tools that make our high performance systems useful. For example, the NERSC Program directly benefited from cluster computing and data management research carried out elsewhere in the Division. On the other hand, the requirements of high-end users often encourage researchers to explore new directions. For example, the Visapult framework described below would not have been developed had not the visualization and distributed computing researchers been directly engaged in addressing the requirements of combustion applications users of the NERSC Facility.

The results of the R&D efforts described below prove that combining a computing facility with research and development in one organization has demonstrable benefits both for the NERSC user community and for the advancement of DOE research programs in applied mathematics, computer science, and computational science.

Applied Mathematics

Applied mathematics research at NERSC involves development of software for high-precision arithmetic, linear algebra algorithms, and adaptive mesh refinement, with applications ranging from quantum mechanics to fluid dynamics to information retrieval. The highlight of this year's R&D was the release of Berkeley Lab AMR at SC2000.

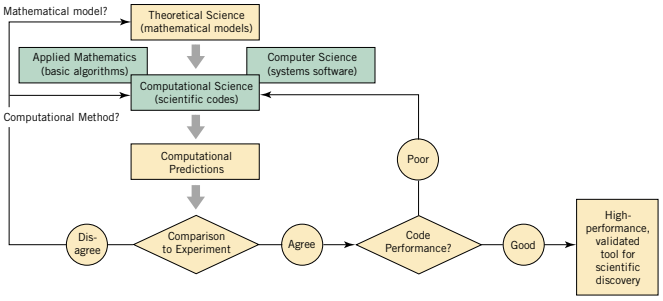


Figure 6. Work flow for the development of scientific modeling and simulation codes (adapted from "Scientific Discovery through Advanced Computing," DOE Office of Science, March 24, 2000).

BERKELEY LAB AMR

Berkeley Lab AMR, a comprehensive library of adaptive mesh refinement software and documentation, is the culmination of more than 15 years of research by members of both the Center for Computational Sciences and Engineering and the Advanced Numerical Algorithms Group. Berkeley Lab AMR is unique among many AMR codes because of its adaptability to a wide range of applications. Scalable parallelism and an object-oriented approach have been built into the design from the very beginning to ensure flexibility and high performance across multiple platforms.

AMR serves as a "numerical microscope," allowing researchers to zoom in on the specific regions of a problem that are most important to its solution. Rather than requiring that the whole calculation have the same spatial resolution, AMR allows different resolution in different regions of the problem. Areas of interest are covered with a finer mesh than the surrounding regions; for time-dependent problems, the finer meshes are also advanced with a smaller time step. Not having to perform the entire calculation at the finest resolution allows scientists to make the most of available computer resources, so that they can then solve bigger, harder problems.

One of the most challenging problems in computational science to which AMR is being applied is the numerical modeling of combustion. Calculations of combustion processes often include a well-defined flame front; focusing the computing power on the flame, where hundreds or thousands of chemical reactions may be taking place,

Amazing Algorithm Pulls Digits Out of π

When the January/February 2000 issue of Computing in Science and Engineering magazine named their top 10 "Algorithms of the Century," the list included the integer-relation algorithm PSLQ, discovered by mathematician and sculptor Helaman Ferguson of Maryland's Center for Computing Sciences, and implemented in practical computer software by David Bailey, NERSC's chief technologist.

David Bailey

As a tool of experimental mathematics, PSLQ's purpose is to discover new mathematical relations among numbers. In a short time it has found polylogarithmic formulas in algebraic number theory, identified a class of multiple-sum constants, uncovered relations in the renormalization procedures of quantum field theory symbolized by Feynman diagrams, and – most surprisingly – found a formula for calculating any digit of π without calculating the digits preceding it.

Before PSLQ, mathematicians had not thought that such a digit-extraction algorithm for π was possible. Using the remarkably simple formula, even a personal computer can calculate π's millionth binary digit in about 60 seconds. Most applications of PSLQ, however, require much more computing power and must employ much greater numerical precision than the standard 16-digit, 64-bit, floating-point arithmetic available on most computers. That's why Bailey has developed software that translates ordinary C or Fortran programs into programs capable of arbitrary precision – calculations accurate to tens, hundreds, or even many thousands of digits.

Some of PSLQ's results have profound implications. The π formula raises questions about the long-held but never proved assumption that π's digits are random. The Feynman-diagram results hint at unsuspected relationships among formulas associated with fundamental particles. These discoveries suggest that experimental mathematics using computers will become increasingly important in this new century.

results in large savings in computing time and memory. As the flame develops and moves through the domain, the finer meshes automatically move with it, allowing researchers to achieve unprecedented temporal and spatial resolution of the internal flame structure. (For an example, see page 78.)

Researchers interested in obtaining a copy of the Berkeley Lab AMR CD can send requests to AMR@lbl.gov. More information about Berkeley Lab AMR is available at http://seesar.lbl.gov/AMR/.

Computer Science

NERSC's R&D efforts in computer science span the entire cycle of scientific data analysis, including data acquisition, secure transmission, storage and retrieval, and visualization. Several projects involve development of components for the DOE Science Grid. Computational grids are persistent environments that enable software applications to integrate instruments, data, computational and information resources that are managed by a number of organizations in widespread locations. Grids give scientists a uniform interface to computational resources similar to the way that a web browser provides a seamless interface to the Internet. With grid technology, the researcher does not need to be concerned with multiple protocols or different commands at individual sites.

In addition to software for massively parallel systems, clusters, and grids, we discuss below NERSC's ongoing involvement in benchmark development and system performance analysis.

SHARING DATA IN PARALLEL: NETCDF

In the latest release of the netCDF software library from Unidata, one of the major improvements is the parallel support developed by NERSC staff for the Cray T3E. A significant limitation of previous netCDF releases was that the software could not be used for collective parallel access to a single file. This limitation made netCDF inefficient and inconvenient for many large-scale simulations, such as high-resolution climate modeling. Since the T3E is currently one of

NERSC's new Distributed Systems Department, a few of whose members are shown here, works on a wide variety of R&D projects for computational grids. Chuck McParland not only helped design the sensors for a neutrino astronomy experiment named AMANDA (Antarctic Muon and Neutrino Detector Array), he traveled to the South Pole to help install the sensors. Dan Gunter is Berkeley Lab's representative in the Grid Performance working group of the Grid Forum, the new standards organization for emerging Grid technologies. Vern Paxson is the creator of BRO, a network security monitoring system. Marcia Perry has developed software for remote camera control and remote videoconferencing control. And Srilekha Mudumbai, the lead developer on the Akenti project, collaborates with researchers from government, industry, and academia on secure authorization and access control systems.

the most popular high-end computing platforms, the new portability enhancements make it possible for a wide range of research programs to access scientific data and share it with collaborators in the netCDF format.

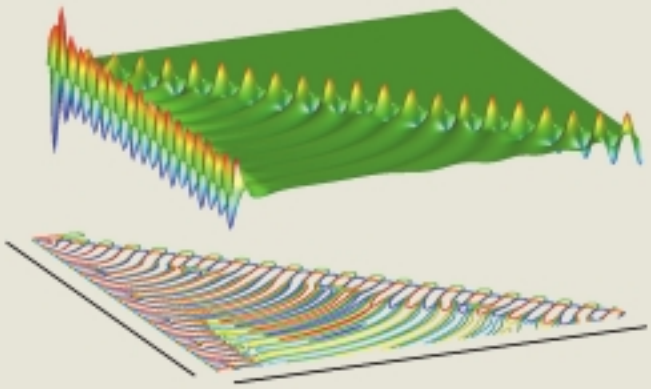
netCDF (network Common Data Form) is a library of input/output software for storing and retrieving scientific data in self-describing, platform-independent files. It was developed primarily for the climate research community by the National Science Foundation-funded Unidata Program Center in Boulder, Colorado; and like many cooperative software efforts, it includes enhancements developed by users (see http://www.unidata.ucar.edu/packages/netcdf/).

NERSC's enhancement effort was initiated to meet the critical needs of climate modeling applications. One of the first applications of parallel netCDF was to speed up the I/O in the Modular Ocean Model (MOM). A similar effort is being planned to port netCDF to the IBM SP platform.

COMMUNICATION FOR CLUSTERS: M-VIA AND MVICH

M-VIA and MVICH are VIA-based software for low-latency, high-bandwidth, inter-process communication. Virtual Interface Architecture (VIA) is an industry standard high performance communication interface for system area networks (SANs). VIA provides protected user-level zero-copy data transfers, enabling low latency and high bandwidth. The communication model includes both cooperative communication (send/recv) and remote memory access (get/put).

M-VIA is a modular implementation of the VIA standard for Linux. It provides a software framework



The ACTS (Advanced Computational Testing and Simulation) Toolkit is a set of computational science tools that were mostly developed at DOE national laboratories and universities. ACTS is an umbrella project of the DOE Office of Advanced Scientific Computing Research which has brought the tools together and is funding developers to provide tool interoperability. The tools in the ACTS Toolkit provide solutions for numerical problems, scientific data representations, data manipulation, visualization, program execution, and distributed computing.

SuperLU, one of the ACTS numerical tools, is a general-purpose library for the direct solution of large, sparse, non-symmetric systems of linear equations on high performance machines. Serial, shared memory, and distributed memory implementations are available. SuperLU was used in a breakthrough quantum mechanical computation done on the NERSC Cray T3E and featured on the cover of the December 24, 1999 issue of Science (see page 82).

ACTS Toolkit Provides Solutions to Common Computational Problems

NERSC provides a centralized source of information about the ACTS Toolkit on the Web (http://acts.nersc.gov), and other services, including:

- outreach to researchers in new and expanding scientific computing areas
- educational programs, including workshops and lectures
- evaluation of current tools for interoperability, robustness, functionality and ease of use
- ongoing support for users in the scientific community.

In September, NERSC presented a three-day workshop on the ACTS Toolkit for graduate students and postdoctoral fellows. Toolkit tutorials were also offered at SC2000 in November.

Here are some of the current tools in the ACTS Toolkit:

Numerical tools
- Aztec
- Hypre
- PETSc
- PVODE
- ScaLAPACK
- SuperLU

Tools facilitating algorithm development
- Global Arrays
- Overture
- PADRE

Support tools
- ATLAS & PHiPAC
- CUMULVS
- Globus
- Nexus
- PAWS
- PETE
- SILOON
- TAU

that eases the development of drivers for new VIA-aware hardware as well as support for legacy network devices. MVICH is an MPICH-based implementation of MPI for VIA. It provides receive-side buffering for short messages and high performance zero-copy RDMA (remote direct memory access) transfers for large messages.

M-VIA and MVICH are the first components of Berkeley Lab Distribution (BLD), a software distribution developed by the Future Technologies Group that will make it easier for scientists to turn a collection of PCs into a usable cluster (see http://www.nersc.gov/research/FTG/bld/index.html). It will provide the key tools for configuring, managing, and running jobs on a cluster, and will support both task-farm and parallel clusters.

GRID SECURITY: AKENTI AND BRO

Akenti, developed by the Distributed Security Research Group under the leadership of Mary Thompson, is an authorization system designed to address the issues raised in permitting access to distributed resources that are controlled by multiple remote stakeholders. Examples of such resources include computing and data storage systems and on-line instruments such as electron microscopes or medical diagnostic systems that have been enabled for remote operation. Access to resources is controlled by a resource gateway, which is typically a secure server such as a secure Web browser, CORBA ORB, Grid gatekeeper, or some distributed application server. These gateways are modified to invoke Akenti to make the authorization, which the gateway then enforces.

Mary Thompson

Akenti enables stakeholders to securely create and distribute policy statements authorizing access to the resources for which they have responsibility. Akenti makes access control decisions based on a set of digitally signed documents that represent these authorization instructions. Public-key infrastructure and secure message protocols provide confidentiality, message integrity, and user authentication, during and after the access decision process. Details and software are available at http://www.itg.lbl.gov/Akenti.

BRO is a standalone system for network security monitoring developed by Vern Paxson of the Networking Group. Named after George Orwell's ever-watching Big Brother, BRO is capable of detecting and shutting down Internet attackers in real time.

The BRO system is designed in layers. The first layer is a general packet filter, which decides which data packets to examine. The second layer is an "event engine," which takes the first-level packets and pieces them together into "events" reflecting different types of activity, such as the beginning of a connection, a successful login, a possible backdoor, or an FTP command request. Next comes the policy layer, which interprets scripts, written in a specialized language, that define how to respond to different events. Should the policy layer detect information amounting to an attempted security breach, the system notifies computer security people in real time. It also can terminate running connections and signal a site's border router to drop traffic coming from an attacker. Finally, it archives summaries of the network traffic into and out of the site in a permanent record.

BRO was used to monitor SCinet at SC2000 and has been continuously monitoring network traffic at Berkeley Lab since April 1996. In that time, it has detected a few hundred formal security incidents, some of which have resulted in law enforcement action. Together with proactive scanning and strategic firewalls, BRO's "reactive firewall" helps provide protection against increasingly sophisticated security threats (see http://www.aciri.org/vern/bro-info.html).

DISTRIBUTED VISUALIZATION OF TERASCALE DATASETS: VISAPULT

Visapult is a prototype application and framework for performing remote and distributed visualization of scientific data. Developed by Wes Bethel of the Visualization Group, Visapult approaches the technical challenges of terascale visualization with a unique architecture that employs high speed WANs and network data

Wes Bethel

Jason Lee

Brian Tierney

Dan Gunter

Arie Shoshani

Alex Sim

Bill Kramer

Grid Applications Win SC2000 Competition

At the SC2000 conference in Dallas, NERSC staff participated in two of the three winning teams in a new competition, the Network Challenge for Bandwidth-Intensive Applications. The competition was initiated to demonstrate high-bandwidth applications and push the limits of network capacity.

Wes Bethel, Jason Lee, Brian Tierney, and Dan Gunter won the "Fastest and Fattest" category for overall best performance for their demonstration of Visapult (see page 21). They recorded a peak performance level of 1.48 Gb/sec over a five-second period. Overall, the Visapult team transferred 262 gigabits of data in 60 minutes from Berkeley Lab to the SC2000 show floor in Dallas, with an average throughput of 582 Mb/sec.

"A Data Management Infrastructure for Climate Modeling Research" took top honors as the "Hottest Infrastructure" application. This application was run by a team from the University of Southern California/Information Sciences Institute, Argonne and Lawrence Livermore national laboratories, and NERSC's Arie Shoshani and Alex Sim. They achieved a peak performance level of 1.03 Gb/sec.

These accomplishments are the culmination of many years of research leadership in high performance networking and data management by Berkeley Lab computer scientists who now work in the NERSC Division. This research included the development of the Distributed-Parallel Storage System (DPSS), which was used in the Visapult demonstration, and STACS (see page 24), which is part of the climate modeling infrastructure.

The SC2000 Network Challenge itself was made possible by SCinet, the temporary but massive communications network which was assembled and operated for the conference under the leadership of NERSC Deputy Director Bill Kramer. With three OC-48 lines and three OC-12 lines, SCinet offered a combined peak speed of 9.4 Gb/sec – more than 167,000 times faster than a typical residential Internet connection and 200 times as fast as the connections used by many universities. For the first time in the history of the SC conference, SCinet also provided wireless networking capability throughout the Dallas Convention Center.

caches such as DPSS (http://www-didc.lbl.gov/DPSS/) for data staging and transmission. High throughput rates are achieved by parallelizing I/O at each stage in the application, and by pipelining the visualization process. Visapult's peak performance level of 1.48 Gb/sec won the top prize in the SC2000 Network Challenge (see sidebar on page 22).

Visapult consists of two components: a viewer and a back end. The back end is a parallel application that loads in large scientific datasets using domain decomposition, and performs software volume rendering on each subdomain, producing an image. The viewer, also a parallel application, implements Image Based Rendering Assisted Volume Rendering, using the imagery produced by the back end. On the display device, graphics interactivity is effectively decoupled from the latency inherent in network applications. Information and downloads are available at http://www-vis.lbl.gov/projects/visapult/.

A MICROSCOPY CHANNEL FOR THE INTERNET: DEEPVIEW

DeepView is a collaborative problem-solving environment for distributed microscopy and informatics. DeepView software allows researchers to seamlessly participate in experiments at online microscopes, acquire expert opinions, collect and process data, and store this information in their electronic notebook. The testbed includes several unique electron and optical microscopes that are located at Lawrence Berkeley National Laboratory, Oak Ridge National Laboratory, and the University of Illinois, with applications ranging from material science to cell biology.

Bahram Parvin

Developed by NERSC's Imaging and Collaborative Computing Group under the leadership of Bahram Parvin, DeepView uses an extensible object-oriented framework built on a foundation of CORBA enabling services. DeepView's Instrument Services provide a layer of abstraction for controlling any type of microscope; Exchange Services provide a common set of utilities for information management and transaction; and Computational Services provide the analytical capabilities needed for online microscopy and problem solving. Key features of the system include scalability and close integration of data collection with online data analysis, annotation, and storage. Further information and DeepView software are available at http://vision.lbl.gov/Projects/DeepView/.

REAL-TIME GRID MONITORING AND ANALYSIS TOOLS: NETLOGGER AND PIPECHAR

High-performance distributed systems are vulnerable to unexpected performance problems, such as low throughput or high latency. Finding the reasons for these problems is challenging because the nature of the systems tends to multiply the number of possible points of failure. To make the optimum use of distributed systems, users also need to know current and maximum bandwidth, current and minimum latency, bottlenecks, burst frequency, and the extent of congestion. Providing new network services such as Quality of Service, in which network capacity can be assigned on a priority basis, also requires network monitoring and analysis. NERSC's Data Intensive Distributed Computing Group, under the leadership of Brian Tierney, has developed a suite of tools to address these problems. (Information about these tools and downloads are available at http://www-didc.lbl.gov/.)

The NetLogger Toolkit enables the real-time diagnosis of performance problems in complex high-performance distributed systems. NetLogger includes tools for generating precision event logs that can be used to provide detailed end-to-end application and system level monitoring, and tools for visualizing log data to view the state of the distributed system in real time. This approach is novel in that it combines network, host, and application-level monitoring, providing a complete view of the entire system. Over the past few years, NetLogger has proven to be invaluable for diagnosing problems in networks and in distributed systems code. NetLogger monitoring allows users to identify hardware and software problems, and to react dynamically to changes in the system.

Deb Agarwal Named One of "Top 25 Women of the Web"

Deb Agarwal, deputy head of NERSC's Distributed Systems Department, was honored in the May 2000 cover story of Upside magazine as one of the "Top 25 Women of the Web." She received the award for her work to provide reliable multicast communication for the Comprehensive Test Ban Treaty monitoring system. Having written her Ph.D. thesis on reliable multicasting, Deb served as an independent expert at the Vienna headquarters of the treaty organization for three months in 1999, and as a member of the U.S. delegation on subsequent occasions.

Deb's report showed that multicasting technology provides the reliability required by the treaty as well as flexible network configuration. Although politics and turf wars prevented official acceptance of Deb's recommendations, one European country is planning to implement a multicast system that Deb hopes will serve as a prototype for other nations.

Deb Agarwal

Deb leads the Collaboration Technologies Group at NERSC, which researches, develops, and deploys the middleware and technologies needed to allow scientists to collaborate on projects. Their research falls roughly into three categories: (1) applications and middleware supporting real-time collaboration between geographically remote researchers, (2) reliable group communication mechanisms, and (3) middleware enabling network-aware applications.

A new and easy-to-use tool for analyzing and monitoring the network itself was also made available this year. This tool, called pipechar, is a sub-service of the Network Character Service Daemon (NCSD) and has been extracted as an individual tool, paired with netest for identifying problem routers. Pipechar is a simple tool that users can run themselves from their desktop computers to query the network for information on bandwidth, latency, and congestion. Unlike SNMP (or Simple Network Management Protocol), pipechar does not require router access privilege, which is not always feasible.

EFFICIENT DISTRIBUTED STORAGE ACCESS: STACS

The Storage Access Coordination System (STACS), developed by the Scientific Data Management Group under Ari Shoshani, streamlines the task of searching and retrieving requested subsets of data files from massive tape libraries. Although STACS was developed for use on a storage system at a single site (the STAR detector at Brookhaven National Laboratory), the DOE Science Grid envisions applying such capabilities to storage systems distributed among multiple sites. This year STACS was expanded to manage data requests over the Earth Science Grid and the Particle Physics Data Grid, two testbeds for the DOE Science Grid, and won honors in the SC2000 Network Challenge (see page 22).

STACS has three components: a specialized index that allows users to specify a request based on the properties of data they are looking for; a Query Monitor, which coordinates such requests from multiple users; and the HRM (for HPSS Resource Manager), which manages, queues, and monitors file transfer requests to the High Performance Storage Systems (HPSS), such as the ones at NERSC, San Diego Supercomputer Center (SDSC), and other labs. To work on distributed systems, STACS needed a way to find out where the desired files were located among participating sites, and, if the files had been replicated from the original site and stored at another, which file could be retrieved the fastest. A new tool, called the Request Manager, was designed to do just that.

In a successful test run, Request Manager accessed files distributed among six sites using Globus software components. The storage sites were Berkeley Lab, SDSC, the National Center for Atmospheric Research, the Information Sciences Institute (ISI), Argonne National Laboratory, and Lawrence Livermore National Laboratory (LLNL). The Request Manager accepted a request from LLNL for a set of climate modeling files; checked the Globus replica catalog to find replicas of each file; selected the best location from which to get the file using Globus Network Weather Service information; and used the secure Globus FTP to move the files to the destination. The HRM module can pre-stage files to a local disk before moving them, which allows researchers to find the files they need and pre-stage them for transfer at a later time to take advantage of Quality of Services network scheduling. (For further information, see http://gizmo.lbl.gov/sm/.)

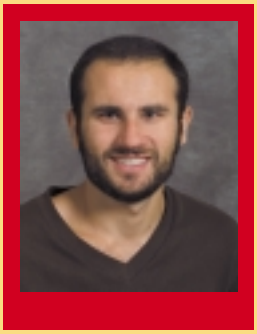
BENCHMARKING AND PERFORMANCE ANALYSIS

There is a growing consensus in the high performance computing community that new benchmarks and performance analysis methods are needed to assess system-level performance running realistic workloads. Theoretical peak performance figures and the scalable Linpack benchmark give little or no insight into system-level efficiency issues.

Recently NERSC embarked on a new focused program in benchmarking and performance analysis to address these challenges. One of our first activities was to create the Effective System Performance (ESP) benchmark suite (http://www.nersc.gov/aboutnersc/pubs/espsc00.pdf). ESP is designed to measure the results of system-level issues such as scheduling efficiency and resource management, job launch times, shutdown-reboot times, and system tools such as backfilling and checkpoint-restart. These factors can make a significant difference in the total throughput of a system.

The ESP suite currently consists of a set of jobs typical of NERSC's workload, which are submitted to a system's batch control facility. The suite includes two full-configuration jobs that test the ability of the system to handle large jobs with high priority. Since the ESP benchmark was first released in November 1999, it has been run on the Cray T3E, the IBM SP, and a Compaq/DEC system. It has provided quantitative data on utilization and scheduling efficacy as well as useful insights on how to manage these systems. The most important conclusion is that certain system functionalities, including checkpoint/restart, swapping, and migration, are critical for efficient scheduling strategies. We plan to modify the ESP test suite so that it can easily be installed and executed on any system.

Future benchmarking activities may include developing a new workload simulation suite, developing an alternative to the Linpack benchmark, collaborating with other centers in developing and applying benchmarks, and developing new models for performance analysis that will help identify hardware and software bottlenecks and contribute to better designs.



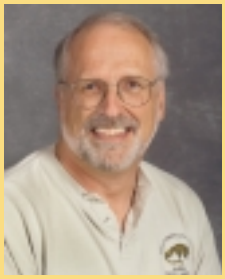
Lenny Oliker

In a related activity, ongoing research into the performance of various architectures and programming models resulted in the Best Student Paper at SC2000, "A Comparison of Three Programming Models for Adaptive Applications on the Origin 2000" by Hongzhang Shan and Jaswinder Pal Singh of Princeton University, Leonid "Lenny" Oliker of NERSC, and Rupak Biswas of NASA Ames Research Center (http://www.nersc.gov/~oliker/papers/sc00.pdf). Lenny (who also won the Best Paper award at SC99) has been advising the principal author on his Ph.D. research.

Computational Science

NERSC staff work closely with scientists in a variety of fields to develop and improve codes for modeling, simulation, and data analysis. For example, Julian Borrill's MADCAP code played a key role in the recent discovery that the Universe is flat (see pages 6 and 67). Other examples are discussed below.

Inna Dubchak, principal author of the VISTA software, will lead the informatics component of a study of comparative genomic analysis of cardiovascular gene regulation under a grant recently awarded by the National Institutes of Health.



David Quarrie

HENP SOFTWARE DEVELOPMENT

NERSC's HENP Computing Group develops software infrastructure for large, international high energy and nuclear physics (HENP) experiments, such as STAR at Brookhaven National Laboratory, BaBar at Stanford Linear Accelerator Center, and ATLAS at CERN. This year David Quarrie, the group's leader, was named chief software architect for ATLAS, an international research program to be carried out at the Large Hadron Collider at CERN beginning in 2005.

The five-story high, 7,000 ton ATLAS detector is designed primarily to find the Higgs boson, which is thought to impart mass to other particles. The Higgs boson is weakly interacting, and will be seen only rarely in the debris of millions of proton collisions. ATLAS will yield up to 1.5 petabytes of data per year for 10 years.

David's job is to establish a coherent vision for the software framework or environment in which scientists will write the physics algorithms they need for the ATLAS experiment, and to manage the development and implementation of the software. The software framework will include on-line data generation and collection, event reconstruction and simulation, and physics analysis.

VISUALIZATION FOR COMPARATIVE GENOMIC ANALYSIS: VISTA

Inna Dubchak of NERSC's Center for Bioinformatics and Computational Genomics led the development of a novel software tool for comparative genomic sequence analysis. Called VISTA (VISualization Tool for Alignments), the software was developed to locate actively conserved regions between species that contain significant genomic synteny. With the required input, two orthologous, contiguous sequences from any two species, and the optional input, the annotation of the one considered the base sequence, VISTA aligns the two sequences and plots the alignment. Color coding identifies the regions of high identity (as defined by the user), the conserved exons, untranslated regions, and non-coding conserved regions.

VISTA combines the GLASS global alignment tool developed at MIT with a visualization and plotting tool developed at Berkeley Lab. Biologists can submit their data and receive their output over the Internet at http:// www-gsd.lbl.gov/vista/. The usefulness of the software has been validated by its appearance in major presentations to the international genomics community.

CALCULATING THE ELECTRONIC STRUCTURE OF LARGE SYSTEMS

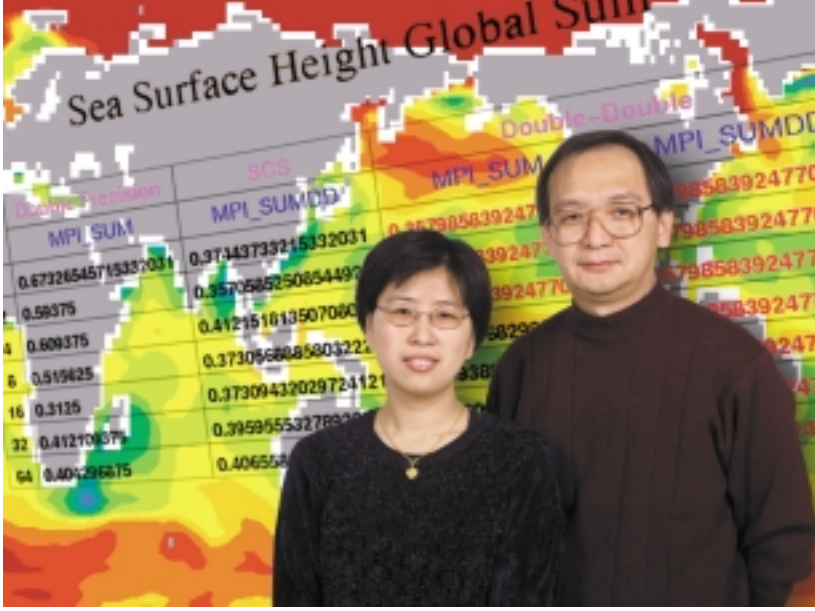
SLCBB, or Strained Linear Combination of Bulk Bands, is a computationally very efficient method for calculating the electronic structure of large systems (up to a million atoms). Lin-Wang Wang of NERSC's Scientific Computing Group helped develop SLCBB while working at the National Renewable Energy Laboratory in Colorado and works with NERSC clients to apply it to their research in the growing field of nanotechnology. He gave an invited talk on the method at the March meeting of the American Physical Society.

As smaller and smaller electronic devices are designed and their sizes shrink from the micron to nanometer scale, certain quantum mechanics effects are introduced. SLCBB allows researchers to calculate the electronic energies and structures of such systems, up to a million or so atoms, and can be run on desktop computers. A related approach (folded spectrum method), also developed by Lin-Wang and coded with the help of Andrew Canning, can also calculate million-atom systems, but requires hundreds of processors.

Before the development of these algorithms, materials scientists could only calculate the electronic structures of systems with hundreds of atoms. Traditionally, there are other methods which allow the calculation of nanometer systems, such as the effective mass method and k.p method, but they are approximated methods that ignore the atomistic features of the wavefunctions. These methods become inappropriate when the size of the system shrinks to a few nanometers.

MERGING THE BEST CLIMATE MODELS

DOE's Climate Change Prediction Program (CCPP, formerly CHAMMP) has awarded an 18-month grant to a multi-agency, multi-laboratory collaboration that aims to develop a modular, performance-portable Climate System Model. Led by Ian Foster of Argonne National Laboratory, the collaboration includes NERSC's Chris Ding and nine other co-investigators from Oak Ridge, Los Alamos, Argonne, and Lawrence Livermore national laboratories and the National Center for Atmospheric Research (NCAR).

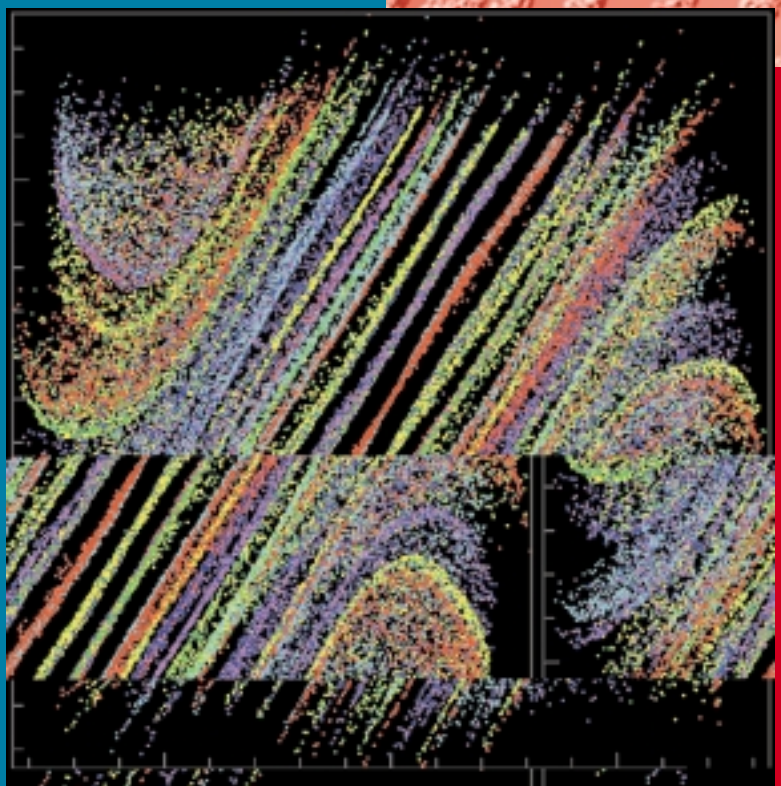


Helen He and Chris Ding have collaborated on two important aspects of climate simulations: input/output performance and numerical reproducibility. Their algorithms for more efficient and reliable codes have been widely adopted by the climate modeling community.

NCAR scientists are working to merge two of the world's most advanced computer climate models, the Climate System Model (CSM) and the Parallel Climate Model (PCM). CSM achieves high performance on parallel vector computers, but was not designed to exploit scalable parallel architectures, and will not scale beyond 64 processors. PCM, developed with DOE support, was designed specifically for parallel systems. The merged CSM-2 will include the best features of both models.

The new R&D work will enable "plug and play" substitution of important modules, making it easier for scientists to improve individual components, and will develop a next-generation "coupler," the top-level model that organizes all the sub-models such as atmosphere, ocean, and sea ice. The result will be a model that performs well on a variety of computer architectures, producing more detailed results in less time. NERSC has two tasks: to optimize input/output and to optimize the code for IBM SP and distributed scalable memory architectures.

Science Highlights



Three-dimensional simulation of a merging-beamlet ion beam injector for heavy-ion fusion. See page 58 for details.

Direct Numerical Simulation of Turbulent Combustion in Homogeneous Charge Compression Ignition (HCCI) Engines

Jacqueline H. Chen, Sandia National Laboratories, Livermore
Hong G. Im, University of Michigan
Ravi Subramanya and Reddy Raghumara, Pittsburgh Supercomputing Center

RESEARCH OBJECTIVES

We are implementing novel computational and modeling tools to understand the effect of mixture composition and temperature variations in homogeneous charge compression ignition (HCCI) engines. A novel turbulence model coupled with autoignition chemistry will be developed and validated against direct numerical simulation (DNS) of turbulent autoignition and flame propagation. The research focus is on the creation of several DNS benchmark databases that will reveal the influence of unsteady stretch effects on turbulent premixed flame propagation and autoignition.

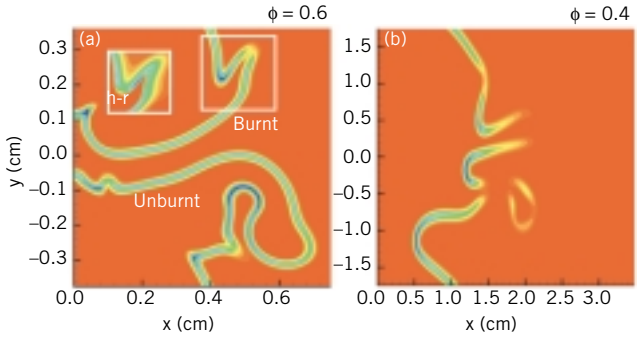
COMPUTATIONAL APPROACH

DNS is used to solve the compressible turbulent reacting governing equations along with boundary and initial conditions. Detailed chemistry and transport models are incorporated including hydrogen-air and hydrocarbon-air mechanisms. Higher-order temporal integration and spatial discretization are used (eighth-order in space, fifth-order in time) along with error monitoring.

ACCOMPLISHMENTS

Five parallel runs were made on the Cray T3E at NERSC. Two runs were made to investigate the effects of unsteady stretch and preferential diffusion on flame propagation in a turbulent premixed hydrogen/air flame. Three runs were made to simulate the effect of turbulent mixing on ignition delay times in a hydrogen/air scalar mixing layer.

We found that strong stretch/preferential diffusion interactions exist in a turbulent flame with computed Markstein numbers varying between –5.34 and 2.85 and area-weighted mean burning velocities 2.4 times the laminar speed. Markstein numbers derived from the DNS data compare favorably with experimental data. We observed that, over a broad range of mixture conditions, as the ratio of the characteristic turbulence to flame time decreases, the flame response to stretch is attenuated. This is consistent with theoretical predictions and recent unsteady counterflow computations.



Isocontour plots of H_2 consumption rate for fuel-air equivalence ratios (a) 0.6 and (b) 0.4. The inset denoted h-r in (a) corresponds to heat release rate.

Autoignition in a hydrogen/air scalar mixing layer in homogeneous turbulence with an initial counterflow of unmixed nitrogen-diluted hydrogen and heated air (1100 °K) showed that away from the steady state ignition turning point, the variation in ignition delay due to variations in turbulence intensity is small. However, for turbulence intensities near the steady ignition limit, ignition delay is sensitive to variations in strain rate.

SIGNIFICANCE

HCCI engines are an attractive alternative to diesel engines, offering the potential for diesel-like efficiencies, while producing extremely low emissions without expensive aftertreatment. The primary technical challenge of HCCI is to control the in-cylinder fluid motion to obtain the desired heat release time profile across the load-speed map of the engine. The computational tools and DNS databases developed in this project will provide a significant jump start to HCCI design efforts.

PUBLICATIONS

J. H. Chen and H. G. Im, "Stretch effects on the burning velocity of turbulent premixed hydrogen-air flames," in *Proceedings of the Twenty-Eighth International Symposium on Combustion* (The Combustion Institute, Pittsburgh, PA, 2000).

H. G. Im and J. H. Chen, "Structure and propagation of triple flames in partially-premixed hydrogen/air mixtures," Combustion and Flame **119**, 436 (1999).

Numerical Simulations of Grain Boundaries, Buried Interfaces, and Catalytic Surfaces

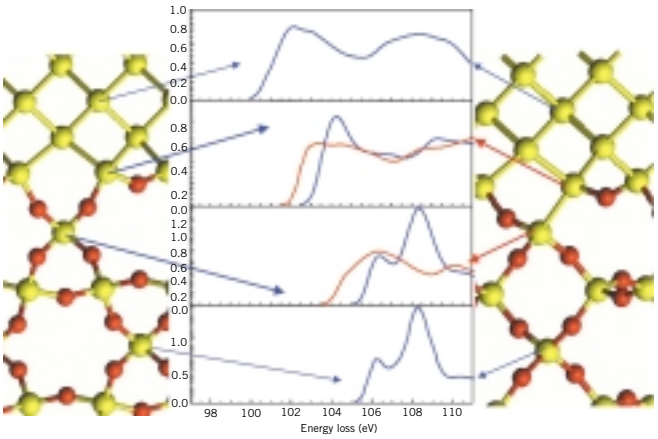
John Cooke, Oak Ridge National Laboratory

RESEARCH OBJECTIVES

This research investigates atomic-scale structures and electronic properties in interfaces, grain boundaries, and catalytic surfaces using self-consistent *ab initio* density functional calculations. Specific projects include (1) the high-efficiency and long-term performance of semiconductor-nanocrystallite catalysts in fixation of CO_2 and conversion to organic compounds; (2) the microscopic origin of the grain boundary potential barrier in yttrium barium copper oxide (YBCO); (3) the relation between quantum yield of perovskite photocatalysts and their microstructure; (4) the structures of Si/SiO_2 and SiC/SiO_2 interfaces; and (5) the growing process of the single-wall nantotube from the Ni-C solution.

COMPUTATIONAL APPROACH

Density functional theory with the local density approximation, pseudopotentials with plane waves or full potential linearized augmented plane waves (FLAPW), and supercells constitute the method of choice for solid state calculations. We have a variety of codes which have been proven successful for our study, including the well-known codes VASP and WIEN97.



Calculations of electron-energy-loss spectra from specific atoms at a model $Si\text{-}SiO_2$ interface with and without suboxide bonds (Si-Si bonds on the oxide side of the interface). Comparison of such curves with experimental spectra leads to a confirmation of the existence of suboxide bonds.

ACCOMPLISHMENTS

Calculations of the surfaces of wurtzite CdSe and zinc-blende CdSe demonstrate that the neutral CO_2 molecules are first chemisorbed into a Se vacancy, where they attract an excess electron. They are then re-emitted taking that electron with them, becoming negatively charged and thus more reactive. The barrier for chemisorption and desorption is small, so the process goes on back and forth and all the CO_2 molecules become charged and reactive. These results account for the observation that CdSe nanocrystals must be Cd-rich (i.e., have Se surface vacancies) to be good catalysts, and for the fact that flat surfaces are not good catalysts.

We have showed that non-stoichiometry is essential for the formation of the grain boundary barrier in $SrTiO_3$, which is a role model for perovskite materials. This non-stoichiometry is confirmed by experiment and is explained by the differences in oxygen vacancy formation energies at the grain boundaries and in bulk as well as by the grain boundary core structure itself.

Using VASP *ab initio* code, we have calculated the momentum angular dependent projected density of states (PDOS) in Si, SiO_2 and SiC materials, and have shown that PDOS of the appropriate symmetry is a good approximation for electron energy loss spectra. These results, together with similar, more exact calculations performed with all-electron LAPW (Wien) code, have resolved a long-standing controversy regarding the role of core excitons in x-ray absorption and electron energy loss spectroscopy.

SIGNIFICANCE

An understanding of atomic-scale structures and electronic properties in interfaces, grain boundaries, and catalytic surfaces is important for both technological and environmental issues, including carbon sequestration; the development of high-temperature superconductors; efficient decomposition of water into H_2 and O_2 to produce energy; building microelectronic devices with oxide thickness less than 30 Å, where the interface dominates the electrical behavior; and carbon nanotubes applications ranging from nano-electronics to super-strong structural materials.

PUBLICATIONS

R. Buczko, S. J. Pennycook, and S. T. Pantelides, "Bonding arrangements at the $Si\text{-}SiO_2$ and $SiC\text{-}SiO_2$ interfaces and the origin of their contrasting properties," Phys. Rev. Lett. **84**, 943 (2000).

M. Kim, G. Duscher, N. D. Browning, S. J. Pennycook, K. Sohlberg, and S. T. Pantelides, "Non-stoichiometry and the electric activity of grain boundaries in $SrTiO_3$," Phys. Rev. Lett. (submitted).

L. G. Wang, S. J. Pennycook, and S. T. Pantelides, "Mechanism for the catalytic activity of CdSe nanocrystals in CO_2 fixation," Phys. Rev. Lett. (in preparation).

Molecular-Based Simulation of Complex Fluids

Peter Cummings, University of Tennessee

RESEARCH OBJECTIVES

Our research is aimed at elucidating the molecular basis for the properties of complex materials and liquid systems, such as lubricants, self-assembling micellar systems, polymers, and high-temperature aqueous solutions.

COMPUTATIONAL APPROACH

We use parallel molecular dynamics codes, developed within our group, running on the NERSC T3E. We use a variety of parallelization strategies, including domain decomposition and data parallel (or replicated data). We have developed our own visualization tool, MDVIZ, which is PVM-based and can be used for remote visualization and steering of ongoing simulations.

ACCOMPLISHMENTS

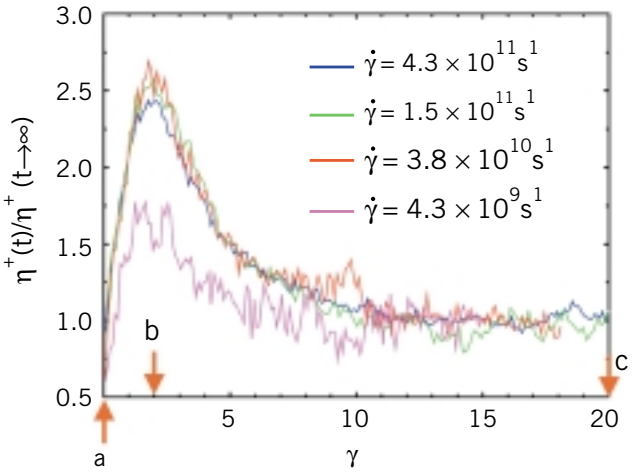
We performed equilibrium and non-equilibrium molecular dynamics simulations of a short polyethylene melt to study the steady-state and transient rheological response of the system to the onset of shear. We computed the diffusion coefficient of the melt under shear — a first for a complex molecular fluid. We developed a novel understanding of the properties of these systems in the non-linear regime based on the measured diffusion coefficients.

We simulated the rheological properties of several lube basestock fluids, and showed that molecular simulation can be used to predict properties such as the viscosity index as well as the pressure-viscosity coefficient at GPa pressures. These properties are used to characterize lubricant performance. The ability to predict these properties via simulation is leading to the molecular design of lubricants. We used molecular simulation of the rheological properties of perfluorobutane to demonstrate that one popular correlation of experimental viscosity data is incorrect.

We performed simulations of dodecane confined to a nanoscale gap between two mica surfaces to identify when solidification of the dodecane could be expected. Our findings are the first to be in full agreement with experiment. In addition, extensive calculations were performed on the self-assembly of reversed micelles in supercritical carbon dioxide. These are the first atomistically detailed simulations to exhibit reversed micellization.

SIGNIFICANCE

This research will lead to better understanding of the basis for the viscous properties of lubricants, leading to the design of improved lubricants in automobile engines, which will, in turn, result in better energy efficiency. We also have significant efforts under way studying the effect of nanoscale confinement on the rheology of lubricants, which has relevance to hard disk drive lubrication. Another focus of our research is aimed at finding new candidates for replacing organic solvents in chemical processes with more environmentally benign alternatives, such as supercritical carbon dioxide. Finally, we perform simulations of supercritical water and aqueous solutions which have relevance to high temperature supercritical water oxidation.



Change in conformation of low (1400) molecular weight polyethylene under shear. The graph shows the phenomenon of stress (or shear) overshoot: when shear is first applied to a polymer, the shear viscosity η initially overshoots its steady state value. The vertical axis is the instantaneous shear viscosity divided by the steady state value, and so it asymptotes to unity in all cases. The horizontal axis is the total shear — the product of the shear rate and the time since shearing began.

PUBLICATIONS

T. Driesner and P. T. Cummings, "Molecular simulation of the temperature- and density-dependence of ionic hydration in aqueous $SrCl_2$ solutions using rigid and flexible water models," J. Chem. Phys. **111**, 5141 (1999).

J. D. Moore, S. T. Cui, P. T. Cummings, and H. D. Cochran, "The transient rheology of a polyethylene melt under shear," Phys. Rev. E **60**, 6956 (1999).

J. D. Moore, S. T. Cui, H. D. Cochran, and P. T. Cummings, "Molecular dynamics study of a short-chain polyethylene melt," J. Non-Newt. Fluid Mech. **93**, 83 (2000).

http://flory.engr.utk.edu

Modeling Molecular Processes in the Environment: In Isolation, in Liquids, on Surfaces, and at Interfaces

T. H. Dunning, B. C. Garrett, M. Dupuis, S. S. Xantheas, D. Feller, K. A. Peterson, L. X. Dang, G. K. Schenter, S. M. Kathmann, E. Arcia, B. Park, and C. J. Burnham,
Pacific Northwest National Laboratory
Y. Borisov, Russian Academy of Sciences
D. Tzeli and A. Mavridis, University of Athens, Greece
G. E. Froudakis, University of Crete, Greece
J. Dai and Z. Bacic, New York University

RESEARCH OBJECTIVES

Research in the Molecular Theory Group is designed to provide a fundamental understanding of how molecular processes in condensed phase systems influence the environment. The goals of the research are:

1. to apply available theoretical techniques to the study of fundamental molecular-level processes that govern the chemistry and physics of natural/contaminated systems and remediation technologies
2. to extend current state-of-the-art methods to treat progressively more complex systems and develop new theoretical techniques that allow us to transcend current computational limitations
3. to incorporate fundamental molecular-level information into models that can simulate dynamical processes in large, complex systems
4. to integrate complementary theoretical approaches for examining multispecies, multiphase systems characteristic of natural and polluted environments and remediation methods
5. to link theory and experiment through collaborative studies.

COMPUTATIONAL APPROACH

Ab initio methodologies developed to model gas-phase molecules and molecular processes are being applied to the study of clusters. The study of clusters provides an effective vehicle for probing the relevant interactions at the molecular level and developing transferable models for different length scales and environments.

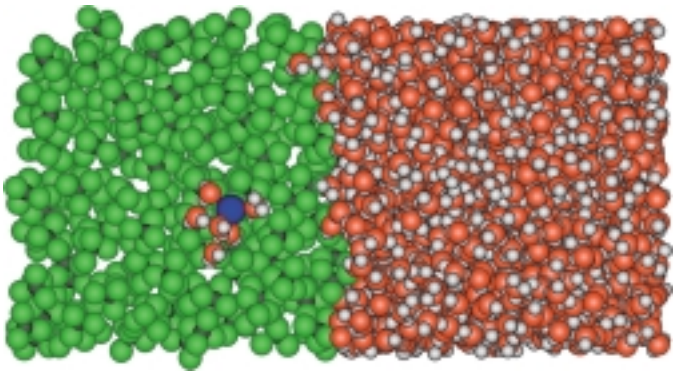
Methods for computing the rates of activated chemical reactions in solution are also being developed, based upon well established gas-phase theories. Because of the importance of reactions involving light atoms that occur in aqueous solutions (e.g., acid and base catalyzed reactions involve proton transfer reactions), the accurate treatment of quantum mechanical effects is a focus of this work.

Equilibrium properties of clusters and solutions, including structural properties (e.g., radial distribution functions and coordination numbers) and average energetics (e.g., enthalpies and free energies) are obtained using classical and quantum statistical mechanics. In addition, time-dependent properties (e.g., correlation functions) are obtained from molecular dynamics (MD) simulations.

Covalently bonded materials such as glasses are being treated using a hybrid of quantum mechanical and classical force fields. Recently, a model potential that treats the bonding and nonbonding interactions separately was developed to specifically address the questions related to the geometric structure of covalently bonded liquids and amorphous materials.

ACCOMPLISHMENTS

Research in the area of aqueous clusters is focused on properties (e.g., structure, energetics, and spectra) of aqueous clusters and aqueous solutions containing inorganic and organic species that occur in natural and contaminated groundwater, and on molecular processes at aqueous interfaces (vapor/liquid, liquid/liquid, and liquid/solid). The goal of this research is to understand the correlation between molecular-scale processes — solvation, association and reaction — and the composition and behavior of species in aqueous environments. Recent results have improved our understanding of the properties of aqueous clusters, infrared spectra of ion-water clusters, thermodynamics of aqueous clusters, acid-base chemistry, chemical reactions of CHCs, benzene-



Transport of Cs^+ ion across a CCl_4/H_2O interface.

water chemistry, chemistry of the water/CHC interface, and ion-ligand binding and ion selectivity of separation agents (e.g., crown ethers and calixarenes).

One of our most significant accomplishments involves the thermodynamics of aqueous clusters. The fundamental definition of a finite temperature cluster was examined in terms of its relation to the measurement of growth kinetics. To accomplish this, a new theoretical approach to the understanding of vapor-phase nucleation was developed. Previous molecular approaches to nucleation focused on the evaluation of the equilibrium distribution of clusters. Our new approach focuses on the evaluation of rate constants for cluster evaporation and condensation. Using variational transition state theory to determine dynamical bottlenecks, a definition of a "physically consistent cluster" naturally falls out of the theory, a result that has eluded the field for the last 30 years.

Research in the area of chemistry and physics of covalently bonded materials (e.g., networked oxides) is focused on studying the properties of amorphous materials involved in waste processing, waste storage, and nuclear fuels. The goal is to provide insight into the long-term performance of materials that contain radionuclides or are exposed to radiation. Recent accomplishments include discoveries concerning the structure and thermodynamics of glasses, the diffusion and reactivity of water in glasses, and radiation damage in oxides.

In addition to the research efforts in the applications areas, we are also developing new theoretical and modeling methods, including basis set development/methods assessment, models for accurate thermochemistry, theory of single molecule chemical dynamics, and interpretation of electron standing wave experiments.

SIGNIFICANCE

Because of the fundamental nature of this research, it supports the mission of the DOE Office of Science to advance basic research that is the foundation for DOE's applied missions in energy resources, environmental quality, and national security. Our research seeks fundamental understanding of chemical transport and reactivity in condensed phases, thermal and nonthermal (i.e., radiation) chemistry, interfacial molecular and ionic transport, and other processes in complex systems related to energy use, environmental remediation, and waste management. A major focus of this research is on fundamental problems in chemical physics that underlie environmental chemistry.

PUBLICATIONS

P. Ayotte, S. B. Nielsen, G. H. Weddle, M. A. Johnson, and S. S. Xantheas, "Spectroscopic observation of ion-induced water dimmer dissociation in the $X^-\bullet(H_2O)_2$ (X = F, Cl, Br, I) clusters," J. Phys. Chem. A **103**, 10665 (1999).

E. R. Batista, S. S. Xantheas, and H. Jonsson, "Electric fields in ice and near water clusters," J. Chem. Phys. **112**, 3285 (2000).

L. X. Dang, "Computer simulation studies of ion transport across a liquid/liquid interface," J. Phys. Chem. B **103**, 8195 (1999).

http://www.emsl.pnl.gov:2080/homes/tms/mt.html

Ab Initio Investigation of Dynamical Electronic Response and Many-Body Effects in Solids

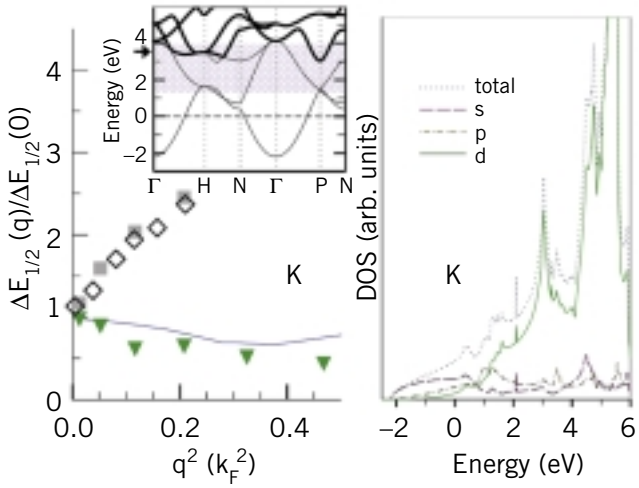
Adolfo Eguiluz, Oak Ridge National Laboratory
James M. Sullivan and Wei Ku, University of Tennessee and Oak Ridge National Laboratory

RESEARCH OBJECTIVES

Our research is devoted to the development of many-body techniques for the *ab initio* study of electron dynamics in strongly correlated materials, including rare-earth metal hydrides, transition metals and transition-metal oxides, narrow-band metals such as Zn and Cd, semiconductors, and insulators. New schemes are being developed for the *ab initio* evaluation of charge- and spin-density response and quasiparticle states in the presence of strong correlations. Two theoretical frameworks are being utilized: time-dependent density-functional theory (TDDFT), and many-body perturbation theory, implemented within the Baym-Kadanoff (BK) method of conserving approximations.

COMPUTATIONAL APPROACH

Our code new.chig.exe calculates the charge and transverse spin response of bulk materials within TDDFT, using realistic all-electron wave functions and band structures. mpi_expSandwich has been used to calculate the charge- and spin-density response of transition metals with shallow semi-core levels, again within TDDFT. PW_GW_jjdiag computes self-consistently the electron self-energy and Green's function within the screened interaction approximation. Two newer codes, pw_SICOEP.exe and genBZ_EXX.exe, evaluate, via the optimized effective potential method, the self-interaction-free potentials which occur in the Kohn-Sham version of DFT. A new code, genBZ_EXX_fxw.exe, is being developed to augment ground state results with the corresponding dynamical exchange-correlation kernel.



Left panel: Plasmon linewidth dispersion obtained upon keeping 3 (triangles) and 6 (squares) valence bands in Kohn-Sham density response function. Inset: LDA band structure of K; the arrow indicates the value of plasmon energy at zero wave vector. Right panel: Calculated density of states (DOS) for potassium — total DOS and contributions from states of *s, p,* and *d* symmetry; the zero of energy is the Fermi level.

ACCOMPLISHMENTS

We have developed a novel technique to perform many-body calculations on the Matsubara time axis, using a non-uniform "power" mesh which allows us to perform fast non-linear interpolation to a uniform mesh. This technique has been implemented to treat, for the first time, the effects of the core electrons on the quasiparticle states in the valence region. We have found novel and important results traced to the effect of the core electrons on the Fock diagram. For example, the impact of this effect on the band gap of Si is of the order of 1 eV, which agrees quite well with experiment.

We have addressed the intriguing lineshape observed in recent electron energy loss experiments on Zn. Our results highlight the enormous impact of *d* band location on the loss spectrum. We have offered a new interpretation of the experimental spectra, identifying the threshold feature as a subtle coherent effect involving *d* electron excitation.

SIGNIFICANCE

TDDFT is a rapidly developing field; its impact on electron dynamics may eventually rival the enormous effect which ground-state DFT has had on materials theory. Most of the recent advances refer to atoms and molecules; our program aims at developing orbital-dependent methods for extended systems.

PUBLICATIONS

W. Ku and A. G. Eguiluz, "Plasmon lifetime in K: A case study of correlated electrons in solids amenable to *ab initio* theory," Phys. Rev. Lett. **82**, 2350 (1999).

A. G. Eguiluz, W. Ku, and J. M. Sullivan, "Dynamical response of correlated electrons in solids probed by inelastic scattering experiments: An *ab initio* theoretical perspective," J. Phys. Chem. Solids **61**, 383 (2000).

A. G. Eguiluz and W. Ku, "*Ab initio* studies of electronic excitations in solids," in *Electron Correlations and Materials Properties*, edited by A. Gonis, N. Kioussis, and M. Ciftan (Plenum, New York, 1999). Available also as cond-mat/9903032.

Computational Studies in Molecular Geochemistry and Biogeochemistry

Andrew R. Felmy, Eric J. Bylaska, James R. Rustad, and T. P. Straatsma, Pacific Northwest National Laboratory

RESEARCH OBJECTIVES

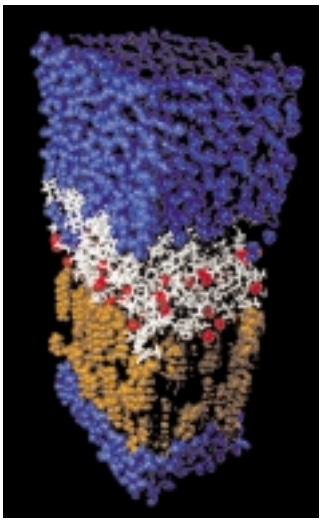
Our effort consists of molecular-level simulations in two key areas of geochemistry and biogeochemistry: (1) microbial surface mediated processes: the effects of lipopolysaccharides present on gram-negative bacteria; and (2) mineral surface interactions: providing a molecular-scale understanding of surface complexation reactions at oxide, oxyhydroxide, and silicate minerals.

COMPUTATIONAL APPROACH

We use a variety of computational chemistry methods, including density functional theory, molecular mechanics/dynamics, Car-Parrinello, and kinetic theories. Besides NWChem, we also use parameterized classical potential models for the interaction of water and hydroxide with Fe/Al surfaces. These models are being used to calculate bulk and surface properties. These models are based upon parameterizations from *ab initio* calculations, and they have been particularly successful in predicting structures, surface charging, and water chemistry of iron-oxide surfaces.

ACCOMPLISHMENTS

Plane-wave pseudopotential methods were used to investigate the structures and total energies of AlOOH and FeOOH in the five canonical oxyhydroxide structures: diaspore (goethite), boehmite (lepidocrocite), akaganeite, guyanaite, and grimaldiite. The local density approximation was used in conjunction with ultrasoft pseudopotentials in full optimizations of both AlOOH and FeOOH in each of these structures. Structures are in reasonably good agreement with experiment, with lattice parameters and bond lengths within 3% of the experimental ones.



Large-scale molecular dynamic simulation of a lipopolysaccharide membrane solvated in a 4.2 nm water box.

An important new code development has benefited from our NERSC computer time. A parallel projector augmented-wave code has recently been completed and is currently in an extensive testing phase. This code will allow us to simulate many new types of materials at a first-principles level, including iron oxides.

An isodesmic procedure based upon density functional theory was developed to predict accurate reaction thermodynamics for important redox half-reactions in the solution phase. This work is an extension of our previous work in which we developed a scheme for predicting the thermodynamics of SN2 reactions in the solution phase.

SIGNIFICANCE

Subsurface microbial processes can control the rates of oxidation/reduction reactions, modify and enhance mineral dissolution and precipitation reactions, and adsorb metals and other ions at the microbial surface. Current theoretical understanding of these processes, which are believed to occur either directly at the microbial surface or at the microbe interface, is very limited.

The ubiquitous occurrence, high specific surface area, and strong binding to a large number of cations, anions, metal ions, and organic chelates makes Fe/Al oxides and oxyhydroxides important adsorbing surfaces. Much of what is known about these adsoption processes on Fe/Al oxides is based upon macroscopic measurements, and relatively little is known at the microscopic level about what types of binding sites exist at oxide surfaces. Difficulties in characterizing the structure and energetics of these sites obstruct the development of improved thermodynamic models for adsorption.

PUBLICATIONS

James R. Rustad, David A. Dixon, Kevin M. Rosso, and Andrew R. Felmy, "Trivalent ion hydrolysis reactions: A linear free-energy relationship based on density functional electronic structure calculations," J. Am. Chem. Soc. **121**, 3234 (1999).

Eric J. Bylaska, David A. Dixon, and Andrew R. Felmy, "The free energies of reactions of chlorinated methanes with aqueous monovalent anions: Application of *ab initio* electronic structure theory," J. Phys. Chem. A **104**, 610 (2000).

James R. Rustad and Kevin M. Rosso, "The structures and energies of AlOOH and FeOOH polymorphs form plane wave pseudopotential calculations," American Mineralogist (submitted, 2000).

Photonic Band Gap Materials

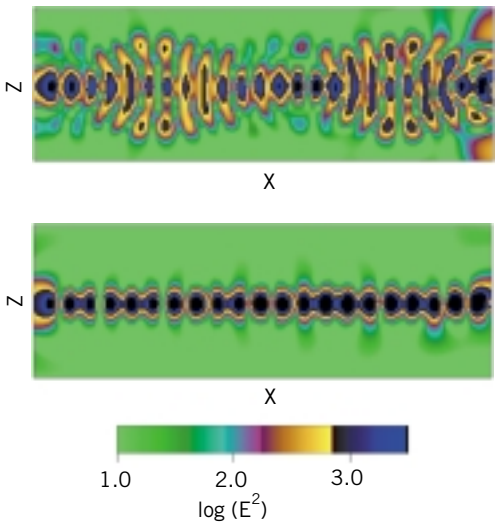
Bruce Harmon, Kai-Ming Ho, Costas Soukoulis, Rana Biswas, Ihab El-Kady, Dave Turner, and Mihail Sigalas, Ames Laboratory

RESEARCH OBJECTIVES

This research project has two concurrent thrusts: (1) The design and development of novel structures and photonic devices in the infrared and optical regimes with full 3D band gaps. This includes unusual colloidal crystals being fabricated by our collaborators here at Ames Laboratory, for photonic gaps at optical wavelengths. (2) Development of applications in the microwave and millimeter wave regime for existing 3D photonic band gap (PBG) crystals. These include novel waveguides that can bend electromagnetic waves with bending radii of the order of a single wavelength.

COMPUTATIONAL APPROACH

(1) The transfer matrix method (TMM) is used to design and study PBG filters. Maxwell's equations are solved to determine the reflection and transmission of electromagnetic waves from a finite thickness PBG material. The method can incorporate realistic absorption and frequency dependent dielectric functions.



Electric field intensity in a 2D photonic crystal waveguide. Top: Dielectric cladding in the z-direction. The photonic crystal is periodic in the x-y direction. The dielectric contrast between the core and the cladding is 12.5/9.5. Bottom: Air cladding in the z-direction. The dielectric contrast between the core and the cladding is 12.5/1.0. The higher contrast achieves better confinement of the wave in the core.

(2) In the finite difference time domain (FDTD) method, Maxwell's equations are discretized on a real-space grid. The time evolution of the electromagnetic fields is calculated by solving the time-dependent Maxwell's equations. This code can simulate the radiation properties of antennas or the bending of light in a photonic crystal.

ACCOMPLISHMENTS

In collaboration with Bilkent University, we fabricated an exceptionally directional antenna utilizing a Fabry-Perot cavity between two photonic crystals. The beam had a half-power width of less than 10°. Very good agreement was achieved between experiment and the finite difference simulation.

We performed calculations to design a microcavity within a PBG crystal with a complete band gap in the infrared. The crystal has been fabricated at Sandia National Laboratories.

Finite-difference simulations were used to design a planar waveguide in the 3D layer-by-layer crystal with a 90° bend with 100% transmission through the bend. A similar L-shaped waveguide was also simulated in a metallic photonic crystal, and only three unit cells thickness were needed for 85% transmission efficiency. The performance of waveguides in 2D PBG structures has been simulated, and the guiding efficiency was optimized as a function of the structural parameters. Such structures are important in all-optical photonic crystal devices.

SIGNIFICANCE

Computational simulation can rapidly test the electromagnetic behavior of new structures and then select the best performing ones for fabrication. Our approach led to novel photonic lattices fabricated at Sandia that can for the first time manipulate 1.5 micron wavelengths used for optical fibers.

PUBLICATIONS

B. Temelkuran, M. Bayindir, E. Ozbay, R. Biswas, G. Tuttle, M. M. Sigalas, and K.-M. Ho, "Photonic crystal based resonant antenna with a very high directivity," Journal of Applied Physics **87**, 603 (2000).

R. Biswas, M. M. Sigalas, G. Subramania, C. M. Soukoulis, and K.-M. Ho, "Photonic band gaps of porous solids," Physical Review B **61**, 4549 (2000).

I. El-Kady, M. M. Sigalas, R. Biswas, and K.-M. Ho, "Dielectric waveguides in two-dimensional photonic bandgap materials," Journal of Lightwave Technology **17**, 2042 (1999).

http://cmp.ameslab.gov/ercap/pbg2001.html

Electronic Structure and Simulation Calculations of Reactive Chemical Systems

Martin Head-Gordon and David Chandler, University of California, Berkeley and Lawrence Berkeley National Laboratory

RESEARCH OBJECTIVES

Our research centers on the development and application of methods that predict the electronic structure of interesting molecules. We seek to open new classes of chemical problems to study via electronic structure theory, including large molecules and extended systems, molecular excited states, and nonadiabatic interactions.

COMPUTATIONAL APPROACH

Our approach includes electronic structure methods of the density functional theory type, Newtonian molecular dynamics, Car-Parinello *ab initio* molecular dynamics (CPMD), and transition path sampling methods.

ACCOMPLISHMENTS

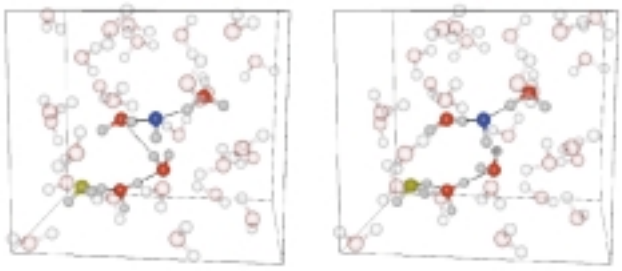
We have shown that a relatively simple theoretical approach, time-dependent density functional theory, can accurately describe electronic excited states of radicals including those with bielectronic character. Previously such states were considered very challenging to characterize.

We have produced calculations of unprecedented accuracy on large polycyclic aromatic hydrocarbon (PAH) cations of the form that arise as intermediates in combustion processes on the way to forming soot particles, and are also believed to play a significant role in interstellar carbon chemistry. Our calculations resulted in a new assignment of the visible spectrum of the perylene cation. Our results open the way to further studies of larger cation radicals.

Through the development of transition path sampling, we have created algorithms that permit the study of rare events without prior knowledge about mechanisms or transition states. The technique is an importance sampling of trajectory space that is based upon our discoveries of statistical mechanics and thermodynamics governing dynamical systems that can be far from equilibrium.

We have created a formulation of transition path sampling that can be interfaced with any trajectory algorithm, to harvest only those trajectories of interest or of importance. We have used this formulation to combine transition path sampling with CPMD.

We have devised a set of statistical methods for interpreting the behavior of transition paths in complex systems. These tools provide a type of pattern recognition for interpreting dynamics in a complex, high-dimensional system.



Configurations of liquid water within a few femtoseconds before (left) and after (right) liquid water has passed through the transition state surface of autoionization. The dashed lines show the hydrogen bond wire that must break in order for the system to cross the transition state surface. Yellow and blue identify the hydroxide and hydronium ions, respectively.

By harvesting reactive trajectories of autodissociation in liquid water, we have discovered molecular details of a fundamental chemical reaction. These results have established coordinates in liquid water that may be important for many bond-breaking and, in particular, proton transfer processes.

SIGNIFICANCE

Electronic structure theory has emerged as a valuable counterpart to direct experiments for the study of reactive species that may not be characterized easily (if at all) in the laboratory. Our research on chemical and conformational transformations of biomolecules is beginning to yield a novel microscopic picture of biochemical dynamics. Our results may have significant implications for the general understanding of solvent roles in chemical and biochemical processes.

PUBLICATIONS

S. R. Gwaltney and M. Head-Gordon, "A second-order correction to singles and doubles coupled-cluster methods based on a perturbative expansion of a similarity-transformed Hamiltonian," Chem. Phys. Lett. **323**, 21 (2000).

P.G. Bolhuis, C. Dellago, and D. Chandler, "Reaction coordinates of biomolecular isomerization," Proceedings of the National Academy of Sciences **97**, 5877 (2000).

P. L. Geissler, C. Dellago, D. Chandler, J. Hutter, and M. Parrinello, "*Ab initio* analysis of proton tranfer dynamics in $(H_2O)_3H^+$," Chem. Phys. Lett. **321**, 225 (2000).

http://www.cchem.berkeley.edu/~mhggrp

Fluid and Tracer Transport in Self-Affine Rough Fractures

Joel Koplik, German Drazer, and Igor Baryshev,
City College of the City University of New York

RESEARCH OBJECTIVES

We are concerned with flow, passive tracer dispersion, and depositional processes in the self-affine fractures often observed in naturally fractured rock. In such systems the rock surfaces have long-range correlations which have been shown to significantly enhance and modify the dynamics of passive tracers used as diagnostic tools. We further wish to examine the motion of solids suspended in flowing fluid in fractured geological formations, and in particular study the evolution of the fractured pore space as deposition occurs.

COMPUTATIONAL APPROACH

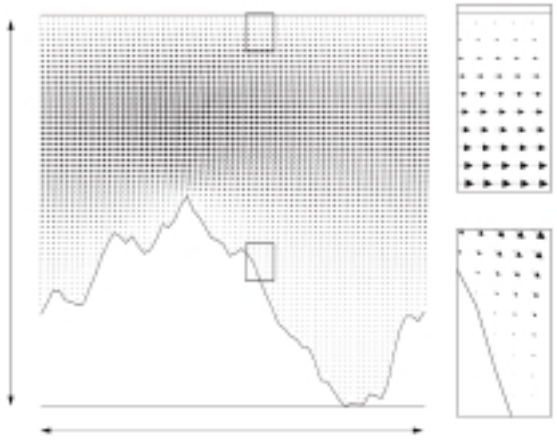
The problems of flow, tracer motion and particle deposition in self-affine fractures require an efficient method for solving the Navier-Stokes and convection-diffusion equations in a highly irregular and evolving geometry. The lattice Boltzmann method is optimal for problems in complicated regions because the core of the calculation is the motion of particles in the region's interior, with an adjustment to the motion when a boundary is reached. As the pore space evolves, only the geometrical specification of the solid region is needed. Other workers have used this method for both active and passive tracer dispersion studies, as well as the somewhat analogous problem of suspension dynamics.

ACCOMPLISHMENTS

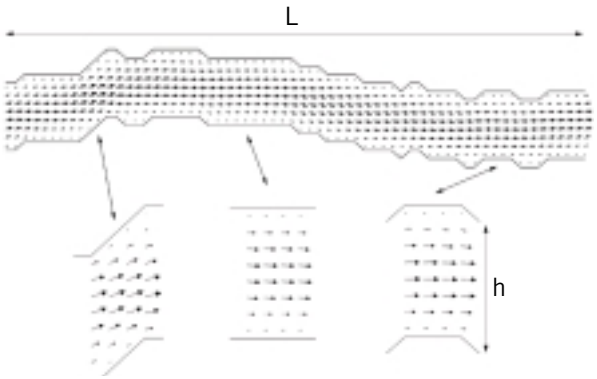
One of last year's two projects, simulations of deep-bed filtration in a statistically homogeneous porous medium, was completed. The second project on fracture flows continued this year. To date, we have developed scaling relations for the permeability of 2D self-affine fractures and verified them by numerical simulations.

SIGNIFICANCE

The efficient extraction of water and hydrocarbon resources from underground reservoirs, as well as the use of underground formations as waste disposal sites, requires a full understanding of the dynamics of the flow of fluids and various suspended matter in the disordered porous media which comprise geological formations and reservoirs. The proposed research looks at the transport and deposition of solid particulates which may clog or perhaps break open new flow channels in these systems, and the effects of subtle correlations resulting from fracture processes on tracer tests. In addition, some of the results are relevant to commercial filtration processes used in purification and manufacture.



Example of the geometry and velocity field in a 2D fracture with one planar and one self-affine surface of roughness exponent H = 0.8. The enlargements show the difference in the velocity decay near smooth and rough boundaries.



Flow field in a narrow self-affine fracture with a constant gap, and Hurst exponent H = 0.8. The vertical aperture is constant everywhere, but the effective local aperture for fluid flow, that is, the local width of the channel normal to the mean flow direction, strongly depends on the local angle between the surface and the mean plane. The enlargements illustrate the effect of the effective aperture on the flow field.

PUBLICATIONS

J. Lee and J. Koplik, "Network model for deep-bed filtration," Phys. Fluids (submitted, 2000).

G. Drazer and J. Koplik, "Permeability of self-affine rough fractures," Phys. Rev. E (submitted, 2000). E-print cond-mat/0006287.

J. Lee and J. Koplik, "Microscopic motion of particles passing through a porous medium," Phys. Fluids **11**, 76 (1999).

Quantum and Classical Simulations of Clusters, Nanostructured Materials, and Friction Control

Uzi Landman, Robert N. Barnett, Charles L. Cleveland, Hannu Hakkinen, and W. David Luedtke,
Georgia Institute of Technology

RESEARCH OBJECTIVES

This project is investigating the microscopic physical and chemical processes underlying the properties of novel materials. These investigations aim at discovering and elucidating size-dependent evolutionary patterns of materials properties, bridging the molecular, cluster, and condensed-phase regimes.

COMPUTATIONAL APPROACH

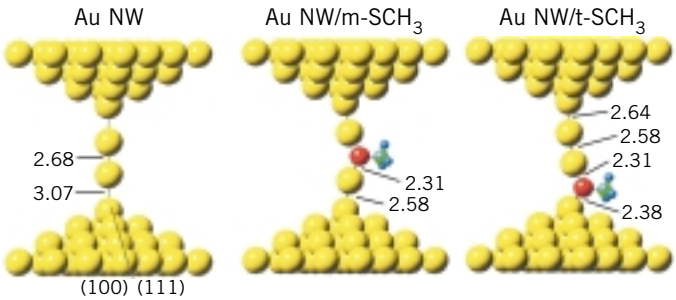
Our computational approaches include large-scale classical molecular dynamics, employing tested many-body interactions, and *ab initio* quantum molecular dynamics (in conjunction with norm-conserving non-local pseudopotentials and a plane-wave basis) based on local-spin density functional theory (LSD) with the inclusion of generalized exchange-correlation gradient corrections. In these *ab initio* simulations, the dynamics of the ions evolve on the concurrently calculated electronic ground state (Born-Oppenheimer, BO) potential energy surface, using the BO-LSD-MD method. We also employ various structural optimization methods (conjugate-gradient and variants thereof, simulated annealing and genetic algorithms), as well as an arsenal of analysis techniques, including animation.

ACCOMPLISHMENTS

We reported on *ab initio* local-density functional investigations of the atomic structure, electronic spectrum, and conductance of a gold nanowire consisting of a four-atom chain connected to gold electrodes. We explored structural and electronic spectral modifications resulting from adsorption of a molecule (methylthiol, SCH_3) to the wire. These results provide a new interpretation of the measured electron microscopy image of the atomic gold wire and suggest a new strategy for formation of organo-metallic nanowires, as well as the use of nanowires as monitoring and chemical sensing devices.

In contrast to the inert nature of gold as bulk material, nanosize particles of gold supported on various oxides, as well as two-monolayer-thick gold islands of up to 4 nm diameter on titania, were found to exhibit an enhanced catalytic activity, in particular for the low-temperature oxidation of CO. We demonstrated the size dependence of the activity of nanoscale gold clusters, with Au_8 found to be the smallest size to catalyze the reaction.

Other studies included photoelectron spectra of aluminum cluster anions; spontaneous symmetry breaking in single and molecular quantum dots; structures, solvation forces, and shear of molecular films in a rough nano-confinement; formation, stability, and breakup of nanojets; and metal-semiconductor nanocontacts.



Equilibrium structures of a bare gold nanowire (AuNW, left) and of wires chemically modified by adsorption of a SCH_3 molecule, with the molecule adsorbed in the middle of the wire (m-SCH_3) or at the vicinity of the tip (t-SCH_3). Yellow spheres correspond to Au atoms, and in the chemically modified wires S, C, and H atoms are depicted by red, green, and blue spheres, respectively. Marked distances are in units of Å.

SIGNIFICANCE

Understanding the microscopic origins of the properties of materials with reduced physical dimensions is essential for the utilization of such materials systems in advanced technologies, including miniaturization of electronic and mechanical devices, development of sensors, design of novel logic gates and information storage strategies using quantum dots, control of friction under extreme conditions, cluster-catalysts, and atomic-scale materials manipulations. *Small is different* — new and often unexpected behavior emerges when the physical size of the materials system is reduced to microscopic scale.

PUBLICATIONS

J. Akola, M. Manninen, H. Hakkinen, X. Li, L.-S. Wang, and U. Landman, "Photoelectron spectra of aluminum cluster anions: Temperature effects and ab initio simulations," Phys. Rev. B **60**, R11297 (1999).

A. Sanchez, S. Abbet, U. Heiz, W.-D. Schneider, H. Hakkinen, R. N. Barnett, and U. Landman, "When gold is not noble: Nanoscale gold catalysts," J. Phys. Chem. A **103**, 9573 (1999).

H. Hakkinen, R. N. Barnett, and U. Landman, "Atomic gold nanowires and their chemical modification," J. Phys. Chem. B **103**, 8814 (1999).

Quantum Monte Carlo for Electronic Structure of Combustion Systems

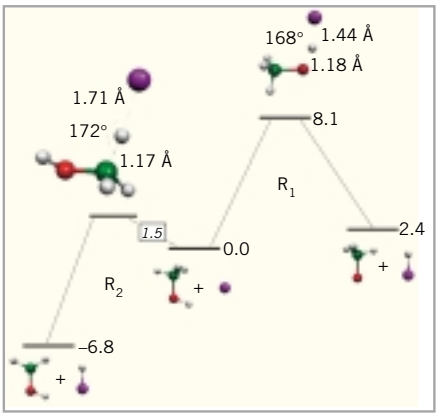
William Lester and Michael Frenklach, University of California, Berkeley, and Lawrence Berkeley National Laboratory
Alan Aspuru-Guzik, Ivan Ovcharenko, and Nigel Moriarty, University of California, Berkeley
Olivier Couronne, John Harkless, and Alexander Kollias, Lawrence Berkeley National Laboratory
Xenophon Krokidis, Institut Français du Pétrole (IFP), Rueil-Malmaison, France
Zhiwei Sun and Ruzeng Zhu, Institute of Mechanics, Academica Sinica, Beijing, China

RESEARCH OBJECTIVES

This research is directed primarily toward high accuracy studies to enable the characterization of the reaction pathways (1) leading to the formation of the first aromatic ring in high temperature environments and subsequent reactions ultimately leading to soot formation and (2) governing combustion reactions of small organic alcohols such as methanol.

COMPUTATIONAL APPROACH

Our dominant computational technique is the quantum Monte Carlo (QMC) method in the diffusion Monte Carlo (DMC) variant. Our version of DMC employs effective core potentials to minimize computational effort. Variational Monte Carlo computations are carried out to test trial functions for DMC constructed as products of independent particle wave functions and correlation functions that depend on interparticle distances.



Two reaction channels are identified for hydrogen abstraction from methanol by chlorine atom. The R1 channel had been found previously, but the R2 channel was found in this group by O. Couronne with the involvement of F. Gilardoni. The R2 channel is consistent with a direct mechanism implied by molecular beam scattering experiments of M. Ahmed, D. S. Peterka, and A. G. Suits at the Berkeley Lab Advanced Light Source.

ACCOMPLISHMENTS

The addition reaction of acetylene and propargyl was investigated using several density functional methods. DMC calculations were performed at optimized geometries for many of the equilibrium structures and transition states. This data serves as input into the determination of RRKM rate constants leading to the formation of the cyclo-C_5H_5 radical. A detailed analysis of the reaction pathways is under way using the bonding evolution theory (BET) concepts applied to electron localization function (ELF).

Hydrogen abstraction from methanol by atoms and radicals yields as products either hydroxymethyl or methoxy radicals in competing reaction channels, depending on which nonequivalent hydrogen of the methyl or hydroxyl group of CH_3OH is abstracted:

$$R1: CH_3OH + Cl \rightarrow CH_3O + HCl$$

$$R2: CH_3OH + Cl \rightarrow CH_2OH + HCl$$

Computations of these reaction channels to date tend to support experimental results showing the importance of a direct reaction mechanism, but QMC calculations, in progress, are needed to resolve the matter.

QMC was used to compute the atomization energy and the heat of formation of the propargyl radical, C_3H_3, and the computed results compare favorably with experimental measurements. The effective core potential and fixed-node approximations were used in the DMC variant. Two generalized gradient approximation density functionals were applied for comparison.

SIGNIFICANCE

(1) Full characterization of the mechanism of soot formation will provide valuable insight on how to reduce a major pollution source. (2) Methanol is an attractive alternative fuel because its combustion generates fewer air pollutants than gasoline. The reaction channels for the formation of CH_3O and CH_2OH are of fundamental importance for combustion and atmospheric chemistry.

PUBLICATIONS

J. C. Grossman, W. A. Lester, Jr., and S. G. Louie, "Quantum Monte Carlo and density functional theory characterization of 2-cyclopentenone and 3-cyclopentenone formation from O(3P) + cyclopentadiene," J. Am. Chem. Soc. **122**, 705 (2000).

X. Krokidis, N. W. Moriarty, W. A. Lester, Jr., and M. Frenklach, "Propargyl radical: An electron localization function study," Chem. Phys. Letters **314**, 534 (1999).

J. C. Grossman, W. A. Lester, Jr., and S. G. Louie, "Cyclopentadiene stability: Quantum Monte Carlo, coupled cluster, and density functional theory determinations," Mol. Phys. **96**, 629 (1999).

Semiclassical Simulations for Reaction Dynamics: Methodological Studies and Biological Applications

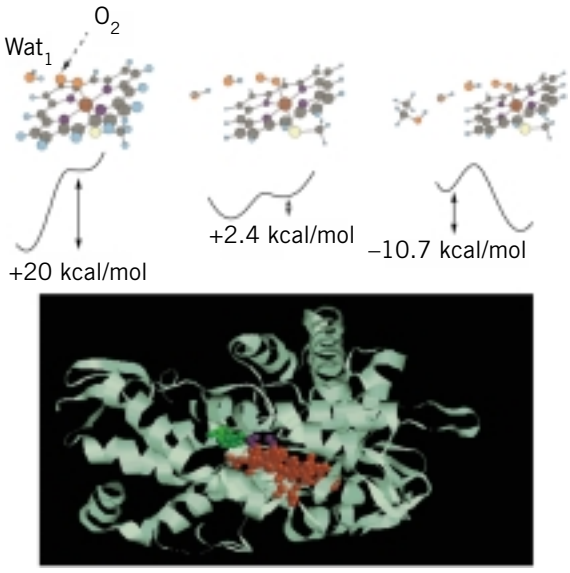
William Miller, Victor Guallar, and Eduardo Coronado,
University of California, Berkeley

RESEARCH OBJECTIVES

This research continues the development of accurate semiclassical methods and extends its application to large biological systems. Specific projects include (1) methodological studies to explore the capabilities of new approximate versions of semiclassical initial value representation (SC-IVR), (2) SC-IVR to study the reaction dynamics in medium-sized systems, (3) semiclassical Wentzel-Kramers-Brillouin (SC-WKB) applications in real-sized enzymatic systems, and (4) *ab initio* potential surface development.

COMPUTATIONAL APPROACH

Semiclassical molecular dynamics are evaluated under two different approaches, SC-IVR and SC-WKB. SC-IVR evaluates a high dimensional integral over initial conditions for semiclassical trajectories, according to a classical equations of motion. SC-WKB evaluates the mean position of the system following a classical trajectory propagation, where tunneling effects are added in terms of sudden transfer at the different classical turning points along the desired reaction coordinate. The potential energy surface used in SC-WKB is calculated on the flight with QM/MM methods, combining quantum chemistry in the active site with classical mechanics for the rest of the system, which accurately generate energies and gradients on an enzyme potential surface.



Semiclassical WKB dynamics combined with density functional theory (DFT) *ab initio* calculations and molecular mechanics were used to describe the initial proton transfer from WAT1 to the distal oxygen (O_2) in the active site of Cytocrome P450eryf. DFT active site studies demonstrates the crucial role of a water channel assisting the proton transfer and leading to an overall exothermic reaction.

ACCOMPLISHMENTS

Accomplishments in FY 2000 include: (1) The most rigorous simulations to date of the excited-state time-dependent proton-transfer dynamics associated with the tautomerization reaction of a relatively large polyatomic system. This computational effort requires a separate calculation of a 140-dimensional integral and demands running approximately 10^7 semiclassical trajectories. (2) The first direct evidence of the protonation in the activation cycle of a member of the P450 family. We have shown that the mechanism only requires an optimal orientation of the bound molecular oxygen and the presence of a dynamically stable hydrogen bond network. (3) We have extended the Meyer-Miller Hamiltonian to describe nonadiabatic processes of systems with more than two electronic degrees of freedom. (4) We have developed a new SC-IVR methodology to describe tunneling with real trajectories by a mapping approach. (5) A log-derivative formulation of the prefactor term appearing in the SC-IVR propagator has been developed to avoid the branch cut problem which has hampered previous formulations.

SIGNIFICANCE

Molecular dynamics simulations that include quantum effects in the description of reaction dynamics offer a powerful approach to elucidate enzymatic reaction mechanisms at the atomic level. Understanding biological chemical processes at the atomic level will have a major impact on the drug and biotechnology industries.

PUBLICATIONS

E. Coronado, V. Batista, and W. H. Miller, "Nonadiabatic photodissociation dynamics of *ICN* in the $\tilde{A}$ continuum: A semiclassical initial value representation study," J. Chem. Phys. **112**, 5566 (2000).

V. Guallar, V. Batista, and W. H. Miller, "Semiclassical molecular dynamics simulation of intramolecular proton transfer in photo-excited 2-2('-Hydroxyphenil)-oxazole," J. Chem. Phys. (submitted, 2000).

V. Guallar, V. S. Batista, and W. H. Miller, "Semiclassical molecular dynamics simulations of excited state double proton transfer in 7-azaindole dimers," J. Chem. Phys. **110**, 9922 (1999).

Magnetic Materials: Bridging Basic and Applied Science

G. Malcolm Stocks, Oak Ridge National Laboratory
Bruce N. Harmon, Ames Laboratory, Iowa State University
Michael Weinert, Brookhaven National Laboratory

RESEARCH OBJECTIVES

The central goal of this project is to develop rigorous approaches to both refining and bridging the models that describe magnetic phenomena on different length scales. An important aspect of this is the development of modeling tools capable of integrating atomic-level understanding of magnetic properties and interactions with structure and microstructure.

COMPUTATIONAL APPROACH

First principles density functional (DFT) methods are applied to calculating fundamental magnetic properties. Constrained density functional theory is used to calculate magnetic moments and constraining (forcing) fields. Spin dynamics (SD) is used to treat the spin degrees of freedom in simulations of finite temperature and non-equilibrium properties. At the macroscopic length scale, micromagnetics calculations based on the Landau-Lifshitz-Gilbert (LLG) equations are used to study microstructural effects on domain wall motion and switching in devices.

ACCOMPLISHMENTS

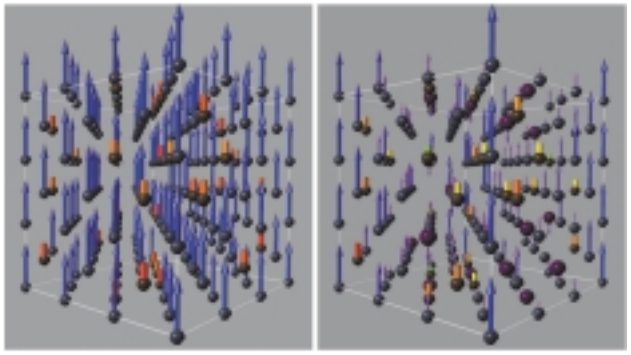
The effects of tantalum on the magnetic structure of permalloy were studied using the coherent potential approximation to Korringa-Kohn-Rostoker (KKR-CPA) method. The effects of Ta additions on local moments were also investigated using the locally-self-consistent multiple-scattering (LSMS) method and large supercell calculations.

Using first principles electronic structure methods, the energy of a magnetic domain wall in CoPt was evaluated for the ideal crystal and for one with an anti-phase boundary (APB) defect. This is the first time first principles calculations have been used to study domain walls and to evaluate quantitatively the consequences of defects on domain wall pinning.

We determined the structural properties and relative phase stability of $NiAl_3$ and $NiAl_2$. The compositional defects for $NiAl_3$ were found to be Ni antisites for the Ni-rich alloys and Ni vacancies on the Al-rich side.

SIGNIFICANCE

Modeling tools capable of integrating atomic-level understanding of magnetic properties and interactions with structure and microstructure would allow the prediction of technologically relevant magnetic properties and the design of improved electromagnetic devices as well as electric motors with reduced weight and improved performance.



Although Ta is a necessary component of permalloy (Py)-based heterostructures used in magnetic random access memory devices, it has deleterious effects on the magnetic properties, resulting in magnetic dead layers and reduced performance. First principles studies of Permalloy show that the addition of Ta not only reduces the average magnetic moment associated with the Ni and Fe atoms but also results in large fluctuations in the Ni site moment magnitude. Left: Length and color of arrows show the magnitude of the magnetic moments associated with the Ni sites (mid-gray spheres) in Py. The much larger moments of the Fe site (dark-gray spheres) have been truncated to aid the visualization. Right: For an alloy of 90% Py and 10% Ta, color and length coding of the magnitude has been preserved to illustrate the reductions and variations in the size of Fe and Ni moments. Interestingly, the magnetic moments associated with the Ta atoms (purple spheres) are antiferromagnetically aligned with respect to the Fe and Ni moments due to hybridization effects.

PUBLICATIONS

B. Újfalussy, X.-D. Wang, D. M. C. Nicholson, W. A. Shelton, G. M. Stocks, Y. Wang, and B. L. Gyorffy, "Constrained density functional theory for first principles spin dynamics," J. Appl. Phys. **85**, 4824 (1999).

V. P. Antropov, B. N. Harmon, and A. V. Smirnov, "Aspects of spin dynamics and magnetic interactions," J. Magnetism and Magnetic Mater. **200**, 148 (1999).

Q. Niu, Xindong Wang, L. Kleinman, Wu-Ming Liu, D. M. C. Nicholson, and G. M. Stocks, "Adiabatic dynamics of local spin moments in itinerant magnets," Phys. Rev. Letters **83**, 207 (1999).

http://theory.ms.ornl.gov/~gms/M4home.html

Computational Semiconductor Physics

Alex Zunger, National Renewable Energy Laboratory

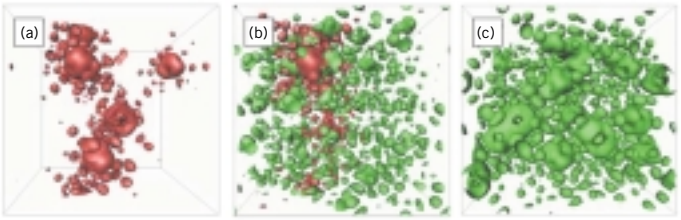
RESEARCH OBJECTIVES

Our calculations allow us to predict in detail the effect of nanoscale atomic structure on the electronic and optical properties of semiconductor systems. Using atomistic models in conjunction with quantum mechanical methods, we are able to interpret excitonic spectra, provide feedback to experiment, and predict new properties to be engineered and investigated. Our studies focus on both one-body electronic structure and properties and many-body (configuration interaction) treatments.

COMPUTATIONAL APPROACH

We use a combination of methods to bridge the length and computational cost scales from the 100–1,000 atom microstructural scale, where we obtain thermodynamic information and compute fully relaxed geometries of complex structures such as impurity complexes and surfaces, to the 100,000–1,000,000 atom nanostructure regime, where the optoelectronic properties are determined by the near gap conduction and valence states.

We use local density approximation based methods for small systems, and empirical pseudopotential based methods, such as the folded spectrum and linear combination of bulk bands methods, for large-scale nanostructures. Our optimized pseudopotential methods allow us to study million-atom systems with quantum mechanical accuracy. Using the single particle wave functions, we are able to treat many-body effects, important for optical properties and effects such as Coulomb blockade, by means of a configuration interaction based approach.



Localized-to-delocalized transition in GaAsN alloys: Calculated conduction band wavefunction isosurfaces of the GaAsN alloy from 14,000 atom supercell plane-wave calculations. At low nitrogen concentrations (a), the lowest energy wavefunctions are highly nitrogen localized (shown in red). With increased nitrogen concentration (b), the lowest energy conduction levels consist of both localized (red) and delocalized levels (green). For higher nitrogen concentrations (c), the conduction levels become increasingly delocalized (green).

ACCOMPLISHMENTS

In FY2000 we successfully studied several classes of nanostructure systems:

1. Alloy dots, arrays: We predicted the excitonic exchange splitting of Si dots, predicted the electron-addition spectra of InP and CdSe dots, and developed a theory of lens-shaped self-assembled InAs/GaAs dots. Using our many-body configuration-interaction approach, we predicted failures of both Hund's rule and the Aufbau principle in quantum dots.

2. Nitrides: We developed the first theory of localization in InGaN alloys and successfully explained anomalous pressure effects in GaAsN.

3. Short-range order: We developed the first quantitative theory of alloy precipitate shape in metal alloys, successfully predicted the size and shape of precipitates versus temperature of Al-Zn, and demonstrated the first quantitative theory of phase-stability of brass (Cu-Zn).

SIGNIFICANCE

The electronic, optical, transport, and structural properties of semiconductor nanostructures (films, quantum dots, and quantum wires) and microstructures in alloys are important because of their potential application to lasers, sensors, photovoltaics and novel quantum devices. These structural features occur on distance scales of ~100–500 Å, thus encompasing 10^4–10^5 atoms. Ours is the only available pseudopotential-based theory which can address this size scale. Understanding the underlying physical phenomena in these systems is essential to designing nanoscale devices with custom-made electronic and optical properties.

PUBLICATIONS

A. Franceschetti and A. Zunger, "Inverse band-structure problem of finding an atomic configuration with given electronic properties," Nature **402**, 60 (1999).

S. B. Zhang, S. H. Wei, and A. Zunger, "Microscopic origin of the phenomenological equilibrium 'doping limit rule' in n-Type III-V semiconductors," Phys. Rev. Lett. **84**, 1232 (2000).

A. Franceschetti and A. Zunger, "Hund's rule, spin blockade, and the Aufbau principle in strongly confined semiconductor quantum dots," Europhysics Letters **50**, 243 (2000).

http://www.sst.nrel.gov

Studies of Decadal Climate Dynamics and Predictability

Tim Barnett and David Pierce,
Scripps Institution of Oceanography

RESEARCH OBJECTIVES

To predict natural climate variability in the North Pacific on the decadal time scale. This is a key region, as it is known that the state of the North Pacific sea surface temperature (SST) field is well correlated with wintertime precipitation and temperatures over the U.S.

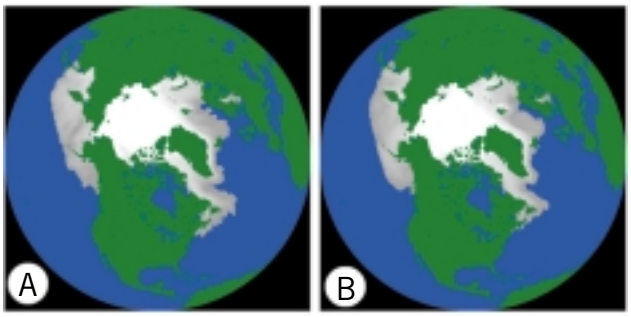
COMPUTATIONAL APPROACH

Our primary tool is the Parallel Climate Model (PCM), a fully coupled general circulation model. The individual ocean and atmosphere components from this model are used separately as well. The resolution of the atmospheric component of the models used in this study is T42, while that of the ocean components varies from about 0.5° to 1°, depending on the latitude and longitude.

ACCOMPLISHMENTS

In previous work we had identified a preferred 20-year timescale of climate variability in the North Pacific in a coupled ocean-atmosphere general circulation model. Our objective in FY 2000 was to show whether or not this preferred timescale arises out of coupled ocean-atmosphere interactions (and therefore might be predictable by a coupled model) or is a manifestation of various so-called "stochastic resonance" mechanisms, which can give enhanced spectral variability at a preferred timescale even in the absence of coupled ocean-atmosphere interactions.

We ran the ocean model component from the fully coupled model in standalone mode, forced first with the saved surface flux fields (heat, fresh water, and momentum) from the fully coupled model. This control run showed that if the ocean model is driven by these saved flux fields, it reproduces the enhanced spectral peak at the same timescale (20 years/cycle) as found in the fully coupled model. Next, we forced the ocean model with the saved flux fields, but scrambled randomly in time by month. We found that the spectral peak is still present in the scrambled forcing run, indicating that most, if not all, of the enhanced energy at a preferred timescale of 20 years/cycle comes from stochastic resonance mechanisms, and therefore is inherently unpredictable. Further analysis of this data in under way.



Initializing a coupled model to observed initial conditions is a difficult challenge, and critical to climate predictions. Sea ice is one of the most sensitive model components to an incorrect initial state. This figure shows a successful initialization of the PCM to observed conditions of 1995. Panel A: Model sea ice cover in January 1995 — the initial state. Panel B: Model sea ice cover in January after running the model forward for 30 years with no constraints. The good agreement with the initial conditions is just one indication that the model drift away from the imposed initial conditions based on observations is small.

SIGNIFICANCE

Determining the levels of natural climate variability, and being able to understand the physical processes responsible for this natural noise, are a requirement of any attempt to make an early detection of mankind's impact on climate.

PUBLICATIONS

T. P. Barnett, D. W. Pierce, M. Latif, D. Dommenget, and R. Saravanan, "Interdecadal interactions between the tropics and midlatitudes in the Pacific basin," Geophys. Rev. Lett. **26**, 615 (1999).

T. P. Barnett, D. W. Pierce, R. Saravanan, N. Schneider, D. Dommenget, and M. Latif, "Origins of the midlatitude Pacific decadal variability," Geophys. Rev. Lett. **26**, 1453 (1999).

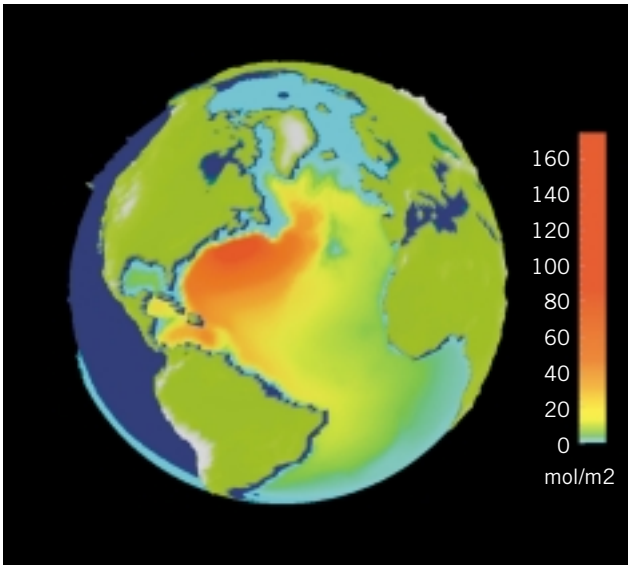
D. W. Pierce, T. P. Barnett, and M. Latif, "Connections between the Pacific Ocean tropics and midlatitudes on decadal time scales," J. Climate **13**, 1173 (2000).

Simulating Ocean Carbon Sequestration

Kenneth Caldeira, Lawrence Livermore National Laboratory
James K. Bishop, Lawrence Berkeley National Laboratory
Kenneth Coale, Moss Landing Marine Laboratory
Paul Falkowski, Rutgers University
Howard Herzog and Sallie Chisholm,
Massachusetts Institute of Technology
Russ Davis, Scripps Institution of Oceanography
Gerard Nihous, Pacific International Center
for High Technology Research

RESEARCH OBJECTIVES

The research objectives of the DOE Center for Research on Ocean Carbon Sequestration (DOCS) are (1) to understand the efficacy and impacts of various strategies proposed for ocean carbon sequestration; (2) to focus research of other groups on the key uncertainties and/or deficiencies in ocean physics and biogeochemical models; and (3) to develop the best numerical simulations of ocean carbon sequestration, both with regard to biological fertilization and direct injection of CO_2 into the deep ocean. This will be accomplished by incorporating the research of other groups into an improved model of ocean physics and biogeochemistry for application to the problem of ocean carbon sequestration.



Results of an injection simulation in which CO_2 was continuously injected off of the coast of New York at a depth of 3,025 m and a rate of 0.1 Pg of carbon per year. The total amount of injected carbon in each column is shown after 100 simulated years.

COMPUTATIONAL APPROACH

For our ocean physics model, we are making a transition from the LLNL verion of GFDL's MOM to the LANL POP model. Some modifications to the POP code will be made to improve the numerics of handling point sources with high spatial concentration gradients. We will use a variety of standard techniques to accelerate convergence on stationary initial conditions for our model experiments (e.g., time step splitting). Because some ocean sequestration strategies involve point sources, and the numerics of the models assume relatively small spatial concentration gradients, we will explore a number of techniques for handling these large gradients within the model. These techniques include testing various tracer advection schemes (e.g., flux corrected transport, the QUICK scheme, etc.), and using results from a high resolution regional model (run at MIT) to initialize the global GCM.

ACCOMPLISHMENTS

We have performed the highest resolution simulations of direct CO_2 injection ever performed on a global scale. Results of simulations of anthropogenic CO_2 uptake in the Southern Ocean were published in Science magazine. We have also performed simulations of iron fertilization of the ocean.

SIGNIFICANCE

The rapid accumulation of CO_2 in the atmosphere could produce adverse environmental impacts. Therefore, we must understand the options available to us to slow this accumulation. To meet this need, DOE has created the DOCS, which is jointly managed by Lawrence Berkeley and Lawrence Livermore national laboratories and works in close collaboration with a variety of researchers in academia and other institutions. Our research will help provide the science base needed to understand the effectiveness and the environmental impacts of various ocean carbon sequestration strategies.

PUBLICATIONS

K. Caldeira and P. B. Duffy, "The role of the Southern Ocean in uptake and storage of anthropogenic carbon dioxide," Science **287**, 620 (2000).

H. Herzog, K. Caldeira, and E. E. Adams, "Carbon sequestration via direct injection," in *Encyclopedia of Ocean Sciences*, edited by J. Steele, S. Thorpe, and K. Turekian (Academic Press Ltd., London, in press).

K. Caldeira and G. H. Rau, "Accelerating carbonate dissolution to sequester carbon dioxide in the ocean: Geochemical implications," Geophys. Res. Lett. **27**, 225 (2000).

http://esd.lbl.gov/DOCS

Theoretical Study on Catalysis by Protein Enzymes and Ribozyme

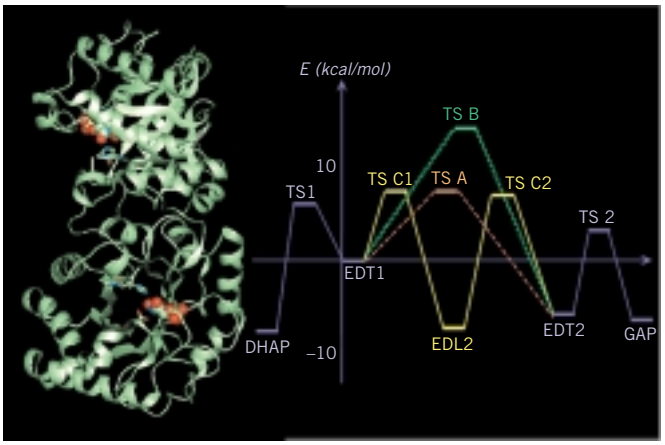
Martin Karplus, Harvard University

RESEARCH OBJECTIVES

The goal of this project is to develop a greater understanding of the mechanisms involved in enzyme catalysis and related protein functions. We are studying two types of enzymes: proteins and a nucleic acid (hammerhead ribozyme).

COMPUTATIONAL APPROACH

For active-site models in the gas phase, *ab initio* or density functional (DFT) calculations are used. A few calculations with continuum dieletric models are carried out to investigate the effect of solvation; the results are compared with those in the enzyme. Those calculations are carried out using mainly Gaussian98 and NWChem. To determine the catalytic mechanism in the presence of the enzyme environment, a combined quantum/molecular mechanics (QM/MM) approach is used, performed with the CHARMM program. To study the effect of tunneling on proton or hydride transfer, variational transition state theory is used.



The energetics were determined for three mechanisms proposed for TIM catalyzed reactions. Results from reaction path calculations suggest that the two mechanisms that involve an enediol intermediate are likely to occur, while the direct intra-substrate proton transfer mechanism (in green) is energetically unfavorable due to the presence of His95 in the active site.

ACCOMPLISHMENTS

Hammerhead ribozymes: Ab initio and DFT calculations have been carried out to study the reaction path in the phosphate ester hydrolysis of an RNA model which represents a minimum active site of the hammerhead ribozyme. Our results help explain the loss of catalytic activity observed experimentally when replacing bridging or non-bridging oxygen atoms from the phosphate group.

Yeast chorismate mutase (YCM): Preliminary calculations have been carried out for the rearrangement of chorismate to prephenate in the gas phase and in solution. Optimized structures for several comformations have been obtained from *ab initio* and DFT calculations. Our calculations identified a pathway for the elimination reaction from chorismate to 4-hydroxybenzoate and explained the experimental observation that the relative rates of the rearrangement and elimination reactions depend on solvents.

Triosephosphate isomerase (TIM): Three catalytic mechanisms proposed in the literature were studied with the combined DFT/MM approach. The two pathways that involve an enediol species were found to be give similar values for the barriers, in satisfactory agreement with experiment. The mechanism that involves intramolecular proton transfer in the enediolate was found to be energetically unfavorable due to the presence of His95. We also applied variational transition state theory (VTST) to investigate the effect of tunneling on two proton transfer steps in TIM. It was found that tunneling has a significant but not very large effect on the rate constants at room temperature, and appears to be consistently more significant in enzyme than for the corresponding reaction in solution.

SIGNIFICANCE

Details of the chemical mechanisms employed by enzymes to serve as catalysts of biochemical reactions remain elusive, largely because the chemical events of bond formation and cleavage are exceedingly short and are currently inaccessible to direct experimental measurement. Theoretical studies, therefore, provide valuable insights into enzyme catalysis.

PUBLICATIONS

P. D. Lyne and M. Karplus, "Determination of the pKa of the 2'-hydroxyl group at the active site of hammerhead ribozyme from *ab initio* calculations with solvation corrections," J. Am. Chem. Soc. **122**, 166 (2000).

Q. Cui and M. Karplus, "Molecular properties from combined QM/MM methods. I. Analytical second derivative and vibrational calculations," J. Chem. Phys. **112**, 1133 (2000).

Q. Cui and M. Karplus, "Molecular properties from combined QM/MM methods. II. Chemical shifts in large molecules," J. Phys. Chem. B **104**, 3721 (2000).

Computational Structural Genomics

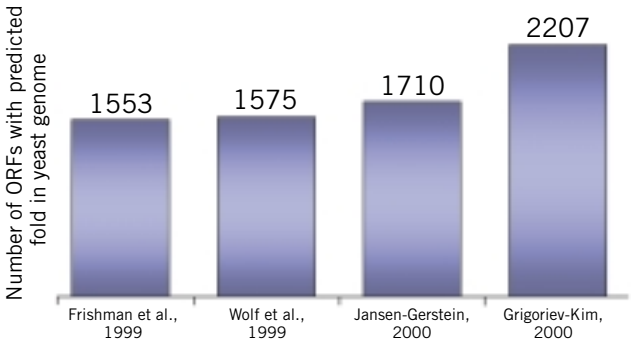
Sung-Hou Kim and Igor Grigoriev,
Lawrence Berkeley National Laboratory

RESEARCH OBJECTIVES

Our objective is to use computation to obtain the basic structural set for the organisms with completely sequenced genomes. We identify and classify protein folds in complete genomes to find new targets for structural analysis, predict protein structures when possible, and use computational tools for solving protein structures from X-ray data.

COMPUTATIONAL APPROACH

We have developed two different methods, Proximity Correlation Method (PCM) and Dual Profile method (DuP), for detecting similarity of protein folds by comparison of protein sequences. In these methods we combine secondary structure predictions with correlation of physical (in PCM) or evolutionary (in DuP) properties of amino acid residues. Segments of sequences rather than single residues are compared, which substantially increases sensitivity in detecting remote homologues, i.e., proteins with similar folds and low sequence similarities that cannot be detected by standard sequence comparison techniques. For *ab initio* structural predictions, we derived new energy potentials for contacts between amino acid residues in protein structures and developed a procedure for efficient structure calculation by torsion angle dynamics.



Using sensitive fold recognition methods (PCM and DuP) allows us to assign folds for more proteins encoded in the complete genome of yeast, *S. cerevisiae*.

ACCOMPLISHMENTS

Pilot projects for protein fold prediction and comparison have been conducted for a few complete genomes. Detailed analysis of protein fold topology allowed us to extend a library of protein fold templates and increase the number of predictions in complete genomes. Results of the pilot projects revealed the folds of several hypothetical proteins in the *Methanococcus jannaschii* genome, clusters of proteins with same the fold or fold pattern in genomes of *Mycoplasmas*, fold population and functional/structural relationship of yeast proteins, and structural relatedness of proteins from other organisms.

We achieved good preliminary results in *ab initio* prediction of protein structure. Statistical analysis of all known protein structures allowed us to derive a new potential of contact energy for pairs of amino acid residues. Using this potential and secondary structure predictions from torsion angle dynamics, we were able to predict contacts between helices in small globular proteins, and we built low-resolution protein structures for 28 out of 36 small helical proteins based only on sequence information — the best results that have been achieved by any method.

SIGNIFICANCE

Structural characterization of proteins can help to understand protein function, especially when protein sequence does not show any significant similarities to proteins of known function. The computational aspect of structural genomics is important because it (1) directs experimental efforts to potential targets, (2) reduces the time for solving protein structures, and (3) predicts fold/structure for proteins whose structure is difficult to determine experimentally (due to low expression, solubility, etc.).

PUBLICATIONS

S.-H. Kim, "Structural genomics of microbes: An objective," Curr. Opin. Struct. Biol. **10**, 380 (2000).

C. Zhang and S.-H. Kim, "Environment-dependent residue contact energies for proteins," Proc. Natl. Acad. Sci. USA, **97**, 2550 (2000).

C. Zhang and S.-H. Kim, "The anatomy of protein β-sheet topology" J. Mol. Biol. **299**, 1075 (2000).

Computational Studies of Protein Folding

Peter Kollman, University of California, San Francisco

RESEARCH OBJECTIVES

(1) To study the early stage of the folding processes of small proteins. (2) To identify important folding intermediates by limited folding simulations using locally enhanced sampling (LES). (3) To refine and evaluate the free energy of structure predictions for small proteins.

COMPUTATIONAL APPROACH

We use the AMBER molecular mechanics simulation program suite and Gaussian quantum mechanical calculation packages.

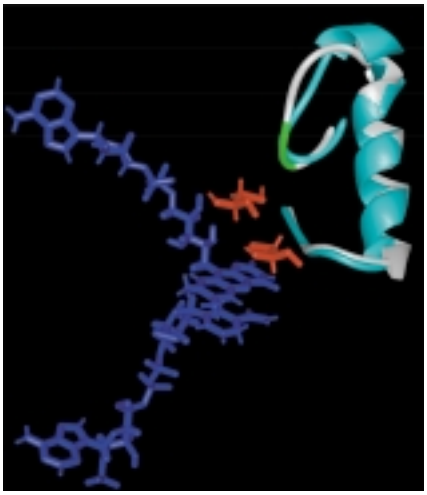
ACCOMPLISHMENTS

We have significantly improved both the single-CPU performance and the scalability of the molecular dynamics code in AMBER. Overall, we have achieved a factor of 6 speedup over the existing code and significantly improved the scalability.

We completed a series of microsecond-scale molecular dynamics simulations on small proteins, including a full microsecond and two 200 ns simulations on the villin headpiece and four 200 nsec simulations on BBA1 that started from fully unfolded states. These simulations marked the beginning of direct simulations of the folding process with detailed all-atom representations of both protein and solvent that may help us to achieve a full elucidation of protein folding mechanisms.

From these simulations, we identified a highly native-like marginally stable folding intermediate as well as other compact intermediate states whose radii of gyration (i.e., size) are comparable to or smaller than that of the native state. This suggests the existence of multiple compact intermediate states that may play roles in the folding process and supports the notion that the barrier separating the folded native state and the unfolded (or partially folded) non-native states may be entropic.

On the other hand, all four folding simulations on BBA1 yielded a similar structure in which the helical secondary structure formed early and was maintained throughout the remainder of the simulations. This observation suggests that the folding process of this protein may follow a simple secondary-tertiary mechanism in which stable secondary helical structures form in the early stages and the completion of the folding process is marked by the formation of the tertiary contacts between these secondary structures. This scenario is distinct from that of the villin headpiece, in which the secondary structures were only partially formed even when the tertiary contacts started to form. Taken together, the results suggest diverse folding mechanisms.



A portion of the mercuric ion reductase enzyme, with electron transfer co-factors shown in blue and an internal pair of cysteine residues in red. A critical aspect of the hypothesized mechanism for the enzyme is the flexibility of the c-terminal helix and loop, which allows an outer pair of cysteines to bind mercuric ion from the surroundings and pass it to the inner pair of cysteines for oxidation. The crystal structure c-terminal helix and loop are shown as the white ribbon structure, with the outer pair of cysteins shaded green. One structure from a molecular dynamics simulation is shown as the cyan ribbon, illustrating the flexibility of this region.

SIGNIFICANCE

Elucidation of the mechanism of protein folding has remained a scientific challenge for decades. Molecular dynamics simulation with full representation of solvent possesses a unique advantage to study protein folding due to their atomic level resolution and accuracy.

PUBLICATIONS

Y. Duan and P. A. Kollman, "Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution," Science **282**, 740 (1998).

C. Simmerling, J. L. Miller, and P. A. Kollman, "Combined locally enhanced sampling and Particle Mesh Ewald as a strategy to locate the experimental structure of a nonhelical nucleic acid," J. Am. Chem. Soc. **120**, 7149 (1998).

Y. Duan and P. A. Kollman, "Toward a realistic simulation of the folding process of small proteins," J. Phys. Chem. (submitted).

http://www.amber.ucsf.edu

Breaking the Scalability Limit of Parallel Molecular Dynamics: Simulating the Fastest Folding Proteins in Atomic Detail

Vijay Pande, Stanford University

RESEARCH OBJECTIVES

We propose a new algorithm for parallel molecular dynamics (MD), which bootstraps on top of current parallel methodologies and uses relatively little communication bandwidth. This allows scaling of current codes to 10 to 1000 times more processors than currently possible. With this new method, we will be able to simulate considerably longer timescales than currently possible — instead of tens of nanoseconds, tens of microseconds. This increase is particularly important for protein folding, since the fastest proteins fold in the microsecond regime. Thus we will be able for the first time to directly fold small proteins using MD simulations in all-atom detail.

COMPUTATIONAL APPROACH

Our computational approach takes advantage of the inherent kinetics in our system. Like other free energy barrier crossing problems, a protein wanders around one free energy minimum, waiting for some rare thermal fluctuation to push it over the barrier. It has been demonstrated that the time to cross the barrier is much less than the overall folding time — most of the time is spent waiting for this fluctuation. Our method parallelizes this "waiting stage": starting from some initial coordinates, we run M parallel simulations (each with different initial velocities). If a system has multiple barriers, this method can be extended to handle them as well.

We bootstrap this method on top of current parallel MD (using NAMD). Each simulation is a 16-processor NAMD simulation. Thus, we get double benefits of parallelization: from traditional MD and from our "ensemble dynamics" method.

Do proteins designed by man fold like their natural counterparts? We compared the folding of two designed zinc fingers to a human zinc finger and found that the folding pathway is similar in all three proteins: all fold via a partially structured, α-helical intermediate state.

ACCOMPLISHMENTS

We have simulated the unfolding pathways of two small proteins (zinc fingers) which fold to the same structure but have very little sequence similarity. We have shown that in spite of little similarity in sequence, these proteins have similar unfolding pathways. This lends strong evidence that native state topology is a primary determinant of the nature of the folding pathway.

SIGNIFICANCE

Computationally, this method should be broadly applicable to any free energy barrier crossing problem, and should be useful for anyone using MD simulations. Scientifically, the folding of a protein in all-atom detail is a holy grail of modern computational biology. Understanding how proteins fold has broad implications in many areas, including protein design, protein misfolding (believed to be related to several diseases), and the design of self-assembling protein-like nanostructures.

PUBLICATIONS

V. S. Pande and D. S. Rokhsar, "Molecular dynamics simulation of unfolding and refolding of a b-hairpin fragment of protein G," Proc. Natl. Acad. Sci. U.S.A. **96**, 9062 (1999).

V. S. Pande and D. S. Rokhsar, "Folding pathway of a lattice model for proteins," Proc. Natl. Acad. Sci. U.S.A. **96**, 1273 (1999).

V. S. Pande, A. Yu. Grosberg, and T. Tanaka, "Heteropolymer freezing and design: Towards physical models of protein folding," Rev. Mod. Phys. **72**, 259 (2000).

http://www.stanford.edu/group/pandegroup

Parallel Implementation of Enhanced Atmospheric and Oceanic Dynamical Cores in the Next-Generation NCAR CCSM

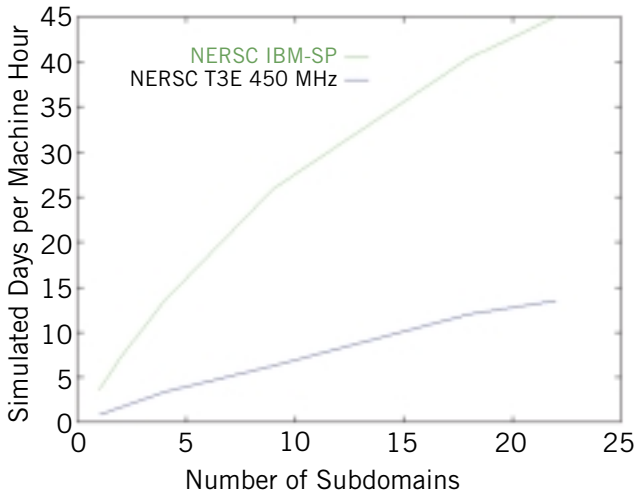
Doug Rotman, Art Mirin, John Tannahill, José Milovich, and Phil Duffy, Lawrence Livermore National Laboratory

RESEARCH OBJECTIVES

In collaboration with NCAR and NASA, as well as Lawrence Berkeley, Los Alamos, Oak Ridge, and Argonne National Laboratories, we are developing, implementing, and enhancing the computational capabilities of the next-generation NCAR Community Climate System Model (CCSM). In particular, we are improving the performance and scalability of the NASA Lin-Rood dynamical core in the Community Climate Model (CCM3/4) and the barotropic solver in the Parallel Ocean Program (POP) model.

COMPUTATIONAL APPROACH

We are expanding the capabilities of the Lin-Rood dynamical core by implementing 2D message passing domain decomposition along with enhanced use of OpenMP within a processing node. The horizontal discretization is built upon the flux form semi-Lagrangian transport algorithms, which have been extended to the shallow water dynamical framework. The piecewise parabolic method (PPM) is used as the 1D building block for multi-dimensional dynamics and transport. Our approach is to use a 2D domain decomposition of latitude and altitude for the dynamics, transposing the data to a 2D latitude/longitude decomposition for column physics calculations, and then transposing back to latitude/altitude for dynamics.



The enhanced Lin-Rood dynamical core shows excellent scaling (~80% efficient) up to 18–20 processors on the NERSC Cray T3E and IBM SP2.

To improve the performance and scalability of the barotropic solver, two approaches are being tried: (1) To parallelize the baroclinic solver (which scales well) using typical 2D ocean domain decomposition, but to carry out the barotropic solver on a small number of processors to reduce the communication latency and time. (2) To implement a new solver that reduces communication needs by reorganizing the calculation to allow maximum use of local cell information to update the barotropic velocities. This new solver uses wave front recursion to eliminate interior variables within each domain. This allows the construction of a reduced system of equations that involves only those variables that are involved in communication across nearest neighbor domain boundaries.

ACCOMPLISHMENTS

We have carried out timing and scalability studies of the Lin-Rood dynamical core on a variety of platforms around the DOE complex. The NERSC IBM SP2 and Cray T3E showed excellent scaling (~80% efficiency) up to 18–20 processors. We began to see a strong reduction in parallel efficiency and performance as we moved to 22 latitude subdomains (the maximum allowed at this resolution). When run in a full climate model, this dynamical core is coupled with a column physics package to update state variables. We tested transpose libraries to ensure accurate results as well as analyzed transpose timings to ensure efficient parallel execution.

In the POP ocean barotropic solver, we have implemented the wave front recursion elimination solver on a single processor using a five-point stencil. Tests have been completed on the use of differing decompositions for the baroclinic and barotropic solvers in the ocean model. By reducing the processors used for the barotropic solver, a 25% increase in POP ocean model throughput has been achieved. This is the result of having fewer but larger messages and hence reducing latency.

SIGNIFICANCE

NASA, NCAR, and DOE are jointly developing a next-generation Community Climate System Model, which will incorporate a higher degree of physical consistency than is realized in the current generation of spectral and finite-difference models. This project aims to enable model simulations on large parallel machines so that longer and greater numbers of long climate simulations can be carried out.

PCMDI: Interpreting Differential Temperature Trends at the Surface and in the Lower Troposphere

B. D. Santer, C. Doutriaux, J. S. Boyle, J. J. Hnilo, K. E. Taylor, and M. F. Wehner, PCMDI,
Lawrence Livermore National Laboratory
T. M. L. Wigley and G. A. Meehl,
National Center for Atmospheric Research
D. J. Gaffen, National Oceanic and Atmospheric Administration
Air Resources Laboratory
L. Bengtsson, M. Esch, and E. Roeckner, Max Planck Institute
for Meteorology, Hamburg, Germany
P. D. Jones, University of East Anglia, Norwich, UK

RESEARCH OBJECTIVES

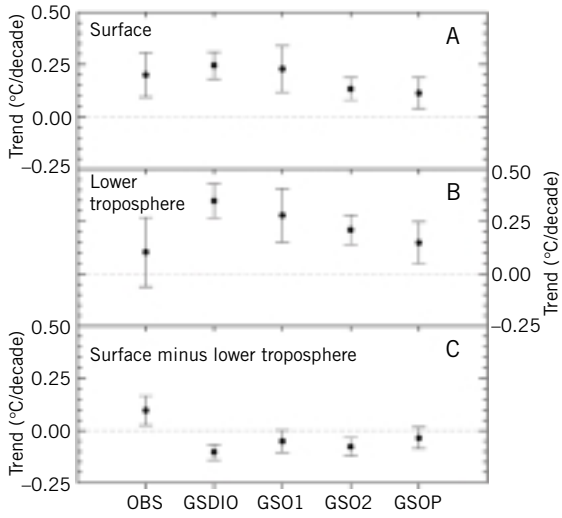
The principal mission of the Program for Climate Model Diagnosis and Intercomparison (PCMDI) is to develop improved methods and tools for the diagnosis, validation, and intercomparison of global climate models and to engage in research on a wide variety of outstanding problems in climate modeling and analysis. In this project, we tested various interpretations of the apparent difference between estimated global-scale temperature trends at Earth's surface (as recorded by thermometers) and in the lower troposphere (as monitored by satellites).

COMPUTATIONAL APPROACH

We performed three tests — NOMASK, VARMASK, and FIXMASK — to investigate the effect of coverage differences on the estimated trend differential between two data sets: the satellite-based Microwave Sounding Unit (MSU) tropospheric temperature data, and the Intergovernmental Panel on Climate Change (IPCC) surface data. We then analyzed data from 300-year control integrations performed with three models — ECHAM4/OPYC, PCM, and CSM — to explore whether the residual trend difference of roughly 0.1 °C per decade could be explained by natural variability of the climate system on decadal time scales. Finally, we used a set of three perturbation experiments to test the effects of external forcing.

ACCOMPLISHMENTS

Our results show that the observed difference between surface and tropospheric temperature changes cannot be fully explained by coverage differences between satellite- and surface-based measurement systems and/or the effects of natural internal climate variability. However, we find that both effects may make substantial contributions to the observed trend difference. A recent model result suggests that the observed warming of the surface relative to the lower troposphere may be a response to combined forcing by well-mixed greenhouse gases, sulfate aerosols, stratospheric ozone, and the effects of the Pinatubo eruption in June 1991. Further simulations of the climate of the past two decades are needed to determine the precise causes of the temperature trend difference.



Least-squares linear trends and associated 95% confidence intervals in modeled and observed surface, 2LT (lower troposphere), and surface – 2LT temperature time series (panels A, B, and C, respectively). Observed trends and confidence intervals over 1979–1998 were computed with annual-mean, spatially-averaged IPCC surface and MSUd 2LT data, using FIXMASK masking. Model-based results are from experiments with anthropogenic and natural forcings performed with ECHAM. Model data were processed such that results are given for only one 20-year period in GSDIO ("1979–1998" in model years) and for one 19-year period in GSO1, GSO2, and GSOP ("1979–1997" in model years). Model results are also based on FIXMASK sampling. Confidence intervals are adjusted to account for temporal autocorrelation in the data.

SIGNIFICANCE

Understanding the difference between surface and tropospheric temperature trends is crucial for modeling climate, explaining and attributing climatic changes, and planning for future climate monitoring.

PUBLICATIONS

B. D. Santer, T. M. L. Wigley, D. J. Gaffen, L. Bengtsson, C. Doutriaux, J. S. Boyle, M. Esch, J. J. Hnilo, P. D. Jones, G. A. Meehl, E. Roeckner, K. E. Taylor, and M. F. Wehner, "Interpreting differential temperature trends at the surface and in the lower troposphere," Science **287**, 1227 (2000).

http://www-pcmdi.llnl.gov/

Large-Eddy Simulations of Marine Boundary-Layer Clouds for Climate Studies and Investigations of Global Aerosols

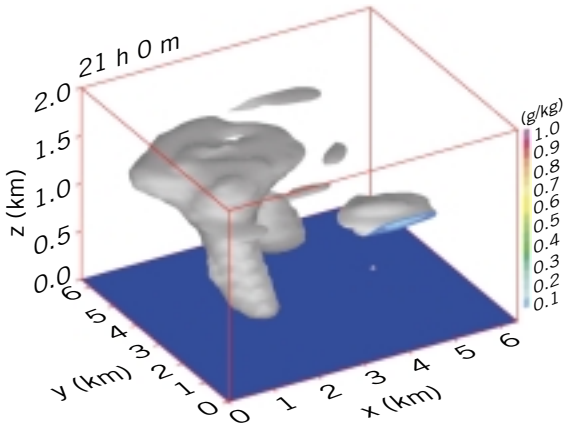
Owen Toon and Peter Colarco, Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder
Andrew Ackerman, NASA Ames Research Center

RESEARCH OBJECTIVES

The technical objective is to build the most advanced explicit aerosol and cloud microphysics model. The scientific objectives are (1) to explore the factors that determine cloud fraction in tradewind cumulus, (2) to use the explicit microphysical model to evaluate the precipitation parameterizations used in simpler models, (3) to investigate the relationship between cloud optical thickness and sea surface temperatures, and (4) to investigate the evolution of the aerosol size distribution and its effect on the atmosphere's radiative budget.

COMPUTATIONAL APPROACH

For the dynamics, we solve the anelastic form of the Navier-Stokes equations in flux form on a rectangular grid using finite difference techniques. The advection algorithm is second-order accurate for smooth flows. Sub-grid scale mixing is solved through either a first-order closure or a turbulent kinetic energy model. For the aerosol and cloud microphysics code, the particle size distributions are resolved on a fixed mass grid with size bins of geometrically increasing width. Condensational growth is treated using the piecewise polynomial method for advection in mass space. For the plane-parallel radiative transfer model, particle scattering and absorption coefficients are calculated using Mie theory. Gaseous absorption and emission are treated using an exponential sum formulation. For the mineral dust simulations, we incorporated the NCAR Model for Atmospheric Transport and Chemistry (MATCH) into the aerosol code.



Isosurface of cloud water depicting the outer boundary of a trade cumulus cloud in simulations for studying the effects of aerosol pollution on clouds and hence global climate. The colored contours denote the liquid water content where the cloud intersects the edge of the model domain.

ACCOMPLISHMENTS

In our simulations of tradewind cumulus, we found that cloud coverage depends strongly on the sub-grid scale (sgs) mixing. Previously we assumed the sgs mixing length scale was fixed by the model resolution, but we obtained realistic results when the sgs mixing length decreased with increasing atmospheric stability if precipitation was included.

In February and March 1999, a dark, murky haze was observed throughout the Indian Ocean, and there were very few clouds. We used our large eddy simulation model to investigate the possibility that solar heating due to the carbonaceous haze evaporated the clouds. Our results indicate that a heating rate at noon of only 1° per day is enough to significantly reduce the cloud coverage, contradicting the widespread assumption that increased aerosol loadings result in more clouds that reflect solar energy back into space.

We also applied the model to a case study of desert dust transport during the ACE-2 experiment. These simulations showed good agreement with observed vertical profiles of dust optical depth and particle sizes, and captured the general radiative properties of the aerosol layer.

SIGNIFICANCE

The indirect effect of aerosols on the radiative heat budget (through their impact on cloud properties) represents a leading uncertainty in the effects of mankind on the global climate. On the coarse grids of general circulation models, the representations of clouds (and their interactions with aerosols) are necessarily crude. Our work aims to advance understanding of cloud processes so that they may be more realistically treated in large-scale models.

PUBLICATIONS

A. S. Ackerman, O. B. Toon, D. E. Stevens, A. J. Heymsfield, V. Ramanathan, and E. J. Welton, "Reduction of tropical cloudiness by soot," Science **288**, 1042 (2000).

P. R. Colarco and O. B. Toon, "Three dimensional model simulations of the desert aerosol lifecycle: Comparisons to ground and satellite observations during ACE2-2," ICTP Summer Conference on Chemistry-Climate Interactions, Trieste, Italy, June 2000.

J. P. Taylor and A. S. Ackerman, "A case study of pronounced perturbations to cloud properties and boundary layer dynamics due to aerosol emissions," Q. J. Roy. Met. Soc. **125**, 2643 (1999).

http://sky.arc.nasa.gov/~ack

Coupled Parallel Climate Model (PCM) Applications to Climate Change Prediction for the IPCC and the National Assessment: High Resolution Studies

Warren Washington and Gerald Meehl,
National Center for Atmospheric Research
Albert Semtner, Naval Postgraduate School
John Weatherly, U.S. Army Cold Regions Research and Engineering Laboratory

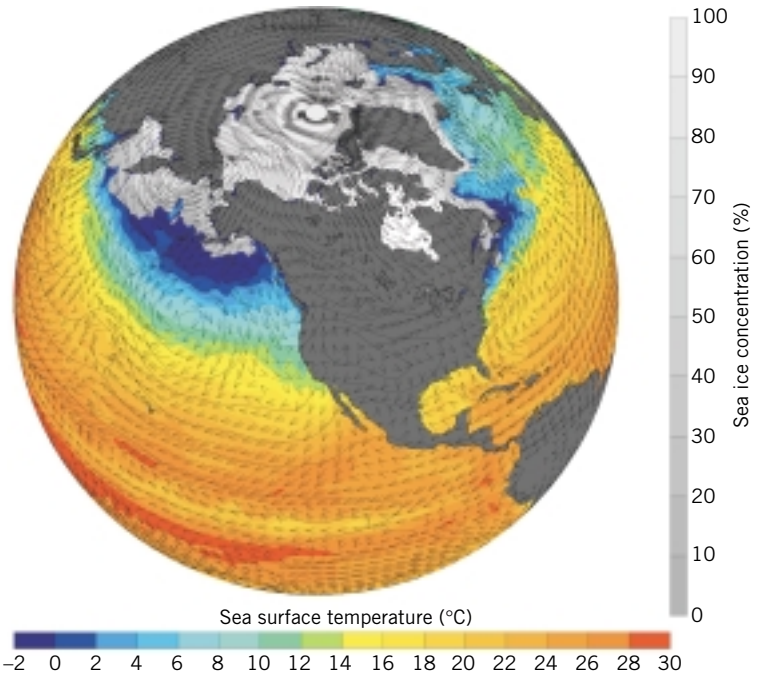
RESEARCH OBJECTIVES

The main purpose of this research is to use the Parallel Climate Model (PCM) for studies of anthropogenically forced climate change simulations with higher resolution and more detailed model components. Because it is difficult to separate anthropogenic climate change from natural climate variability, it is necessary to carry out ensembles of simulations in order to statistically find the climate change signal.

COMPUTATIONAL APPROACH

Ocean model component: Through collaboration among LANL, NPS, and NCAR, we have developed an ocean component that uses the finite difference Parallel Ocean Program (POP) with a displaced north pole. This model was modified from the original average resolution of 2/3° latitude and longitude to allow increased latitudinal resolution near the equator of approximately 1/2°. Also, because of the displaced pole, there is high horizontal resolution in the eastern North Pacific, in the Arctic Straits near northern Canada and Greenland, and the Gulf Stream area. The continents and bottom topography were carefully modified so as to obtain the correct volume transport flow in many regions through the globe. The model in its present form yields an extraordinary simulation of the Arctic Ocean, tropical Pacific, and boundary currents such as the Gulf Steam, with eddies. POP has recently been reformulated in terms of data structures and the use of memory and cache for enhanced performance in a message passing environment. Also, many additional model physics options have been added to improve the model's fidelity.

Sea ice model component: This model component is entirely new. The thermodynamic part of the model uses the physics from the C. Bitz University of Washington ice model, which allows an unusually high level of detail, including five or more ice thickness categories and elaborate surface treatment of snow and sea ice melt physics. Recently, an option for using elastic-viscous-plastic physics has been added, using the E. Hunke and J. Dukowicz approach to the solution of the ice dynamics.



This figure depicts a short simulation from the Parallel Climate Model (PCM). For the atmosphere, the figure shows vectors depicting the winds in the lowest model layer, and shows the sea level pressure as lines of constant pressure. The surface temperatures are shown in color, and the sea ice is shown in grayscale.

Atmospheric model component: The atmospheric component is the massively parallel version of the NCAR Community Climate Model version 3 (CCM3). This state-of-the-art model includes the latest versions of radiation, boundary physics, and precipitation physics. We use a T42 version (approximately 2.5°) for most of the simulations and will run a single simulation at T85 (approximately 1.25°) — the highest resolution climate change simulation ever attempted.

Flux Coupler: The method of tying the components together and allowing the exchange of fluxes and variables is the flux coupler. Since the component grids are different, there is an interpolation scheme for passing information between the atmosphere component grid and the ocean/sea ice grid that has been developed by P. Jones of LANL. The interpolation algorithm has been designed to run efficiently on distributed memory architectures. Recently, Chris Ding and collaborators at NERSC have improved the performance of the flux coupler on parallel computers.

Higher resolution atmosphere, ocean, and sea-ice models give more realistic simulations, but the computer time required for climate change scenarios prohibits conducting many experiments at higher resolution. We anticipate that the component interfacing through the flux coupler will be sufficiently flexible to handle a large range of resolutions. Therefore, the ongoing DOE-supported high resolution models can use the same structure as lower resolution models.

ACCOMPLISHMENTS

Over the past year we have accomplished a large ensemble set of simulations, showing the global climate changes due to increased greenhouse gases and changes in sulfate aerosols, for the years 1870–2100. We believe this is the largest set of climate change simulations with state-of-the-art models in the United States. We have continued to develop improved versions of the coupled model that include better sea ice and ocean components and a river transport component. Because of the importance of regional aspects of climate change, we have developed a higher resolution atmosphere model component that has a better definition of the continents-ocean boundaries as well as an improved treatment of mountains.

The archive of PCM simulation data sets has been made available to a wide community, including the Intergovernmental Panel on Climate Change (IPCC), the U.S. National Assessment of the Potential Consequences of Climate Variability and Change, and various model intercomparison projects, including the DOE Program for Climate Model Diagnosis and Intercomparison (PCMDI).

SIGNIFICANCE

The DOE Climate Change Prediction Program is focused on developing, testing and applying climate simulation and prediction models that stay at the leading edge of scientific knowledge and computational technology. The intent is to increase dramatically both the accuracy and throughput of computer model-based predictions of future climate system response to the increased concentrations of greenhouse gases on decadal and longer time scales. In the Kyoto protocol, there are several climate change scenarios with emissions of greenhouse gases and sulfate aerosols that must be completed for the U.S. National Assessment.

PUBLICATIONS

A. J. Semtner, "Ocean and climate modeling on advanced parallel computers: Progress and prospects," Communications of the ACM (in press).

W. M. Washington, J. W. Weatherly, G. A. Meehl, A. J. Semtner, T. W. Bettge, A. P. Craig, W. G. Strand, J. Arblaster, V. B. Wayland, R. James, and Y. Zhang, "Parallel Climate Model (PCM) control and transient simulations," Climate Dynamics (in press).

http://www.cgd.ucar.edu/pcm/

Numerical Tokamak Turbulence Project

B. Cohen, A. Dimits, G. Kerbel, D. Shumaker, and W. Nevins, Lawrence Livermore National Laboratory
W. Lee, M. Beer, G. Hammett, and Z. Lin, Princeton Plasma Physics Laboratory
J.-N. Leboeuf and R. Sydora, University of California, Los Angeles
V. Lynch, Oak Ridge National Laboratory
Y. Chen, S. Parker, and C. Kim, University of Colorado
P. Snyder, R. Waltz, Y. Omelchenko, and J. Candy, General Atomics
J. Cummings, Los Alamos National Laboratory
W. Dorland and S. Novakovski, University of Maryland
D. Ross, University of Texas

RESEARCH OBJECTIVES

The primary research objective of the Numerical Tokamak Turbulence Project (NTTP) is to develop a predictive ability in modeling turbulent transport due to drift-type instabilities in the core of tokamak fusion experiments, through the use of three-dimensional kinetic and fluid simulations and the derivation of reduced models.

COMPUTATIONAL APPROACH

We are utilizing three main classes of simulation algorithms to study core tokamak microturbulence: gyrokinetic particle-in-cell (GK PIC), 5D Eulerian gyrokinetic (EGK), and gyro-Landau-fluid (GLF). In each case, the simulation domain can be either global or annular (flux tube). (1) The GK PIC simulations are based on PIC methods for the self-consistent solution of Poisson's equation (or Maxwell + Poisson in electromagnetic extensions) and plasma equations of motion, and domain decomposition methods to run efficiently in parallel. (2) The EGK algorithm solves for the 5D distribution function and Maxwell's equations on a mesh that includes two velocity space coordinates (energy and magnetic moment). Four of five dimensions are distributed among PEs for efficient parallel operation. The linear terms are treated implicitly. (3) The GLF algorithm is most similar to conventional fluid dynamics approaches, since a set of fluid moments of the gyrokinetic equation are solved together with Maxwell's equations. Typically, four to six velocity space moments per plasma species are evolved explicitly.

ACCOMPLISHMENTS

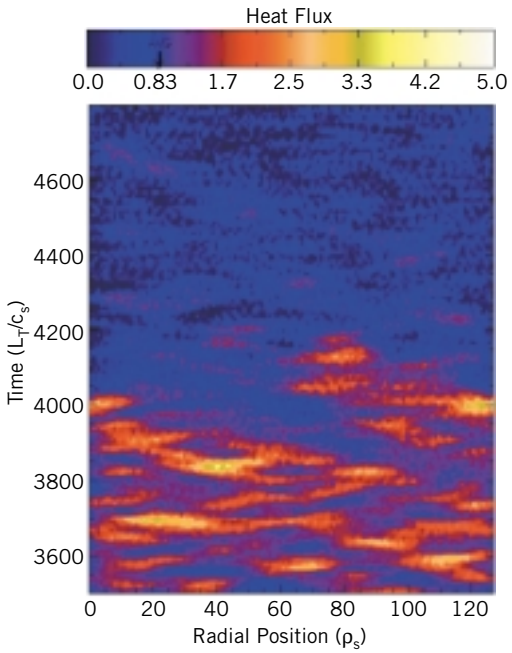
In the last year, we have reached several important milestones:

We completed a lengthy exercise benchmarking turbulence simulations from several different codes.

We completed: (1) A much more extensive systematic set of nonlinear gyrokinetic simulation parameter scans of ITG-driven transport than before, exploring parts of the physical parameter space new for such nonlinear studies as well as revealing new dependences in regions that had been explored before but not as systematically. (2) Nonlinear investigations of the effect of radial velocity shear on ITG-driven transport, leading to the discovery of significant new parameter dependences that were hidden in previous nonlinear studies and in widely used transport models. (3) A study in which the transport reduction by radial profile gradient variation was discovered to be highly sensitive to the various terms in the equilibrium radial force balance. (4) A characterization of the event-size dependence of the thermal flux carried by transport events of a given size in ITG-driven turbulence. This provides a key qualitative constraint on any theory that claims to describe ITG-driven turbulence.

We elucidated the importance of zonal flows in ITG-driven turbulence, particularly near marginal stability, and particularly in the weakly collisional limit that is appropriate for the fusion program.

We presented the first toroidal electromagnetic simulations of tokamak microturbulence at invited talks at three major scientific



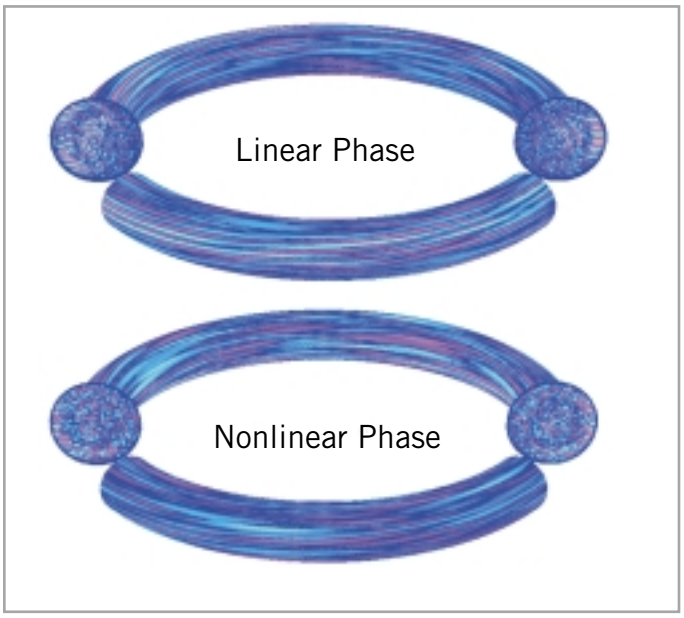
This figure shows the radial heat flux from the pg3eq 3D PIC code as a function of the radial coordinate, x (in units of the ion gyroradius) and time, t (in units of the thermal transit time). There is a transition at about t = 4100 from low-confinement regime to a high-confinement regime. The low confinement regime (t < 4000) is characterized by large heat pulses excited by plasma microturbulence which propagate both up (toward smaller values of x) and down (toward larger values of x) the ambient temperature gradient. These large heat pulses are absent in the high confinement regime (t > 4200).

meetings, and in the literature. We also discovered that electron scale turbulence (ETG modes) can, under some conditions, cause transport comparable to that resulting from ITG modes.

We discovered a collisionless instability which regulates zonal flows in ITG turbulence.

We presented detailed simulations of plasmas in experimental devices including UCLA's Electric Tokamak, TEXTOR-94, DIII-D, and Alcator C-Mod.

We have begun to exercise our 3D toroidal hybrid model having gyrokinetic ions and drift fluid electrons. We have been exploring Alfvénic ITG-driven turbulence at moderate plasma β with this code. In addition, we have recently developed a drift-kinetic electron model using the canonical parallel momentum formulation and the split-weight scheme. We have been using this code to study kinetic electron effects and electron transport at low plasma β. Finally, we have been actively developing a PIC method compatible with finite elements and unstructured grids that can be used for adding a kinetic pressure term to the nonlinear MHD code NIMROD. We have developed a new parallel algorithm for PIC which we call "domain cloning," which augments a 1D domain decomposition without the complications of a second dimension. Multiple copies of the sub-domains are spread across processors, allowing for a much larger number of particles and many more processors than grid cells in the domain-decomposed direction.



3D rendering of potential fluctuations from global gyrokinetic particle calculations of ITG-driven turbulence in a large aspect ratio tokamak plasma without and with externally imposed localized sheared toroidal flow.

SIGNIFICANCE

Experiments have shown that control of drift-type modes in tokamak fusion experiments leads to major improvements in plasma energy confinement and, hence, fusion conditions. NTTP simulations are leading to a deeper understanding of anomalous transport in current experiments. Since controlling the energy transport has significant leverage on the performance, size, and cost of fusion experiments, reliable simulations can lead to significant cost savings and improved performance in future experiments.

PUBLICATIONS

A. M. Dimits, M. A. Beer, G. W. Hammett, C. Kim, S. E. Parker, D. E. Shumaker, R. Sydora, A. J. Redd, J. Weiland, M. T. Kotschenreuther, W. M. Nevins, G. Bateman, C. Bolton, B. I. Cohen, W. D. Dorland, A. H. Kritz, J. E. Kinsey, L. L. Lao, and J. Mandrekas, "Comparisons and physics basis of tokamak transport models and turbulence simulations," Phys. Plasmas **7**, 969 (1999). http://stripe.colorado.edu/~sparker/andris.ps

F. Jenko, W. Dorland, M. Kotschenreuther, and B. N. Rogers, "Electron temperature gradient driven turbulence," Phys. Plasmas **7**, 1904 (2000).

Z. Lin, T. S. Hahm, W. W. Lee, W. M. Tang, and R. B. White, "Gyrokinetic simulations in general geometry and applications to collisional damping of zonal flows," Phys. Plasmas **7**, 1857 (2000).

http://www.acl.lanl.gov/GrandChal/Tok/tokamak.html
http://www.acl.lanl.gov/GrandChal/Tok/BlueBook.html
http://www.er.doe.gov/production/cyclone/

LLNL Magnetic Fusion Energy Supercomputing

Bruce Cohen, Maria Caturla, Harry McClean, Andris Dimits, Thomas Rognlien, Xueqiao Xu, Alice Koniges, Gary Kerbel, Dan Shumaker, Tomas Diaz Delarubia, Lynda LoDestro, Marvin Rensink, Brian Wirth, Babak Sadigh, Simon Woodruff, and Susanne Ramsey, Lawrence Livermore National Laboratory

RESEARCH OBJECTIVES

(1) Edge plasmas: Continued development of simulation tools to model the boundary region of magnetic fusion energy (MFE) devices between the hot core plasma and material walls. (2) Core transport: Gyrokinetic simulation of drift-wave driven transport of energy and plasma across magnetic field lines. (3) Neutron interactive materials: Simulations to determine the effect of irradiation on materials properties. (4) Simulation of spheromak plasmas: Extend previous simulations of spheromak plasmas to study formation, stability, and field line closure.

COMPUTATIONAL APPROACH

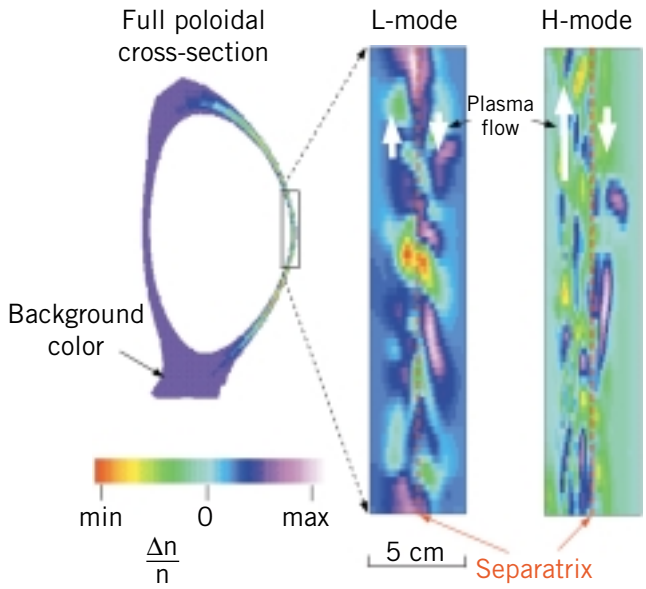
(1) A 2D fluid transport code (UEDGE) includes multispecies impurities to calculate edge-plasma profiles, and a 3D fluid turbulence code (BOUT) predicts the anomalous turbulence-induced transport of particles and energy across the confining magnetic field in the edge region. (2) In five-phase-space dimensional particle simulations of plasma, the Vlasov-Maxwell or Vlasov-Poisson systems of equations are solved on a grid in configuration space and with a Monte Carlo sampling technique in velocity space. (3) A molecular dynamics (MD) code (MDCASK) uses empirical, alloy interatomic potentials to simulate the evolution of displacement cascades, fundamental kinetics properties of defects in metals, and the interactions between defects.

ACCOMPLISHMENTS

(1) Substantial progress has been made in understanding edge turbulence and its role in formation of the edge H-mode transport barrier. BOUT shows that X-point damping, coupled with the change in boundary flow direction, plays a crucial role in determination of the fluctuation levels.

(2) Gyrokinetic results are described in the NTTP abstract (previous page).

(3) Simulations of the evolution of recoils in Cu and W for energies from a few eV to 100 keV show damage in the form of vacancy and interstitial clusters in Cu, but few vacancy clusters in W. Large-scale MD simulations to study irradiation-produced defects in face-centered cubic metals show that a key mechanism in the formation of defect-free dislocation channels is the absorption of stacking fault tetrahedra by moving dislocations.



3D BOUT edge plasma simulations show large density fluctuations on the outboard side for L-mode |Δn/n|~0.2 and for H-mode |Δn/n|~0.05. H-mode structures are broken up and their amplitudes are suppressed by flow shear.

(4) A series of simulations indicate how the gun potential, the plasma resistivity, and the dimensions of the gun and flux conserver influence formation and magnetic flux buildup in the spheromak plasma.

SIGNIFICANCE

Edge plasmas and core turbulence are both critical issues for improved performance of MFE devices. The application of fusion as a viable energy source also depends on developing structural materials that can withstand the harsh radiation conditions of the fusion environment without experiencing severe degradation.

PUBLICATIONS

A. M. Dimits et al., "Comparisons and physics basis of tokamak transport models and turbulence simulations," Phys. Plasmas **7**, 969 (1999).

M.-J. Caturla, N. Soneda, E. Alonso, B. D. Wirth, T. Diaz de la Rubia, and M. Perlado, "Comparative study of radiation damage accumulation in Cu and Fe," J. Nucl. Mat. **276**, 13 (2000).

X. Q. Xu, R. H. Cohen, T. D. Rognlien, and J. R. Myra, "L-H transitions simulations in divertor geometry," Phys. Plasmas **7**, 1951 (2000).

Simulation of Intense Beams for Heavy-Ion Fusion

Alex Friedman, William M. Sharp, and David P. Grote,
Lawrence Berkeley National Laboratory

RESEARCH OBJECTIVES

To model beam dynamics for heavy-ion fusion (HIF) in accelerators and in fusion chambers.

COMPUTATIONAL APPROACH

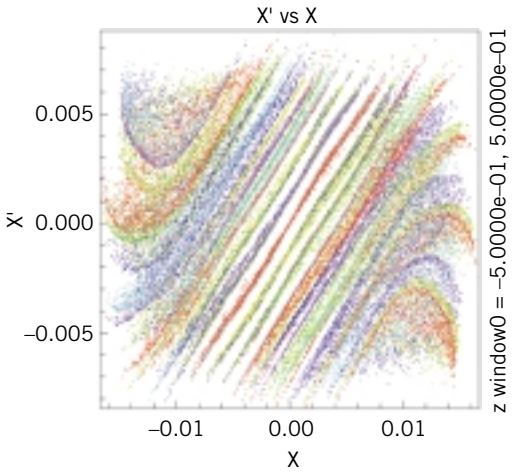
A hierarchy of codes is used for numerical simulation of beam dynamics in heavy-ion accelerators. IBEAM is first used for global optimization. A truncated-moment beam module in WARP is then used for refining the physics design and for generating the fields for beam acceleration, compression, and longitudinal control. Detailed accelerator simulations are done mainly with the electrostatic particle-in-cell (PIC) modules of WARP. A 2D semi-Lagrangian Vlasov code is being developed to study halo questions and to corroborate the PIC modeling. Beam transport modeling in the fusion chamber has been done with the axisymmetric PIC code BICrz and with the more modern 2D/3D code BPIC. Mission Research Corporation will use their parallel 2D/3D electromagnetic PIC code LSP to investigate chamber-transport scenarios using higher plasma densities.

ACCOMPLISHMENTS

Much of the simulation work has focused on developing the physics design of the Integrated Research Experiment (IRE). WARP simulations of several representative layouts were made to quantify such phenomena as emittance growth, beam halo formation, and transverse-longitudinal coupling. Simulations examining the effects of quadrupole offsets and idealized steering were also carried out. Initial steps have been taken to design and simulate a drift compression system.

Simulations were also used to design and support other experiments. For beam injectors, extensive simulation was carried out to determine the sensitivity of the 2 MeV injector to changes in geometry, voltage, and source uniformity. A multiple-beamlet injector has been designed that promises a higher average current density than a single large-aperture beam injector. Simulations were also done for the High-Current Experiment (HCX), the Scaled Final-Focus Experiment, the University of Maryland electron ring experiment, and the European approach to heavy-ion fusion.

Finally, simulations were used to study several basic questions of accelerator physics, including beam splitting and halos. A preliminary investigation of impedance effects on longitudinal instability found that a coasting beam under some conditions remains stable through hundreds of accelerating modules.



3D WARP simulation of a novel merging-beamlet ion beam injector. On the order of 100 beamlets are independently accelerated to 1.2 MeV and then merged to form a single beam with a current of 0.5 A. The transverse phase space, (*x, vx/vz*) projection at 1.6 m after the end of the accelerating Pierce column, shows the beamlets as they begin to mix. The particles are color coded based on the beamlet they began in.

SIGNIFICANCE

The U.S. HIF program is charged with developing an accelerator and focusing system capable of igniting inertial-fusion targets at a cost that is competitive with other long-term energy sources. Achieving this objective requires significant advances in beam physics and in accelerator science and technology. Numerical simulations are essential for all aspects of the HIF program, including interpreting the results of ongoing experiments and developing and optimizing the designs of future experiments.

PUBLICATIONS

A. Friedman, D. P. Grote, E. P. Lee, and E. Sonnendrucker, "Beam simulations for IRE and driver: Status and strategy," 13th International Heavy-Ion Fusion Symposium, 12–17 March 2000, San Diego, CA; Nuclear Instruments and Methods in Physics Research (in press).

D. P. Grote, A. Friedman, G. D. Craig, W. M. Sharp, and I. Haber, "Progress toward source-to-target simulations," 13th International Heavy-Ion Fusion Symposium, 12–17 March 2000, San Diego, CA; Nuclear Instruments and Methods in Physics Research (in press).

E. Sonnendrucker, A. Friedman, J. J. Barnard, D. P. Grote, and S. M. Lund, "Simulation of heavy ion beams with a semi-Lagrangian Vlasov solver," 13th International Heavy-Ion Fusion Symposium, 12–17 March 2000, San Diego, CA; Nuclear Instruments and Methods in Physics Research (in press).

http://fusion.lbl.gov/US_HIF.html

University of Maryland Fusion Energy Research

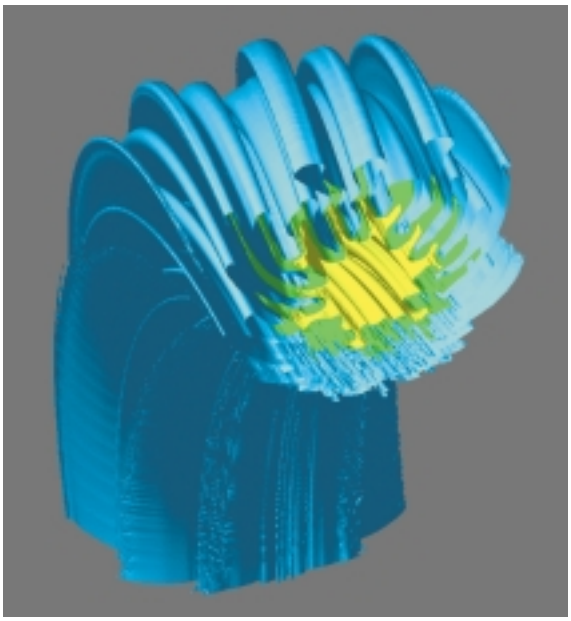
Parvez Guzdar, Bill Dorland, James Drake, Adil Hassam, Robert Kleva, and Barrett Rogers, University of Maryland

RESEARCH OBJECTIVES

The Maryland Theory and Computational Physics Magnetic Fusion Energy Program focuses on (1) 3D simulation of particle, ion, and electron energy transport in the core and edge region of tokamak plasmas using two-fluid Braginskii equations, gyrofluid equations, and Vlasov codes; (2) 3D simulation of high-β disruptions, sawtooth crashes, and pellet injection for tokamak plasmas using a toroidal resistive magnetohydrodynamics (MHD) and two-fluid code; (3) 2D and 3D simulations of novel centrifugal confinement devices using MHD codes; (4) 2D and 3D hybrid simulations of collisionless reconnection; and (5) 3D gyrokinetic simulations of the levitated magnetic dipole experiment (LDX) at the Massachusetts Institute of Technology.

COMPUTATIONAL APPROACH

For all the two-fluid Braginskii and MHD codes for the studies listed above, the basic codes solve a coupled system of convection-diffusion equations. For these codes we use a leapfrog trapezoidal algorithm for the time stepping and a fourth-order up-wind finite differencing scheme for the spatial convective derivatives.



The figure shows three isosurfaces of the pressure from a 3D resistive MHD simulation of high β.

The second-order accurate gyrokinetic algorithm we use is comprised of an implicit treatment of the linear dynamics; an explicit, pseudo-spectral treatment of the nonlinear terms; and an Adams-Bashforth integrator in time. Parallelization is accomplished with MPI and SHMEM libraries. The gyrokinetic problem involves the usual 3D spatial grid, as well as a 2D velocity space grid, for a total of five dimensions. We divide four of the dimensions at a time over processors to achieve a high degree of parallelization with good load balancing.

Our 3D gyrofluid code shares all communication and pseudo-spectral evaluation modules with the gyrokinetic code, as well as the design philosophy. Two of the spatial domains are spread among processors at any given time.

ACCOMPLISHMENTS

During the last year we made progress on three areas which required large-scale computation: (1) nonlinear simulations of η_i modes to understand electron energy transport, (2) hybrid simulation for the study of collisionless reconnection, and (3) high-resolution simulations of high β disruptions of tokamaks.

SIGNIFICANCE

These studies focus on the most important problems for understanding the issues of anomalous confinement and MHD, and collisionless disruption and reconnection processes that limit the parameter space of stable operation of tokamak devices. The codes that we have developed may also be used to study the confinement and stability properties of non-tokamak devices, such as the levitated dipole and centrifugal confinement configurations. The physics that we learn from these studies has motivated the study of a novel centrifugal confinement device which incorporates the positive features of shear flow suppression of microinstabilities and the associated reduction in anomalous transport.

PUBLICATIONS

M. Shay, J. Drake, B. Rogers, and R. Denton, "The scaling of collisionless, magnetic reconnection for large systems," Geophys. Res. Lett. **26**, 2163 (1999).

A. Zeiler, J. Drake, and B. Rogers, "Magnetic reconnection in toroidal η_i mode turbulence," Phys. Rev. Lett. **84**, 99 (2000).

R. G. Kleva and P. N. Guzdar, "Nonlinear stability limit in high β tokamaks," Phys. Plasmas **7**, 1163 (2000).

http://www.ipr.umd.edu/Theory/research.htm

3D Modeling of Fusion Plasmas

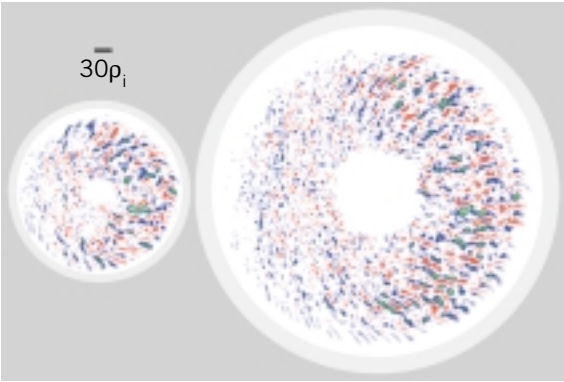
Stephen Jardin, W. W. Lee, Z. Lin, D. Stotler, D. Mikkelsen, W. Park, M. A. Beer, and G. W. Hammett,
Princeton Plasma Physics Laboratory

RESEARCH OBJECTIVES

(1) Improving the existing 3D gyrokinetic particle code to study neoclassical and turbulent transport in tokamaks and stellarators, as well as to investigate hot-particle physics, toroidal Alfvén modes, and neoclassical tearing modes. (2) Enhancing the existing 3D beam equilibrium, stability, and transport (BEST) code, used for two-stream and filamentation instability studies for space-charge dominated beams in accelerators and chamber transport in heavy ion fusion research. (3) Developing a new δf simulation model with application to studies of laser-plasma interaction physics in the areas of turbulence and transport. (4) Study of tokamak and stellarator plasma turbulence with a 3D flux-tube gyrofluid code (GRYFFIN), which has recently been extended to include equilibrium-scale sheared E × B flows. (5) Coupling the DEGAS 2 Monte Carlo neutral transport code and the UEDGE fluid plasma transport code to analyze experimental results and predict scrape-off layer conditions in future devices.

COMPUTATIONAL APPROACH

(1 and 2) The particle-in-cell method is used for transport studies. (3) A 2D δf ion + fluid electron code has been developed to study laser-driven acoustic turbulence. A 1D Monte-Carlo code for studying nonclassical drive and transport of electrons in laser-plasma instabilities is being developed. (4) The gyrofluid code is a 3D nonlinear pseudo-spectral code, using both spectral and grid representations with fast Fourier transforms between the two representations. (5) The DEGAS 2 code has been parallelized and demonstrates excellent scaling. While UEDGE has been parallelized, some testing and possible improvements remain. Since the two codes parallelize differently, schemes for parallelizing the coupled code system must be established and evaluated.



Results from large-scale full torus gyrokinetic particle simulations of plasma microturbulence in tokamak with device-size scans for realistic parameters show that Bohm-like transport can be driven by microscopic-scale fluctuations with isotropic spectra. These simulation results resolve some apparent physics contradictions between experimental observations and turbulent transport theories.

ACCOMPLISHMENTS

Using the GTC code, we demonstrated the importance of nonlinearly generated zonal flow for the reduction of ion thermal transport, as well as the role played by ion-ion collisions for the bursting behavior observed in the tokamak experiments. We used the BEST code to study two-stream instabilities in space-charge-dominated beams in accelerators, helping to explain the beam loss observed in various machines. We demonstrated numerically the existence of non-KV equilibria in a periodic focusing lattice. We devised an efficient numerical particle scheme for treating electrons; it enables us to circumvent the parallel Courant condition.

We extended the nonlinear gyrofluid code to include equilibrium-scale-sheared E × B flows, with a coordinate system that shears in time to follow the flow. The nonlinear effects of this equilibrium-sheared flow can be significantly different than the linear effects. Our gyrokinetic continuum simulations show that electron-temperature-gradient turbulence can be much larger than suggested by naive scaling from ion-temperature-gradient turbulence because of the different adiabatic species response. We used the DEGAS 2 code to interpret CMOD data and have completed the initial coupling of the UEDGE and DEGAS 2 codes.

SIGNIFICANCE

Three-dimensional modeling helps build a fundamental understanding of plasma turbulence processes and turbulence suppression methods, improves the analysis of experimental results, and suggests new ways to improve magnetic and heavy ion fusion reactor designs.

PUBLICATIONS

Z. Lin, T. S. Hahm, W. W. Lee, W. M. Tang, and R. B. White, "Gyrokinetic simulations in general geometry and applications to collisional damping of zonal flows," Phys. Plasmas **7**, 1857 (2000).

Z. Lin, M. S. Chance, T. S. Hahm, J. A. Krommes, W. W. Lee, I. Manuilskiy, H. E. Mynick, H. Qin, G. Rewoldt, W. M. Tang, and R. B. White, "Numerical and theoretical studies of turbulence and transport with E × B shear flows," Nuclear Fusion **40**, 737 (2000).

I. Manuilskiy and W. W. Lee, "The split-weight particle simulation scheme for plasmas," Phys. Plasmas **7**, 1381 (2000).

http://w3.pppl.gov/theory/

Magnetohydrodynamic Codes for Fusion Applications

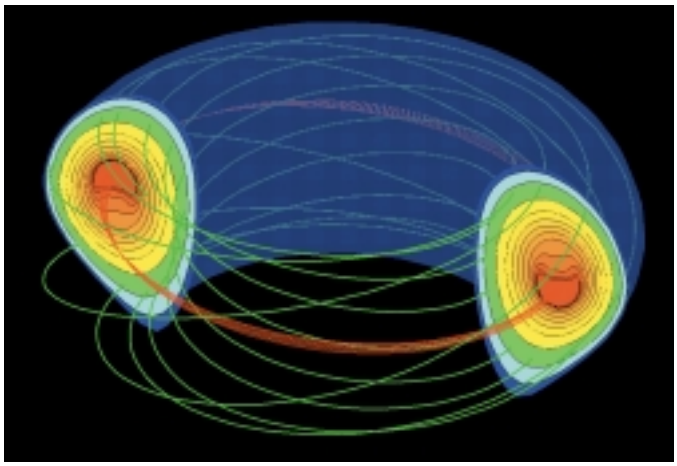
Dalton Schnack and Scott Kruger,
Science Applications International Corp.

RESEARCH OBJECTIVES

Our research objective is to provide computational tools and support for the study of macroscopic instabilities in the magnetic confinement fusion community, especially the DIII-D Tokamak. All of our codes use the (extended) magnetohydrodynamic (MHD) equations, and are nonlinear, initial-value codes.

COMPUTATIONAL APPROACH

We use NERSC resources for several of our MHD codes, but the primary one is NIMROD, a relatively new code built to take advantage of new computer architectures. It uses a combined finite element/Fourier series spatial representation with a time-split, semi-implicit advance. The semi-implicit time advance requires the inversion of matrices which are extremely ill-conditioned due to the anisotropy caused by the magnetic field and the disparate time scales of the instabilities we wish to study (Alfvén wave time scale much shorter than instability growth time). The matrices are inverted using either a NIMROD-developed conjugate gradient solver or the AZTEC software package. NIMROD is an extremely sophisticated code that works in axisymmetric geometries and for problems requiring the extended MHD equations (MHD + 2-fluid + advanced closures).



Numerical simulation of the nonlinear MHD evolution of shot 86144 in the DIII-D Tokamak at General Atomics, San Diego. Deformed pressure surfaces and magnetic field line trajectories are shown. The simulation used a realistic value of the plasma resistivity and was performed with the NIMROD code on the NERSC T3E.

ACCOMPLISHMENTS

Many numerical problems were discovered in trying to simulate a high-β disruption of the DIII-D tokamak. Most of the problems were found to be due to the preprocessing of the DIII-D data. Now that the problems are solved, work is under way to compare theory to experiment.

A tearing mode unstable case was identified in a simple (cylindrical) geometry, and efforts were made to run this case as high in magnetic Lundquist number as possible. Linearly, NIMROD was able to achieve converged solutions at a Lundquist number of 10^9 due to grid packing. Nonlinearly, converged solutions were achieved at 10^7. Efforts are now under way to extend this type of parameter pushing to realistic DIII-D equilibria.

SIGNIFICANCE

The codes described here are designed to do nonlinear, initial-value simulations of long-wavelength phenomena in fusion-relevant plasmas. These types of motions severely constrain the operating regime of fusion experiments. By developing and applying powerful computational tools to the study of these problems, our understanding of these operational limits will increase, leading to better design and operation of fusion experiments. Our primary focus is support of the DIII-D Tokamak, the largest fusion experiment in the U.S. program.

PUBLICATIONS

Zoran Mikić, Jon A. Linker, Dalton D. Schnack, Roberto Lionello, and Alfonso Tarditi, "Magnetohydrodynamic modeling of the global solar corona," Phys. Plasmas **6**, 2217 (1999).

A. H. Glasser, C. R. Sovinec, R. A. Nebel, T. A. Gianakon, S. J. Plimpton, M. S. Chu, D. D. Schnack, and the NIMROD Team, "The NIMROD code: A new approach to numerical plasma physics," Plasma Phys. and Control. Fus. **41**, A747 (1999).

http://haven.saic.com/

Plasma Confinement, Stability, Heating and Optimization in Stellarators and Tokamaks

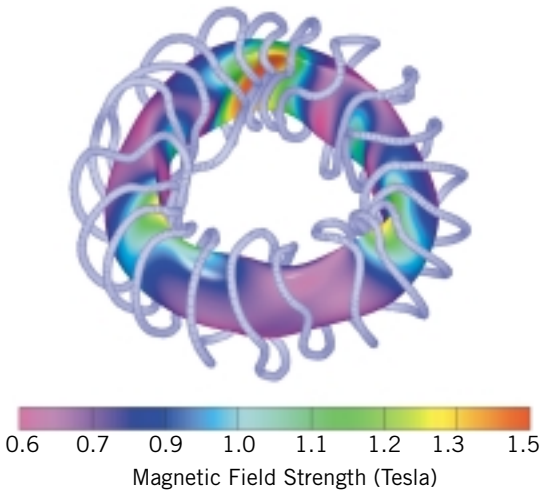
Don Spong, Vickie Lynch, Steven Hirshman, Ben Carreras, and Donald B. Batchelor, Oak Ridge National Laboratory

RESEARCH OBJECTIVES

The ORNL Fusion Theory Group is pursuing computational research in four areas: stellarator optimization and physics, stellarator transport and heating, plasma turbulence and its effects on transport, and radio frequency (rf) heating of plasmas.

COMPUTATIONAL APPROACH

Stellarator optimizations are carried out using a steepest-descent method to minimize a variational form for the 3D plasma equilibrium. The plasma optimization is then carried out with a Levenberg-Marquardt algorithm. We are using a 3D model of turbulence-induced transport to evaluate the role of avalanches. We have also used particle tracers to estimate the anomalous diffusion exponent associated with this model. Coupled partial differential equations for the ion density, parallel velocity, and temperature are evolved in time in the presence of a noise source in the temperature equation (to simulate heating). Finite differences in radius and Fourier expansions in the toroidal and poloidal angles are used. The time stepping scheme is time-implicit for the linear terms and time-explicit for the nonlinear terms. Rf heating uses combinations of Fourier and finite difference representations for the time-varying electric and magnetic fields used in plasma heating and current drive. Maxwell's equations coupled with various forms of the plasma dielectric tensor are then solved over the plasma volume.



Outer magnetic flux isosurface (color contours show magnetic field strength) and magnet coils (light blue) for a three-field-period, high β (23%), quasi-poloidal stellarator. This configuration has been numerically designed using physics-based stability and confinement optimization targets.

ACCOMPLISHMENTS

New stellarator physics optimization targets added during the past year include the self-consistent bootstrap current, ballooning stability, direct targeting of transport coefficients using the DKES code, and confinement improvement using the longitudinal adiabatic invariant. These targets have allowed us to design compact stellarators with higher ballooning stability limits (up to 23% stable β) and improved neoclassical confinement properties. The stellarator Monte Carlo code has been extensively used to analyze the transport physics of existing stellarator configurations and new designs generated by the stellarator optimization code. This code identified the greatly improved transport that can be achieved for the recently developed high β (23%) configurations.

Using a 3D model of turbulence-induced transport, we are evaluating the role of avalanches. We have also used particle tracers to estimate the anomalous diffusion exponent associated with this model. Preliminary results indicate that the tracer particle transport is closer to ballistic than diffusive.

The 2D spectral code for rf heating can now solve up to 120,000 dense equations with direct (noniterative) ScaLAPACK solutions and achieves up to 0.6 teraflop/sec performance. This code is being used to analyze high harmonic fast-wave propagation in the National Spherical Torus Experiment at Princeton Plasma Physics Laboratory.

SIGNIFICANCE

The development of new compact stellarators allows larger-volume plasmas to be designed at a fixed cost. Larger-volume plasmas are less edge-dominated, lose less energy from charge exchange, and as a result allow better science to be carried out. Compact plasmas could also lower development costs and allow smaller, more modular devices to be built. If successful, this could significantly improve the economics of fusion power. Improved understanding of transport and rf heating could lead to smaller, more reliable, and less costly fusion reactors.

PUBLICATIONS

B. A. Carreras, V. E. Lynch, and D. E. Newman, "Self-organized criticality and plasma fluctuation dynamics," J. Plasma Fus. Res. **2**, 8 (1999).

S. P. Hirshman, D. A. Spong, et al., "Physics of compact stellarators," Phys. Plasmas **6**, 1858 (1999).

E. F. Jaeger, L. A. Berry, and D. B. Batchelor, "Full-wave calculation of sheared poloidal flow driven by high-harmonic ion Bernstein waves in tokamak plasmas," Phys. Plasmas **7**, 3319 (2000).

http://www.ornl.gov/fed/theory/Theory_Home_page.html

Equilibrium, Stability, and Transport Studies of Toroidal Magnetic Confinement Systems

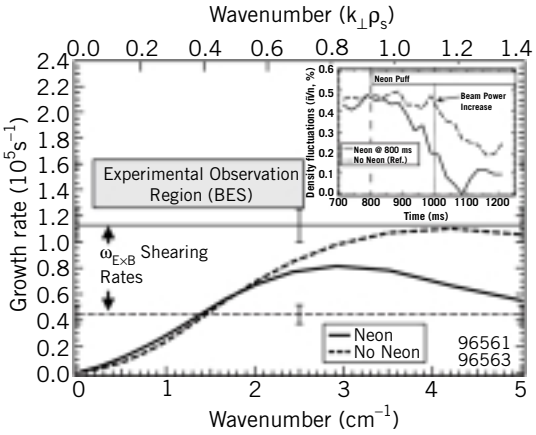
A. D. Turnbull, J. Candy, M. S. Chu, J. R. Ferron, L. L. Lao, Y. Omelchenko, P. B. Snyder, G. Staebler, R. E. Waltz, and the DIII-D team, General Atomics, Inc.
A. M. Garofalo, Columbia University
E. J. Kinsey, Lehigh University
W. Wang, University of California, Irvine

RESEARCH OBJECTIVES

The aim of this research is fourfold: (1) Provide support calculations for the DIII-D National Fusion Facility, including experimental predictions and analysis and interpretation of data. (2) Establish an improved theoretical and computational scientific basis for the physics of fusion plasmas. (3) Optimize presently known Advanced Tokamak (AT) configurations for high performance and identify potential new configurations. (4) Explore and optimize alternative magnetic confinement configurations, and elucidate the relationships between these and tokamak configurations.

COMPUTATIONAL APPROACH

The principal codes used are EFIT and TOQ (equilibrium); GATO, TWIST-R, MARS, and BALOO (MHD stability); GLF23, BALDUR, TRANSP, ONETWO, CORSICA, MCGO, and P2D (transport and fueling); CQL3D, CURRAY, and TORAY (current drive); and UEDGE and DEGAS (edge physics). New computational tools being developed and tested include linearized MHD stability codes (ELITE, TWIST-R) and the highly parallelized simulation codes GYRO, GRYFFIN, and FORTEC.



Calculations done at NERSC explain the observed large improvement in the confinement of tokamak discharges seeded with neon impurities. The figure shows a fully kinetic linear growth rate calculation for ITG modes over a range of wavenumbers, with reduced growth rates at the high end of the wavenumber spectrum for the discharge with neon. The inset shows the evolution of the measured electron density fluctuations; the discharge with injected neon impurities shows a marked drop in fluctuation level, indicating reduced turbulence.

ACCOMPLISHMENTS

Gyrokinetic growth rate calculations analyzing the drift-wave stability of a variety of tokamak plasmas found that discharges with neon injection had improved energy confinement due to the suppression of ion temperature gradient (ITG) mode turbulence. The growth rates were reduced both directly by the neon and by $E \times B$ shear, which was synergistically enhanced. These neon-injection discharges have significant potential as a new option for improved confinement in ATs.

New electromagnetic gyrofluid simulations of tokamak plasmas, which quantified the transition from electrostatic to electromagnetic turbulence with increasing β, call into question the validity of the electrostatic approximation commonly employed in turbulent transport studies. The new simulations found that microturbulence takes on an electromagnetic character even at low values of β, and that significant electromagnetic effects on turbulent transport occur.

A new working model of edge localized modes (ELMs), which have been observed but poorly understood for two decades, was developed and shown to describe the DIII-D ELM behavior well. Analysis of results from resistive wall mode (RWM) closed loop feedback experiments in DIII-D showed the first clear evidence that the n = 1 RWM can be controlled by an applied external magnetic field.

SIGNIFICANCE

Recent progress in fusion has been accelerated as a result of a strong coupling between theory, computation, and experiments. Previous calculations over the past decade identified several extremely promising AT configurations that are now the focus of the U.S. tokamak program. AT and alternative configurations need to be optimized further to guide future experiments.

PUBLICATIONS

G. M. Staebler, G. L. Jackson, W. P. West, S. L. Allen, R. J. Groebner, M. J. Schaffer, and D. G. Whyte, "Improved high-mode with neon injection in the DIII-D Tokamak," Phys. Rev. Lett. **82**, 1692 (1999).

J. Kinsey, R. E. Waltz, and J. C. DeBoo, "Perturbative tests of theoretical transport models using cold pulse and modulated electron cyclotron heating experiments," Phys. Plasmas **6**, 1865, (1999).

R. E. Waltz and R. L. Miller, "Ion temperature gradient turbulence and plasma shape," Phys. Plasmas **6**, 4265, (1999).

http://fusion.gat.com

Lattice Boltzmann Simulations for Divertor Physics and Turbulence

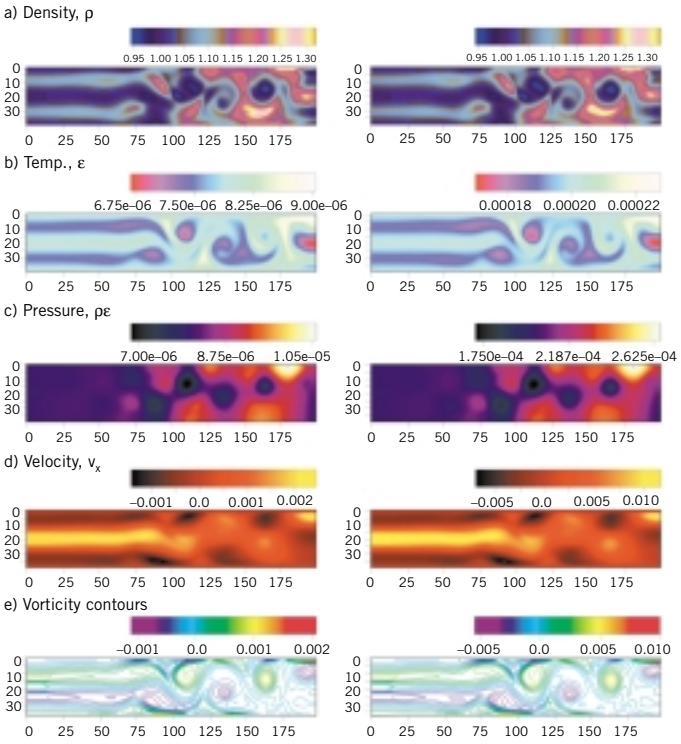
George Vahala, College of William and Mary
Linda Vahala, Old Dominion University
Pavol Pavlo, Institute of Plasma Physics, Czech Academy

RESEARCH OBJECTIVES

Thermal lattice Boltzmann modeling (TLBM) is an ideal MPP computational tool to study nonlinear macroscopic systems. In the divertor regime, where the neutral collisionality roams from very collisional (fluid) to weakly collisional (Monte Carlo), TLBM can give a unified framework, and thus avoid the stiff problem of coupling UEDGE to Monte Carlo. Our current codes run at a kinetic Courant number CFL = 1, so that no numerical diffusion or dissipation is introduced.

COMPUTATIONAL APPROACH

TLBM discretizes the Bhatnagar-Gross-Krook (BGK) kinetic equation to solve the system on a lattice. In its simplest form, the algorithm advances the distribution at time t to time $t + 1$ by (a) free-streaming (at CFL = 1) the distribution function from one spatial node to nearest lattice neighbors (this only involves the shift operation and use of MPI to handle boundary points in the domain decomposition); (b) recomputing local mean variables (simple summations, a local operation); (c) linear collisional relaxation (local operation).



Evolution of 2D jet flow at Mach number 0.5 between constant temperature walls using the 9-bit energy-dependent lattice: a long-time comparison of the scaling integrity of the flows for Reynolds number = 20,000. Left column: timestep = 300 K, temperature = 8.0 (–6), v0 = 2.0 (–3). Right column: timestep = 60 K, temperature = 2.0 (–4), v0 = 1.0 (–2).

ACCOMPLISHMENTS

In order to improve the numerical stability of TLBM, we are employing an energy-dependent octagonal 2D lattice. Initial results are very promising as we investigate jet flow into highly stratified background.

As we move to modeling divertor physics, we have been investigating two-fluid equilibration. In particular, we have been examining the effects of velocity shear turbulence of a lighter species on a laminar heavier species. We have verified the Morse 1967 theory dealing with the rate of species temperature equilibration to velocity equilibration. This is dimension independent for weakly turbulent systems.

SIGNIFICANCE

In the standard computational fluid dynamics approach to solving the nonlinear equations, one must handle the nonlinear Riemann problem and over 30% of the CPU time is spent in resolving the nonlinear convective derivative. In TLBM, one sidesteps the Riemann problem altogether and can use Lagrangian streaming to handle the linear advective derivative. In essence, by embedding the nonlinear system into higher dimensional phase space (i.e., by going to a linearized kinetic description), we can choose a simplified system (e.g., a BGK collision operator) to recover the desired equations. This concept is similar to using multi-scale perturbation theory to solve singular problems in applied math/physics. TLBM codes are very well suited for parallel machines.

PUBLICATIONS

L. Vahala, D. Wah, G. Vahala, J. Carter, and P. Pavlo, "Thermal lattice Boltzmann simulation for multispecies equilibration," Phys. Rev. E **62**, 507 (2000).

G. Vahala, J. Carter, D. Wah, L. Vahala, and P. Pavlo, "Parallelization and MPI performance of thermal lattice Boltzmann codes for fluid turbulence," in *Parallel Computational Fluid Dynamics '99*, edited by D. Keyes et al. (Elsevier Science, Amsterdam, 2000).

L. Vahala, G. Vahala, J. Carter, and P. Pavlo, "Velocity and temperature equilibration for multi-species gases using thermal lattice Boltzmann simulations," Czech. J. Physics (in press).

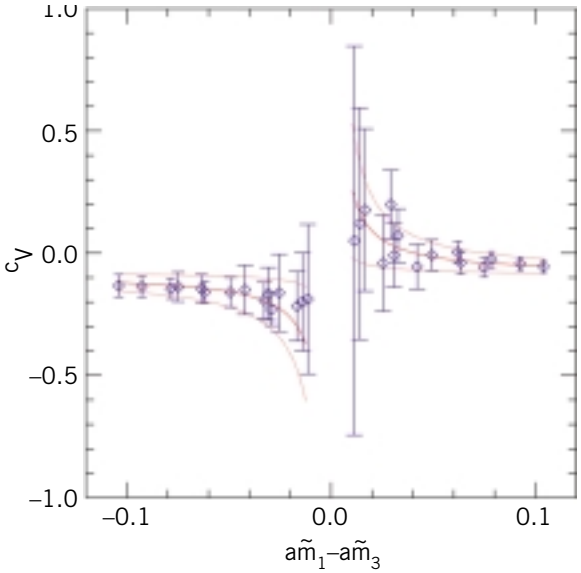
http://www.physics.wm.edu/~vahala/july00.html

Phenomenology with $O(a)$ Improved Lattice QCD

Rajan Gupta, Tanmoy Bhattacharya, and Weonjong Lee,
Los Alamos National Laboratory
Stephen Sharpe, University of Washington

RESEARCH OBJECTIVES

Lattice QCD provides the most promising non-perturbative approach to solving the highly non-linear behavior of quarks and gluons, the building blocks of strongly interacting particles. A price paid for discretizing QCD onto a space-time grid to make it amenable to numerical simulations is the introduction of errors associated with the granularity of the lattice. These errors can be reduced by improving the discretization scheme and by calculating the quantum corrections on the operators. Our goal is to carry out a study of phenomenologically interesting quantities using a theory for which all corrections needed to remove the leading discretization errors, i.e., linear in the lattice spacing, have been determined non-perturbatively.



A fit of the form $c_V = c_V^{(0)} + c_V^{(1)}/(\tilde{m}_1 - \tilde{m}_3)$ to extract the improvement constant for the vector current. The pole term is an artifact of lattice discretization, and resolved in our calculation.

COMPUTATIONAL APPROACH

The basic tools for studying field theories like QCD on a computer are: (1) Monte Carlo methods for generating importance sampled background gauge configurations. We have used a combination of Metropolis and over-relaxed algorithms. (2) The calculation of quark propagators by inverting the Dirac matrix. We have done this using a stabilized bi-conjugate gradient iterative solver. Using these gauge and quark degrees of freedom, physical quantities are extracted by constructing gauge-invariant correlation functions.

ACCOMPLISHMENTS

We have extended the method based on using axial and vector Ward identities to calculate the renormalization and improvement constants for lattice QCD. These include all the scale independent renormalization constants, the mass dependence of the renormalization constants for all quark bilinear operators, the improvement constants for currents, and the coefficients of the equation of motion operators that arise at $O(a)$. Precise results have been obtained for two values of the lattice scale at which calculations by many lattice collaborations have been done. One of the highlights of our approach — the result for the improvement constant for the vector current — is shown in the figure. The uncertainty in this quantity was reduced by a factor of 4 compared to an earlier method used by the ALPHA collaboration.

SIGNIFICANCE

In order to obtain the full improvement in scaling behavior of quantities calculated using better discretization schemes for the lattice action, it is also necessary to determine all the renormalization and improvement constants. Our results will be used by all lattice collaborations using the Symanzik $O(a)$ improved lattice theory and will also form the basis for our future calculations.

PUBLICATIONS

T. Bhattacharya, S. Chandrasekharan, R. Gupta, W. Lee, and S. Sharpe, "Non-perturbative renormalization constants using Ward identities," Phys. Lett. B **461**, 79 (1999).

T. Bhattacharya, S. Chandrasekharan, R. Gupta, W. Lee, and S. Sharpe, "Order *a* improved renormalization constants," Phys. Rev. D (submitted); hep-lat/0009038 (2000).

Continuing Studies of Plasma-Based Accelerators

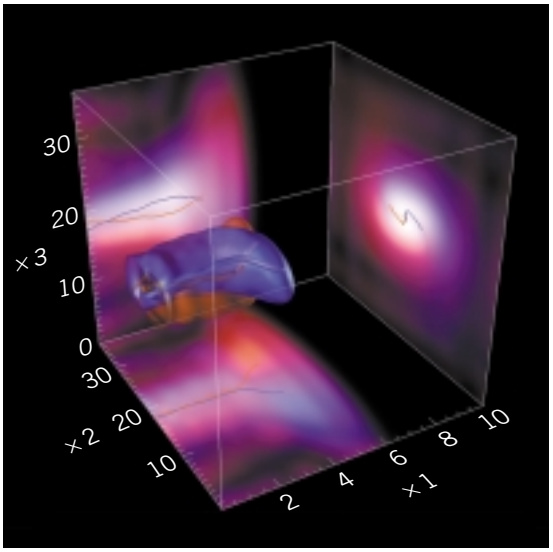
Chan Joshi and Warren Mori,
University of California, Los Angeles

RESEARCH OBJECTIVES

This research attempts to test the feasibility of various plasma-based accelerator concepts, to model full-scale plasma-based accelerator experiments, and to help develop new advanced accelerator concepts. The main objectives are to support the beat wave experiments at the Neptune Lab at UCLA, to support the E-157 experiment at SLAC (a collaborative experiment between SLAC/UCLA/USC), and to study basic physics in intense particle and laser plasma interactions.

COMPUTATIONAL APPROACH

We apply particle-based models including fully explicit particle-in-cell (PIC) codes, ponderomotive guiding center PIC codes, and new photon kinetic codes. We are integrating all these algorithms into an object-oriented framework that supports massively parallel processing.



Isosurfaces and centroids of the electric field of two lasers (red and blue) show the laser braiding. The projections of the laser centroids are also shown on the walls. This simulation shows that in a plasma, one light beam can influence another beam's propagation by affecting the properties of the medium. The attractive force originates from relativistic mass increase of the plasma electrons in a strong laser field. This braiding effect might be useful in optical steering applications and might occur in nature when intense photon fluxes filament as they emanate from supernovas and powerful celestial gamma ray sources.

ACCOMPLISHMENTS

Full-scale 2D and 3D modeling of the E-157 experiment has allowed us to study the role of the focusing and the acceleration wakefields. These simulations have been instrumental in the proper interpretation of this plasma wakefield experiment and have been a guide for finding new physics in GeV beam plasma interactions.

We have begun to understand hosing of short pulse plasma wakefield drivers. We have also performed the first simulations of the wakes generated by positron drivers. And we have begun to study the feasibility of adding a 100 GeV afterburner wakefield stage at the end of E-157. This involves understanding both electron and positron drivers, beam loading and hosing.

We have verified in simulations our prediction that there is a mutual attraction between two laser beams in a plasma. The simulations showed the beams actually form a braided pattern. We have also performed simulations of a new asymmetric spot size self-modulation instability of intense lasers in plasmas. In addition, we have developed and tested a new ponderomotive guiding center parallel PIC code which will allow us to carry out full-scale 2D simulations of the Neptune beat wave experiments.

SIGNIFICANCE

In plasma-based acceleration, electrons "surf" on relativistic space charge plasma waves. In such waves electrons can be accelerated with gradients orders of magnitude larger than current technology. If plasma-based accelerator technology is successfully developed, then multi-GeV stages could be miniaturized to fit on a tabletop. Tabletop accelerators could have impacts in fields as diverse as high-energy physics, synchrotron radiation sources, medicine, and biology. If the simulations indicate that an afterburner concept is found to be viable, then this work could lead to much larger and broader R&D effort.

PUBLICATIONS

C. Ren, R. G. Hemker, R. A. Fonseca, B. J. Duda, and W. B. Mori, "Mutual attraction of laser beams in plasmas: Braided light," Phys. Rev. Lett. **85**, 2124 (2000).

R. G. Hemker, W. B. Mori, S. Lee, and T. Katsouleas, "Dynamic effects in plasma wakefield excitation," Phys. Rev. ST Accel. Beams **3**, 61301 (2000).

M. J. Hogan, R. Assmann, F.-J. Decker, et al., "E-157: A 1.4-m-long plasma wake field acceleration experiment using a 30 GeV electron beam from the Stanford Linear Accelerator Center Linac," Phys. Plasmas **7**, 2241 (2000).

http://www.ee.ucla.edu/labs/laser-plasma/

Cosmic Microwave Background Data Analysis: The BOOMERANG Long Duration Balloon Flight

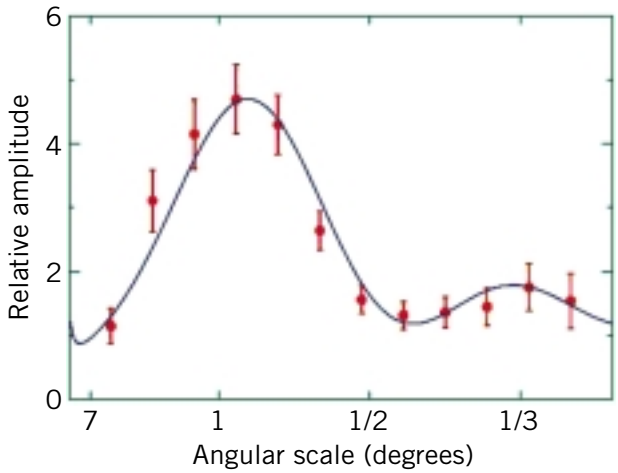
Andrew Lange, California Institute of Technology
John Ruhl, University of California, Santa Barbara
Andrew Jaffe, Center for Particle Astrophysics, University of California, Berkeley
Julian Borrill, NERSC, Lawrence Berkeley National Laboratory

RESEARCH OBJECTIVES

In January 1999 the BOOMERANG Long Duration Balloon flight spent 10.5 days in the Antarctic stratosphere measuring the temperature of the cosmic microwave background (CMB). The resulting data set is the most significant measurement of the tiny fluctuations in the CMB temperature since they were first detected by the COBE satellite. This research project is devoted to analysis of this data set.

COMPUTATIONAL APPROACH

The analysis of a massive CMB data set can be recast as a problem in the solution of linear systems involving very large, dense, symmetric matrices. First we convert the time-ordered CMB data to a pixelized map, triangular-solving a linear system with a single right hand side to obtain the maximum of the map likelihood function. Then we apply a Newton-Raphson iterative method to locate the peak of the CMB power spectrum likelihood function (which has no closed-form solution) given this map. Each iteration requires triangular-solving many linear systems, each with as many right hand sides as there are pixel-pixel correlation matrix rows and columns. The entire analysis algorithm has been implemented in parallel as the Microwave Anisotropy Dataset Computational Analaysis Package (MADCAP).



The strength of the CMB fluctuations on different angular scales as measured by BOOMERANG: points with error bars are the data, while the curve corresponds to the best-fitting cosmological model.

ACCOMPLISHMENTS

This year saw the first publication of results from the BOOMERANG LDB experiment. Using the MADCAP code on the T3E at NERSC, we analyzed the output of individual detectors and confirmed the high quality of the data. We then produced a map of 30,000 pixels. We excised the highest-quality central region of this map, comprising 8,000 pixels. From this region, we estimated a single gold-plated angular power spectrum over a broad range of angular scales, also with MADCAP. We analyzed this power spectrum and determined such cosmological parameters as the geometry of the Universe and the power spectrum of density perturbations. We have shown that these are consistent with expectations of the inflationary paradigm for the origin of structure in the Universe.

SIGNIFICANCE

The CMB is the earliest photon-picture of the Universe we can ever obtain, showing the state of the Universe 300,000 years after the Big Bang. It is our best window onto the early Universe and the most powerful discriminant between competing cosmological models. The tiny fluctuations in the CMB temperature correspond to the very first density perturbations in the Universe. Their pattern contains detailed information about all the fundamental parameters of cosmology — the Universe's geometry, expansion rate, number of neutrino species, ionization history, and the energy density in baryons, dark matter, and cosmological constant.

PUBLICATIONS

A. H. Jaffe et al., "Cosmology from MAXIMA-1, BOOMERANG and COBE/DMR CMB observations," Phys. Rev. Lett. (submitted); astro-ph/0007333 (2000).

A. E. Lange et al., "First estimations of cosmological parameters from BOOMERANG," Phys. Rev. D (submitted); astro-ph/0005004 (2000).

P. de Bernardis et al., "A flat Universe from high-resolution maps of the cosmic microwave background radiation," Nature **404**, 955 (2000).

http://www.physics.ucsb.edu/~boomerang/
http://cfpa.berkeley.edu/~borrill/madcap.html
http://www.nersc.gov/news/boomerang4-26-00.html

Lattice QCD Monte Carlo Calculation of Hadron Structure

Keh-Fei Liu, Terrence Draper, and Shao-Jing Dong,
University of Kentucky

RESEARCH OBJECTIVES

We plan to study chiral condensate, decay constants, hadron and quark masses, chiral logs, nucleon form factors, and the sea quark contributions, such as the strangeness content in the nucleon. Our goal is to push the calculation of various fundamental physical quantities to the continuum limit, the chiral limit, and the large-volume limit within the quenched approach.

COMPUTATIONAL APPROACH

We have implemented Neuberger's overlap fermion to test chiral symmetry and scaling via the calculation of hadron masses, quark masses, and the chiral condensate. The new overlap fermion action involves a matrix sign function. We approximate the square root of the matrix by the optimal rational polynomial approach, and we invert the matrix with conjugate gradient with a multiple mass algorithm. To speed up the convergence, we project out some of the smallest eigenvalues and treat the sign function of these states exactly. The overall inversion of the quark matrix to obtain the quark propagator is also done with conjugate gradient with multiple quark masses.

ACCOMPLISHMENTS

Our calculation on small volumes and three different lattice spacings yields encouraging results. The chiral symmetry breaking due to numerical implementation is limited to less than 1% for the smallest quark mass, and the scaling of hadron masses shows that there is no $O(a)$ error, and even the $O(a^2)$ error is small. We have implemented the overlap fermion on large quenched lattices (20^4 with a = 0.15 F) with 24 different quark masses, with the smallest one close to the physical quark mass. This requires a delicate balance between projecting enough small eigenvalues for chiral symmetry and faster convergence in the matrix inversion, and not exceeding the memory on 64 nodes.

We have accumulated data on 40 gauge configurations. We have fairly accurate results on the pion mass, and from it we have extracted the chiral log reasonably reliably. The other hadron masses are still noisy, but we begin to see a trend that the isovector scalar meson mass and that of the axial-vector meson cross over for light quark masses. This is consistent with experiments and is now seen on the lattice for the first time.

SIGNIFICANCE

Chiral symmetry is a fundamental symmetry in QCD that governs low-energy hadron structure and dynamics. The lack of lattice formulation of this symmetry has so far hindered reliable extrapolation of lattice results to the physical pion mass region. With the advent of Neuberger's overlap fermion, physical observables sensitive to this symmetry should be calculated more reliably, and they can be compared with experiments more readily and directly.

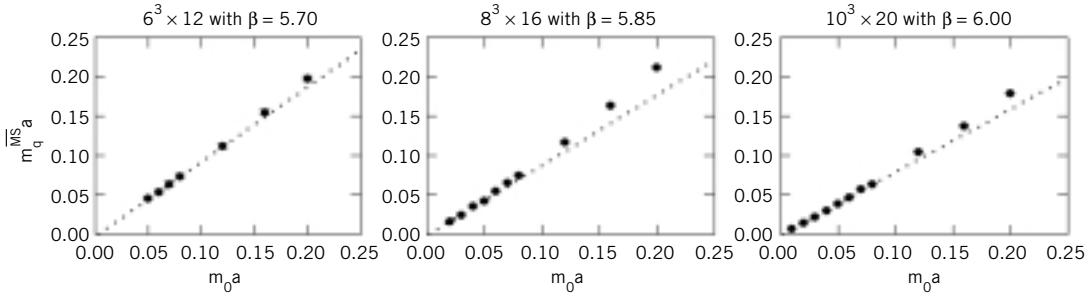
PUBLICATIONS

J. Christensen, T. Draper, and Craig McNeile, "Renormalization of the lattice heavy quark effective theory Isgur-Wise function," Phys. Rev. D **62**, 114006 (2000); hep-lat/9912046.

L. Lin, K. F. Liu, and J. Sloan, "A noisy Monte Carlo algorithm," Phys. Rev. D **61**, 74505 (2000).

K. F. Liu, S. J. Dong, T. Draper, D. Leinweber, J. Sloan, W. Wilcox, and R. M. Woloshyn, "Valence QCD and quark model," Phys. Rev. D **59**, 112001 (1999).

http://www.pa.uky.edu/~liu



Renormalized quark mass in the $\overline{MS}$ scheme vs. the bare quark mass on the three lattices with the same physical volume but different lattice spacings, calculated with the overlap fermion action, which is known to possess lattice chiral symmetry. The renormalized quark mass does not have an additive part due to chiral symmetry. In fact, the linear fit including the smallest 5, 7, and 8 quark masses for β = 5.7, 5.85, and 6.0, respectively, shows that the intercepts are of the order of 10^{-3} and are consistent with zero. This is proof that the overlap fermion action resembles the continuum physics and overcomes the difficulty faced by the previous Wilson-type fermion actions.

Exploration of Hadron Structure Using Lattice QCD

John Negele, Stefano Capitani, Patrick Dreher, Alvaro Montero, Andrew Pochinsky, Dru Renner, James Steele, and Uwe-Jens Wiese, Massachusetts Institute of Technology
Robert Edwards, Thomas Jefferson National Accelerator Facility
Thomas Lippert, Hartmut Neff, and Klaus Schilling, Universität Wuppertal

RESEARCH OBJECTIVES

The major focus of the work is on two key issues: Understanding the role of instantons and their associated quark zero modes in nucleon structure, and using the quark zero modes to calculate the sea quark content of the nucleon. This project has several physics objectives. By calculating the spectrum for an ensemble of configurations, we expect to understand the spectrum and the degree of separation of physical modes from unphysical doublers, and to check the relation between the density of fermion modes with low virtuality and the chiral condensate. We will carry out a high statistics study of the degree to which hadron propagators are dominated by zero modes. We will reconstruct the topological charge density from the quark eigenmodes to obtain an unambiguous determination of the instanton content of the QCD vacuum. We will use the zero modes to calculate the disconnected diagrams corresponding to the strange quark content of the nucleon.

COMPUTATIONAL APPROACH

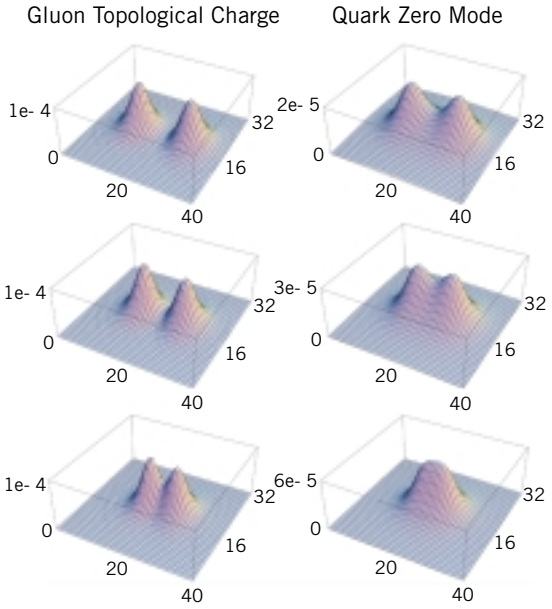
We calculate the low eigenmodes of the Dirac operator using the k-step Arnoldi method. This method has compelling advantages for our work. First, since it works in a fixed dimension space, there is no degradation of orthogonality and corresponding loss or duplication of modes. Second, its insensitivity to the quark mass makes it extremely useful near the chiral limit of low pion mass. We have two complementary implementations. One is an exploratory code in which we can control the region of eigenvalues at will. The other uses the robust and well-optimized PARPACK package from ORNL.

ACCOMPLISHMENTS

We began major production late in the fiscal year, but we have verified that our truncated eigenvector approach provides a statistically superior signal to that obtained with conventional stochastic estimators. Results show that using 300 low-lying eigenmodes significantly increases the signal.

SIGNIFICANCE

This project will develop a new method that can attain a higher level of statistical accuracy than existing methods and will provide the essential quark zero modes necessary for these calculations. In addition to elucidating the physics for timely parity-violating electron-scattering experiments, this new method should also enable the evaluation of the disconnected diagrams encountered in deep inelastic electron scattering.



Topological excitations of the gluon field in QCD, which play an important role in generating quark masses and interactions, can be identified by localized quark zero modes. The gluon topological charge (left) and corresponding quark zero mode (right) are shown for a meron pair on a lattice in an ongoing project to explore the role of these configurations in producing quark confinement.

We will perform the first quantitative study of the eigenmodes for a full ensemble of configurations. This will enable us to perform the most precise explorations to date of the eigenvalue spectrum of the Dirac operator, the separation of physical modes from lattice artifacts, the instanton content of the QCD vacuum, the quantitative accuracy of the 't Hooft interaction, and the relation between the density of eigenvalues and the chiral condensate, as expected from the Banks-Casher formula.

PUBLICATIONS

James V. Steele and J. W. Negele, "Meron pairs and fermion zero modes," Phys. Rev. Lett. (in press); hep-lat/0007006.

John W. Negele, "Instantons, the QCD vacuum, and hadronic physics," Nucl. Phys. B (Proc. Suppl.) **73**, 92 (1999); hep-lat/9810053.

O. Jahn, F. Lenz, J. W. Negele, and M. Thies, "Center vortices, instantons, and confinement," Nucl. Phys. B (Proc. Suppl.) **83**, 524 (2000).

http://www-ctp.mit.edu/~negele/
http://www.fzJuelich.de/nic/Forschungsgruppen/Elementarteilchen/

STAR Detector Simulations and Data Analysis

Douglas L. Olson,
Ernest Orlando Lawrence Berkeley National Laboratory
John Harris, Yale University

RESEARCH OBJECTIVES

The STAR detector (Solenoidal Tracker at RHIC) at Brookhaven National Laboratory is a large acceptance collider detector designed to study the collision of heavy nuclei at very high energy in the laboratory. Its goal is to investigate nuclear matter at extreme energy density and to search for evidence of the phase transition between hadronic matter and the deconfined quark-gluon plasma (QGP). STAR and the RHIC accelerator began operation in June 2000. The complete STAR detector will contain a set of time projection chambers for charged particle tracking, a silicon detector for vertexing, electromagnetic calorimeters, and a number of other systems.

COMPUTATIONAL APPROACH

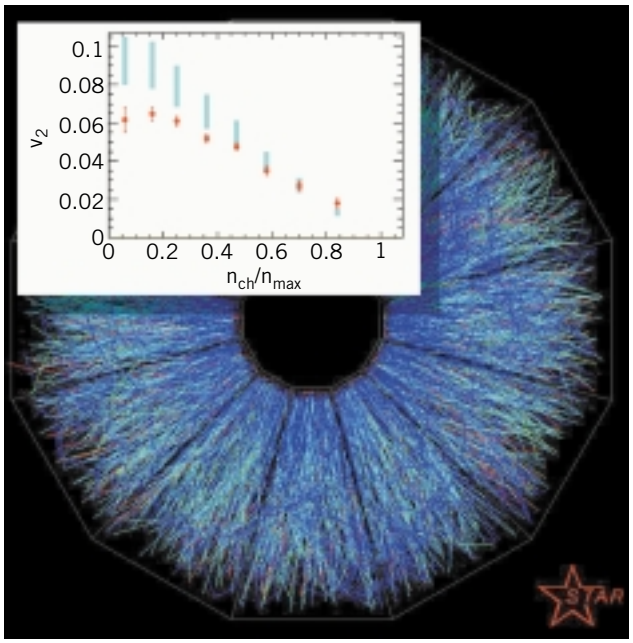
The basic method of deriving physics results in experimental relativistic heavy ion collisions is to carry out statistical analysis of large numbers of events (collisions of individual atomic nuclei). The theoretical models are implemented as Monte Carlo codes that describe the final state of each of the thousands of particles that are produced in these collisions. We use a number of these theoretical codes (VENUS, HIJING, RQMD, and others) to produce large samples of events. A simulation code called GEANT is used to propagate each of these thousands of particles through the material of the STAR detector and compute the reactions and energy deposition that occurs throughout the detector. These theoretical model codes and the detector simulation code are run on MPP systems utilizing the natural parallelism of the problem, namely that each event is independent, so that different events are computed in parallel on the various processor nodes.

ACCOMPLISHMENTS

In FY 2000 STAR detector simulations were run on the T3E with both the VENUS and HIJING event generators in various configurations of impact parameters and physics conditions, all for the detector configuration that STAR is presently running with at RHIC. In total (as of July 2000) we produced 26K events of Au + Au collisions with about 1.5 TB total volume.

SIGNIFICANCE

The existence of the QGP is predicted by lattice QCD calculations, and this state of matter is thought to be important in the dynamics of the early universe and the core of neutron stars. The most violent nuclear collisions at RHIC will generate approximately 10 thousand secondary particles. STAR aims to detect and characterize a large fraction of these secondaries in order to reconstruct a meaningful picture of each individual collision.



Elliptic flow in Au + Au collisions at $\sqrt{s_{NN}}$ = 130 GeV, measured by the STAR experiment in August 2000. The figure shows elliptic flow (solid points) as a function of centrality defined as n_{ch}/n_{max}. The vertical bars show a range of values expected for v_2 in the hydrodynamic limit, scaled from the initial space eccentricity of the overlap region of the two colliding gold nuclei. The scaling factor used ranges from 0.19 to 0.25.

PUBLICATIONS

D. Hardtke (for the STAR Collaboration), "Inclusive particle spectra and exotic particle searches using STAR," in *Proceedings of ISMD99* (Brown University, Rhode Island, Aug. 9–13, 1999).

H. Caines (for the STAR Collaboration), "The year-one physics capabilities of STAR," in *Proceedings of the Relativistic Heavy Ion Mini-Symposium at the APS Centennial Meeting* (Atlanta, GA, March 21–26, 1999).

W. B. Christie, "Data sets for high pT physics with the STAR detector," in *Proceedings of Hard Parton Physics in High Energy Nuclear Collisions Workshop* (Brookhaven National Laboratory, Upton, NY, March 1–5, 1999); report No. BNL-52574.

http://www.star.bnl.gov/

Variational and Green's Function Monte Carlo Calculations of Light Nuclei

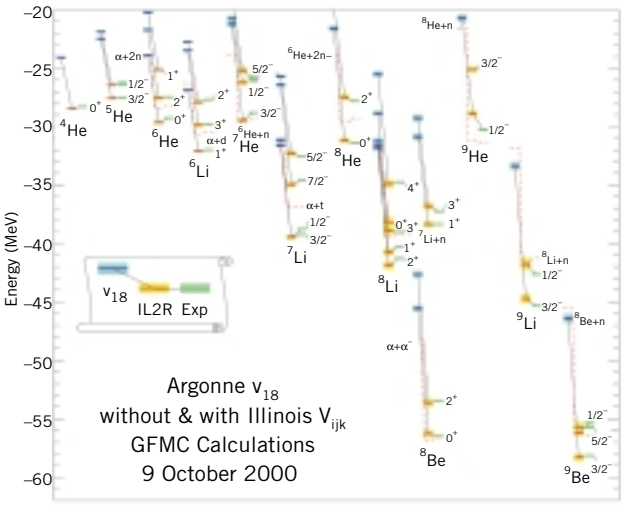
Steven Pieper, Argonne National Laboratory

RESEARCH OBJECTIVES

This project uses Green's function (GFMC) and variational (VMC) Monte Carlo (generically known as quantum Monte Carlo) methods to compute ground-state and low-lying excited state expectation values of energies, densities, structure functions, astrophysical reaction rates, etc., for light nuclei and neutron drops. Realistic two- and three-nucleon potentials are used. We are developing new computational techniques, optimizing them for different computer architectures, and improving the nuclear Hamiltonian used in the calculations. An area of increasing interest is the use of our wave functions to compute reaction rates of astrophysical interest.

COMPUTATIONAL APPROACH

This project uses variational and Green's function Monte Carlo methods. The variational wave function contains noncentral two- and three-body correlations corresponding to the operator structure of the potentials. The GFMC systematically improves these wave functions to give the exact (within statistical errors) energy for the given Hamiltonian. We have demonstrated the reliability of constrained path methods for overcoming the well-known Fermion sign problem.



Comparison of computed energy levels of light nuclei with experimental (green) values. The blue lines are computed without a three-nucleon potential and fail to reproduce experiment. The red lines have a modern three-nucleon potential. The A = 9 calculations shown here are the first *ab initio* calculations of nine-body nuclei and were made on NERSC's IBM SP.

The Quantum Monte Carlo methods work very efficiently on parallel processors. We use MPI and see speed-up efficiencies better than 97% for up to 512 CPUs on NERSC's IBM SP. The program has achieved 100 Gflops on the SP.

ACCOMPLISHMENTS

This year we finished our study of new three-nucleon potential terms. Our previous work had demonstrated that the Hamiltonian that has been used successfully for more than a decade in studies of *s*-shell nucleon is inadequate in the *p*-shell. Some of the possible new potential terms, whose forms are derived from meson-exchange arguments, result in considerable additional complications in the Green's function propagator. We have developed potential models that reproduce all of the known stable or narrow-width levels of up to nine-body nuclei with an average error of only 300 keV.

SIGNIFICANCE

One of the principal goals of nuclear physics is to explain the properties and reactions of nuclei in terms of interacting nucleons (protons and neutrons). There are two fundamental aspects to this problem: (1) determining the interactions between nucleons, and (2) given the interactions (i.e., the Hamiltonian), making accurate calculations of many-nucleon systems. We work in both areas and have made the only calculations of six- through ten-nucleon systems that use realistic interactions and that are accurate to 1% for the binding energies. The resulting wave functions can be used to compute properties measured at electron and hadron scattering facilities (in particular Jefferson Laboratory), and to compute astrophysical reaction rates, many of which cannot be measured in the laboratory.

PUBLICATIONS

R. B. Wiringa, Steven C. Pieper, J. Carlson, and V. R. Pandharipande, "Quantum Monte Carlo calculations of A = 8 nuclei," Phys. Rev. C **62**, 14001 (2000).

D. Van Neck, M. Waroquier, A. E. L. Dieperink, S. C. Pieper, and V. R. Pandharipande, "Center-of-mass effects on the quasi-hole spectroscopic factors in the ^{16}O(e,e' p) reaction," Phys. Rev. C **57**, 2308 (1998).

B. S. Pudliner, V. R. Pandharipande, J. Carlson, S. C. Pieper, and R. B. Wiringa, "Quantum Monte Carlo calculations of nuclei with A ≤ 7," Phys. Rev. C **56**, 1720 (1997).

http://www.phy.anl.gov/theory/research/forces.html

Advanced Computing for 21st Century Accelerator Science and Technology

Robert Ryne, Los Alamos National Laboratory
Kwok Ko, Stanford Linear Accelerator Center

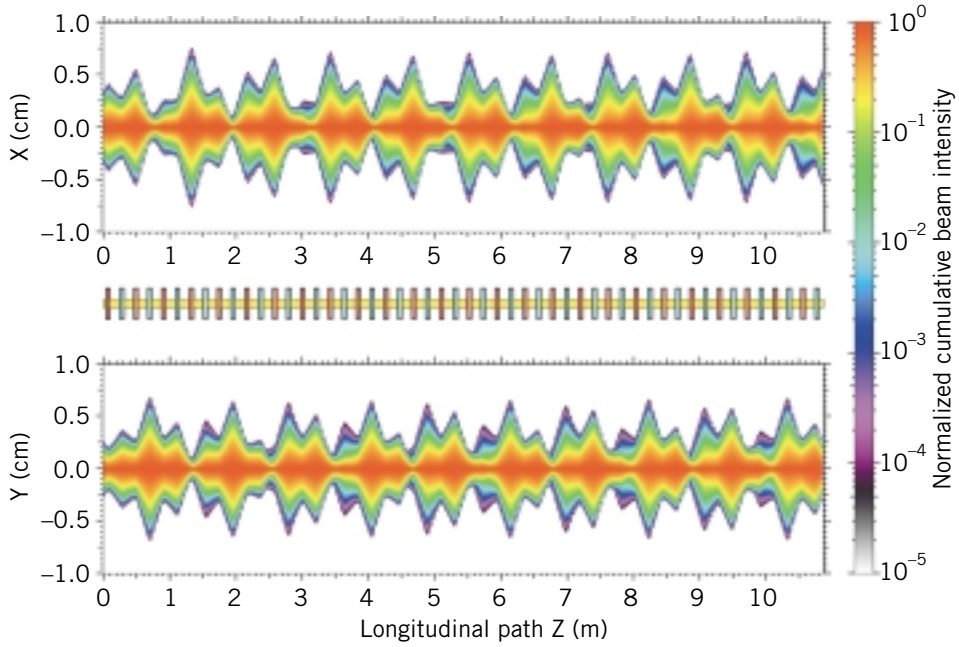
RESEARCH OBJECTIVES

The primary purpose of this project is to develop a comprehensive, coherent terascale accelerator simulation environment (ASE) for the U.S. particle accelerator community, and to apply it to projects of national importance. The ASE will contain a suite of codes and modules that will enable accelerator physicists and engineers to collaborate and solve the most challenging problems in accelerator design, analysis, and optimization. Omega3P, Tau3P, and IMPACT (all developed under the Accelerator Grand Challenge) are examples of codes that will be contained within the ASE. But new codes and modules remain to be developed: electromagnetic codes for modeling lossy structures and wakefields, parallel statics codes for electric and magnetic component design, and parallel codes for modeling intense beams in injectors and circular machines. In addition, modules need to be developed to include such physical phenomena as collisions, synchrotron radiation, and surface emissions. Furthermore, the software components in the ASE suite must be able to work together to enable the difficult calculations involving tight coupling of beam dynamics and electromagnetics.

This project has three main areas: Beam Systems Simulation (BSS), Electromagnetic Systems Simulation (ESS), and Beam/Electromagnetic Systems Integration (BESI).

COMPUTATIONAL APPROACH

BSS uses parallel particle-in-cell (PIC) techniques, particle managers, dynamic load balancing, FFT-based Poisson solvers, and techniques from the field of magnetic optics. Split-operator methods are used to combine magnetic optics and parallel PIC techniques in a single framework, and to establish efficient particle advance algorithms. ESS uses unstructured mesh generation, domain decomposition, adaptive mesh refinement, finite elements and sparse linear solvers (for eigenmode codes), and unstructured Yee grids (for time-domain codes). BESI, involving particles in electromagnetic structures, will use hybrid grids, with a structured mesh near the beam and an unstructured grid near the structure boundaries.



Horizontal (top) and vertical (bottom) cumulative density profiles from a simulation of a mismatched beam in the Low-Energy Demonstration Accelerator (LEDA) beam halo experiment at Los Alamos. A schematic layout of the alternately focusing and defocusing quadrupole magnets that comprise the beamline is shown between the graphs. This simulation was performed with 100 million macroparticles. The contours display the amount of charge that would be intercepted by a scraper placed at a given location. In this case, the amount of charge intercepted has been calculated with a resolution of about 1 part in 10 thousand.

ACCOMPLISHMENTS

BSS accomplishments:

We have developed LANGEVIN3D, which we used to perform the first-ever self-consistent 3D Fokker-Planck simulations.

We have extended our former beam dynamics work so that we can model space-charge effects in beams in circular accelerators. We soon will be able to "flip a switch" to turn on/off space-charge effects in the major beam transport codes.

We have enhanced our parallel Poisson solver capability by extending our treatment of boundary conditions from two to six cases. The most important cases (conducting wall boundary conditions transversely, with open or periodic boundary conditions longitudinally) are an important new capability needed for simulations in support of upcoming experiments related to the physics of beam halo formation.

We have increased our interaction with the accelerator community. IMPACT is now being used to perform large-scale simulations of high intensity beams in rf accelerators at LANL, Fermilab, BNL, ORNL, PSI (Zurich), and CERN (Geneva) for major projects, including the Spallation Neutron Source, the neutrino factory/muon collider, the CERN Superconducting Proton Linac, and the LEDA beam halo experiment.

ESS accomplishments:

Previously, we developed the parallel eigensolver Omega3P and used it to design the cells in the Next Linear Collider (NLC) RDDS. This year we ported Omega3P to the IBM SP and succeeded in modeling a six-cell RDDS structure involving up to 18 million degrees of freedom. These runs are used to generate the dispersion curves for the lowest three dipole bands needed for wakefield calculations. We have also simulated the APT CCL cavity on the T3E and obtained results that agree well with measurement.

We have incorporated an adaptive mesh refinement strategy within Omega3P to find accurate peak-power loss densities in complex rf cavities needed for designing cooling channels to control thermal effects in high power operations.

We successfully used Tau3P to model a 10-cell structure corresponding to the output end of an NLC RDDS, and we calculated the external loading of the dipole modes at the output end due to the fundamental output coupler. Measurements were in excellent agreement with numerical calculations.

We implemented a rigid beam in the serial version of Tau3P to model wakefields, and the parallel version is in progress. In parallel, an effort to construct a RDDS 206-cell mesh is under way so that it can be used in a Tau3P simulation with beams to find the dipole wakefield.

We have begun development of a new 3D parallel electromagnetic solver called Phi3P, which is based on fields instead of potentials for higher accuracy.

We have increased our interaction with the accelerator community. Omega3P is becoming the most widely used code for performing large-scale eigenmode calculations of electromagnetic cavities for accelerator applications. It is used at SLAC, LBNL, LANL, and Fermilab for major projects, including the NLC, the PEP-II B-factory, the Advanced Light Source, Accelerator Production of Tritium, and the Spallation Neutron Source.

Tau3P simulation of the Next Linear Collider (NLC) Rounded Damped Detuned Structure (RDDS) output end, modeled on NERSC's T3E.

SIGNIFICANCE

The design and construction of the next generation of accelerators will involve greater complexity than ever before, and will require unprecedented precision in accelerator design and beam control. Examples include the Next Linear Collider, in which beams will be manipulated on a scale ranging from millimeters in the main linac to tens of nanometers at the collision point; a muon collider, in which a 100 MW-class proton beam will be used to produce short-lived muons that must be cooled by six orders of magnitude and brought into collision in a matter of microseconds; and a fourth-generation light source, requiring nanometer-scale smoothness in the beam pipe to successfully control an electron beam and produce intense, ultrashort X-ray pulses for imaging ultrafast phenomena. For all these accelerator systems, terascale simulation will play a key role by facilitating important design decisions, increasing safety and reliability, optimizing performance, and ultimately, helping to ensure project completion within budget and on schedule.

PUBLICATIONS

J. Qiang, R. Ryne, S. Habib, and V. Decyk, "An object-oriented parallel particle-in-cell code for beam dynamics simulation in linear accelerators," J. Comp. Phys. **163**, 434 (2000).

J. Qiang, R. Ryne and S. Habib, "Self-consistent Langevin simulation of coulomb collisions in charged-particle beams," in *Proceedings of SC2000* (Dallas, TX, Nov. 4–10, 2000).

J. Qiang and S. Habib, "A second-order stochastic leap-frog algorithm for multiplicative noise Brownian motion," Phys. Rev. E (in press).

http://laacg1.lanl.gov/

QCD at Finite Baryon Number Density and Finite Isospin Density

Donald Sinclair, Argonne National Laboratory

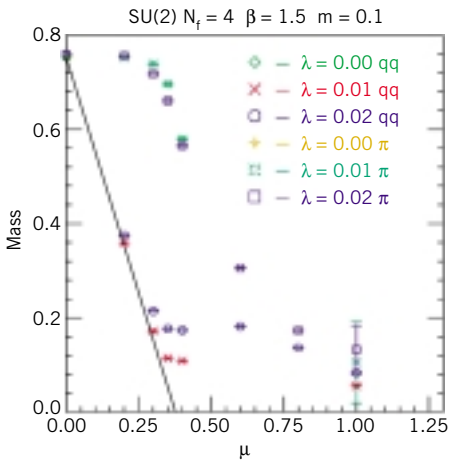
RESEARCH OBJECTIVES

At high baryon number density, QCD is believed to enter a phase with a diquark condensate. Since no practical method is known to simulate QCD at finite chemical potential for baryon/quark number, we are studying a simple model — QCD where the color group is $SU(2)$ rather than $SU(3)$ — at finite quark number chemical potential (μ). We will simulate true ($SU(3)$) QCD with two flavors at finite isospin density. A charged pion condensate is expected to form for large enough chemical potential, giving rise to spontaneous breakdown of isospin and Goldstone bosons.

COMPUTATIONAL APPROACH

To measure diquark condensation, we include a small diquark source term, as well as the standard Dirac mass term. This model has a real positive fermion pfaffian, allowing us to simulate it by standard techniques. We perform simulations to determine the phase structure of this model with four flavors of light dynamical quarks. We measure the diquark and chiral condensates and the quark number density on a small (8^4) lattice to map the phase diagram. This is repeated on a larger ($12^3 \times 24$) lattice, where we also measure the spectrum of Goldstone bosons in each phase. Upon completion of these simulations at a relatively large quark mass, we will move to a smaller quark mass with its more complex spectrum of Goldstone and pseudo-Goldstone bosons.

We use standard hybrid molecular-dynamics simulations with gaussian noise for the (pseudo) fermions, allowing us to "tune" the number of flavors of staggered fermions. This works since the Dirac matrix with Majorana and Dirac masses is positive definite and the pfaffian is always positive. We use the Verlet method, modified to keep the errors $O(dt^2)$ in the presence of the noisy fermions, to discretize molecular dynamic time. We perform the required inversion of the Dirac matrix at each update using the standard conjugate gradient algorithm.



Pion (π) and scalar diquark (qq) masses as functions of μ. The straight line is *mass* = $m_\pi - 2\mu$. The arrow is at $\mu = m_\pi/2$.

ACCOMPLISHMENTS

Our preliminary simulations have identified a phase with a diquark condensate, as has recently been suggested for QCD. It does, however, have differences. At zero chemical potential its "baryon," a two-quark state, is in the same multiplet as the pion and is a Goldstone boson in the chiral limit. In addition, the observed diquark condensate is a color singlet, and gives rise to a true Goldstone mode. In the QCD case, any diquark condensate has color, and thus the "broken" symmetry would be expected to be realized as a Higgs phenomenon.

SIGNIFICANCE

Studies of QCD at finite baryon/quark number density (nuclear matter) have potential relevance to the physics of neutron stars. Studies of QCD at finite density and finite temperature are relevant to the physics of the early Universe. In addition, they are expected to be relevant to the physics of relativistic heavy ion collisions, which will soon be observed at RHIC, and which are already being observed at CERN, yielding preliminary evidence for a quark-gluon plasma. Finally, they would greatly enhance our knowledge of the structure of QCD.

QCD at finite chemical potential for isospin maps one surface of the phase diagram for nuclear matter (nuclei and neutron stars). It should exhibit spontaneous breakdown of isospin symmetry with a charged pion condensate and a true Goldstone boson. Our simulations will probe this behavior. Some of these properties are expected to survive to finite chemical potential for baryon number: nuclear matter has both finite isospin density and finite baryon number density.

PUBLICATIONS

S. J. Hands, J. B. Kogut, S. E. Morrison, and D. K. Sinclair, "Two-color QCD at finite fundamental quark-number density and related theories," Argonne National Laboratory technical report No. ANL-HEP-CP-00-110; hep-lat/0010028 (2000).

J.-F. Lagaë and D. K. Sinclair, "High temperature meson propagators with domain-wall quarks," Nucl. Phys. B (Proc. Suppl.) **83–84**, 405 (2000).

J.-F. Lagaë and D. K. Sinclair, "Domain wall fermions at finite temperature," Nucl. Phys. B (Proc. Suppl.) **73**, 450 (1999).

http://www.hep.anl.gov/dks/NERSC2000/

Weak Matrix Elements with Domain Wall Fermions

Amarjit Soni, Thomas Blum, Christopher Dawson, and Matthew Wingate, Brookhaven National Laboratory

RESEARCH OBJECTIVES

We are in the process of setting up a comprehensive framework for using lattice gauge methods with domain wall quarks (DWQ) for the calculation of weak matrix elements.

COMPUTATIONAL APPROACH

The basic ingredient in the method is a calculation of the quark propagator. The novel element in our method is that it requires introduction of a fictitious fifth dimension. The T3E-900 machine at NERSC is being used for these computations.

ACCOMPLISHMENTS

We have demonstrated that DWQ start to exhibit the crucial symmetries of the continuum theory (chiral symmetries) with a modest extent in the fifth dimension, i.e., even when the number of lattice sites in the extra dimension is as few as about 16 at $\beta \geq 6.0$. In the work finished so far, we have obtained a number of key results, including a calculation of the important matrix elements, B_K, $B_7^{3/2}$ and $B_8^{3/2}$, and the value of the strange quark mass. Furthermore, our results show that DWQ have significantly improved scaling behavior: the discretization errors are $O(a^2)$ and not $O(a)$. Our data indicate that the good scaling and chiral behavior of DWQ more than compensate for the added cost of the extra dimension.

SIGNIFICANCE

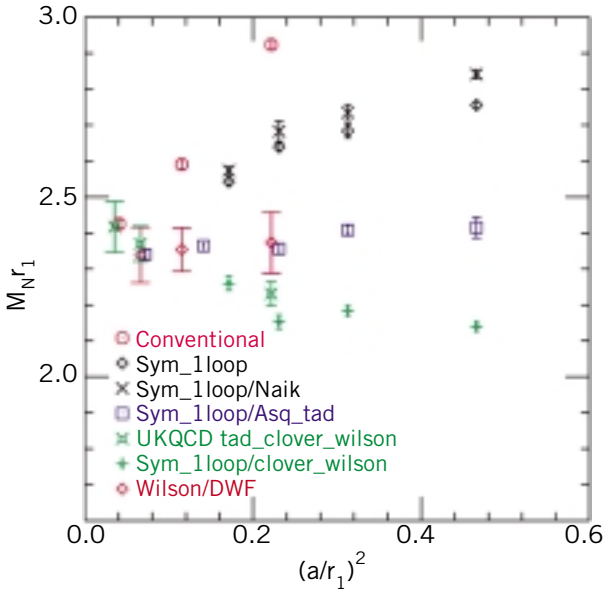
This work opens up an entirely new method for attacking some of the basic challenges in particle physics computations. For the past many decades, we have not been able to understand the strength of some simple reactions such as K decays to $\pi\pi$. Consequently, we have been unable to test the Standard Model of elementary particles through existing data and with improved experiments that are now under way. Using DWQ, we are now in the process of calculating the crucial charge parity violation parameter ε'/ε. Successful completion of this calculation should enable us to test for clues for the new physics that lies beyond the Standard Model.

PUBLICATIONS

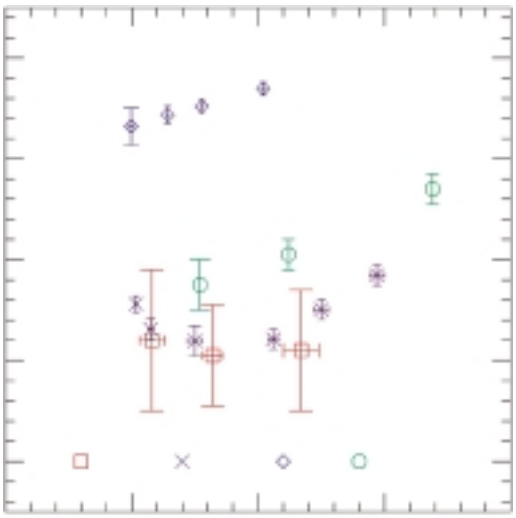
A. Soni, "Hadronic matrix elements from the lattice," hep-ph/0003092 (2000).

T. Blum, A. Soni, and M. Wingate, "Light quark masses with doman wall quarks," Phys. Rev. D **60**, 114507 (1999).

T. Blum and A. Soni, "QCD with domain wall quarks," Phys. Rev. D **56**, 174 (1997).



Comparison of the scaling of the nucleon mass from DWF with the other discretizations.



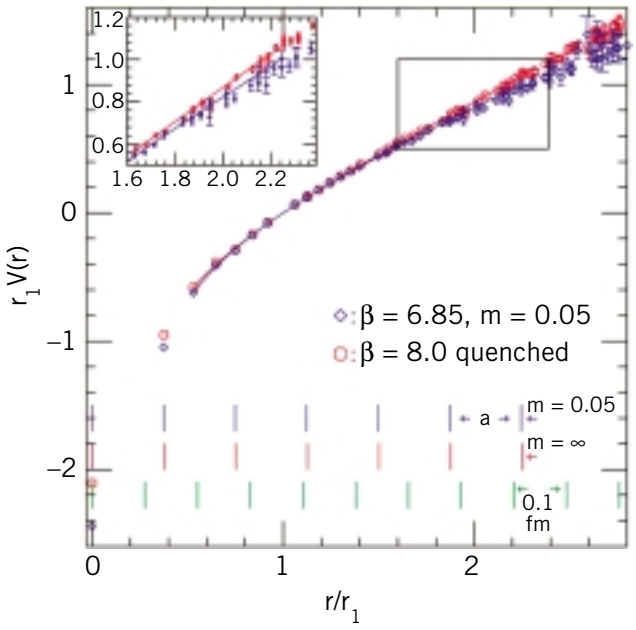
The strange quark mass after matching to the $\overline{\text{MS}}$ scheme.

High-Temperature QCD with Three Flavors of Improved Kogut-Susskind Quarks

Doug Toussaint, University of Arizona
Claude Bernard, Washington University
Tom DeGrand, University of Colorado
Carleton DeTar, University of Utah
Steven Gottlieb, Indiana University
Urs Heller, Florida State University
Bob Sugar, University of California, Santa Barbara

RESEARCH OBJECTIVES

We are using improved actions for Kogut-Susskind fermions in lattice QCD to study the effects of strange quarks on the phase transition, or crossover, between the low-temperature state of ordinary hadronic matter and the high-temperature quark-gluon plasma.



This figure shows the static quark potential for quenched (red) and three flavor (blue) QCD. The solid lines are fits to the Coulomb plus constant plus linear form. The upper two rulers show the lattice spacing in the two cases, and the lowest ruler shows units of 0.1 fm. The inset expands the area shown by the box. It can be seen that the potentials have different shapes. In particular, the potential with the quarks "turned on" is steeper at short distances. This agrees with theoretical expectations, which predict that the effective QCD coupling should decrease more slowly at short distances when the quark effects are included. Thus, the coupling constant with the quarks included should be larger, meaning a larger force or a steeper potential.

COMPUTATIONAL APPROACH

In lattice gauge theory calculations, one uses importance sampling techniques to numerically evaluate the Feynman path integrals from which all physical quantities in the theory can be obtained. We use the standard "refreshed molecular dynamics" algorithm to generate sample configurations of the gluon fields, or "lattices," with a probability proportional to their weight in the imaginary time quantum chromodynamics partition function. Then, expectation values of quantum mechanical operators can be estimated by averaging the operators over the set of gluon field configurations.

ACCOMPLISHMENTS

We studied the scaling properties of the Kogut-Susskind action by calculating the light hadron spectrum in the quenched approximation. We found that the masses of particles such as the π, ρ, and nucleon are nearly independent of lattice spacing in the range $a = 0.1$ fm to $a = 0.48$ fm. We have used our improved action to generate an extensive set of $20^3 \times 64$ lattices with three flavors of dynamical quarks at fixed lattice spacing, $a = 0.14$ fm.

SIGNIFICANCE

At very high temperatures, one expects to observe a phase transition or crossover from ordinary strongly interacting matter to a plasma of quarks and gluons. To observe such a transition or crossover successfully, it is important to determine the nature of the transition, the properties of the plasma, including strange quark content, and the equation of state. Lattice gauge theory has proven to be the only source of first-principle predictions about this form of matter in the vicinity of the phase transition, and our group has played an important role in the worldwide effort. Specifically, our improved actions for Kogut-Susskind fermions in lattice QCD, when combined with a well-understood improved action for the gluon fields, greatly reduce the effects of nonzero lattice spacing artifacts on physical results.

PUBLICATIONS

C. Bernard, T. Burch, T.A. DeGrand, C. DeTar, S. Gottlieb, U. M. Heller, J. Hetrick, K. Orginos, R. Sugar, and D. Toussaint, "The static quark potential in three flavor QCD," Phys. Rev. D **62**, 34503 (2000).

Kostas Orginos, Doug Toussaint, and R. L. Sugar, "Variants of fattening and flavor symmetry restoration," Phys. Rev. D **60**, 54503 (1999); hep-lat/9903032.

Kostas Orginos and Doug Toussaint, "Testing improved actions for dynamical Kogut-Susskind quarks," Phys. Rev. D **59**, 14501 (1999).

High Precision Arithmetic and Applications

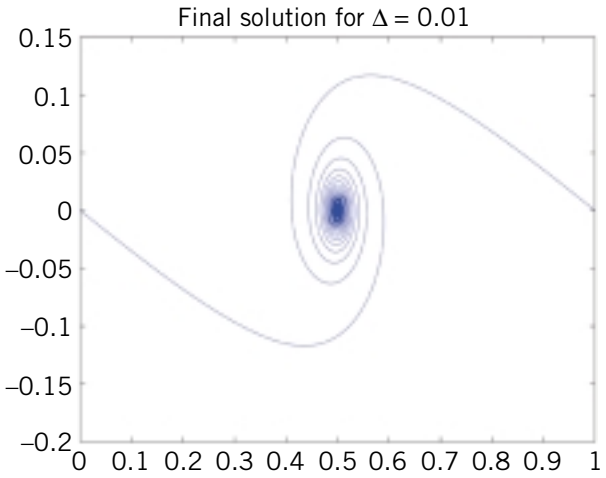
David Bailey and Xiaoye Sherry Li, NERSC,
Lawrence Berkeley National Laboratory

RESEARCH OBJECTIVES

We are working to develop high-precision arithmetic software. One of the key applications that we wish to pursue is a vortex roll-up simulation we developed utilizing a new "quad-double" arithmetic package written by Yozo Hida. In addition, we plan to pursue applications of the PSLQ integer relation detection program.

COMPUTATIONAL APPROACH

This research employs advanced techniques for performing arithmetic with more than the standard 16-digit IEEE floating-point arithmetic that is available on most technical computers today. During the past year we have developed "double-double" and "quad-double" software packages, which enable ordinary C or Fortran computer programs to perform arithmetic with 32 and 64 decimal digit accuracy, respectively. In addition, we use a separate package, written by the PI, which performs arithmetic to an arbitrarily high level of numeric precision. Another key technique used is the PSLQ integer relation detection algorithm developed by the PI and mathematician-sculptor Helaman Ferguson of the Center for Computing Sciences in Maryland.



The result of a vortex sheet roll-up computation, performed on 256 processors of the IBM SP. An integro-differential equation that describes the sheet motion is solved by regularization with a smooth approximation. The smaller the smoothing parameter Δ, the more accurate is the solution. But severe roundoff errors occur with small values of Δ. In this computation, we use our newly developed quad-double (256 bits) arithmetic to overcome this difficulty.

ACCOMPLISHMENTS

During the past year five technical papers were completed based on calculations mostly performed using NERSC systems. Other accomplishments include the completion of two new extended precision software packages — a "double-double" package, which provides approximately 32 decimal digit accuracy, and a "quad-double" package, which provides approximately 64 decimal digit accuracy. These software packages also include bindings that permit ordinary C and Fortran programs to use these packages with only minor changes to the source code.

In addition, we have developed a new variant of the PSLQ integer relation detection algorithm that is suitable for highly parallel computer systems.

SIGNIFICANCE

Our vortex roll-up simulation research explores an unresolved question regarding the behavior of vortices — namely, whether they always form a nice exponential spiral. Until now, researchers in the field have assumed that this always happens, but our initial runs show that beyond a critical value of a certain parameter, the exponential spiral develops chaotic irregularities. We need to make more runs to firmly establish and better understand this phenomenon.

The PSLQ integer relation finding program explores relationships between constants that arise in certain fields of mathematics and physics. For example, PSLQ has unearthed a simple formula for calculating any binary digit of π without calculating the digits preceding it. In January 2000, the PSLQ algorithm was named one of ten "algorithms of the century" by the editors of Computing in Science and Engineering. We hope to uncover some new facts of mathematics and physics with this program.

PUBLICATIONS

David H. Bailey, "Integer relation detection," Computing in Science and Engineering **2**, 1 (2000).

Yozo Hida, Xiaoye S. Li, and David H. Bailey, "Quad-double arithmetic: Algorithms, implementation, and application," Lawrence Berkeley National Laboratory technical report LBNL-46996 (2000). Condensed version submitted to 15th IEEE Symposium on Computer Arithmetic.

Helaman R. P. Ferguson, David H. Bailey, and Stephen Arno, "Analysis of PSLQ, an integer relation finding algorithm," Mathematics of Computation **68**, 90 (1999).

http://www.nersc.gov/~dhbailey
http://www.nersc.gov/news/bailey1-20-00.html

Numerical Simulation of Turbulent Reacting Flows

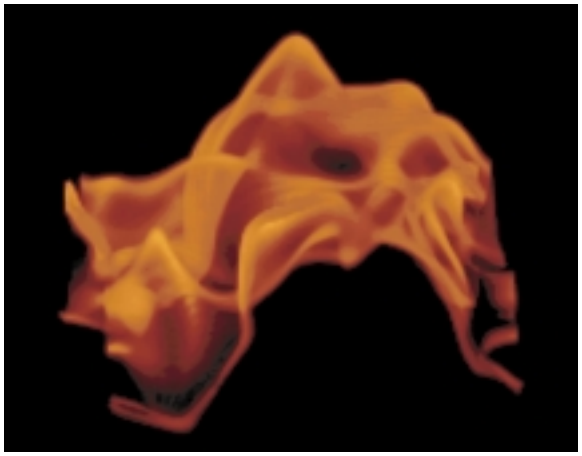
Center for Computational Sciences and Engineering and Applied Numerical Algorithms Group, NERSC, Lawrence Berkeley National Laboratory

RESEARCH OBJECTIVES

The goal of this project is to develop a high-fidelity numerical simulation capability for turbulent combustion processes such as those arising in furnaces and engines. The key issue in modeling turbulent combustion is the interplay between kinetics and the small-scale turbulent eddies in the flow. In practical, engineering combustion settings, the fuel typically consists of hydrocarbons whose chemical behavior is described by a complex reaction mechanism. Our objective is to develop and validate models for turbulent reacting flow that can accurately represent both the chemical and fluid-mechanical behavior of combusting hydrocarbons in a turbulent environment.

COMPUTATIONAL APPROACH

The principal computational tool for this project is the low Mach number adaptive mesh refinement (AMR) algorithm developed by CCSE. This methodology provides an accurate and efficient approach for modeling reacting flows in the regime that is appropriate for engineering applications. The algorithm uses a fractional step discretization that easily facilitates the inclusion of complex kinetics mechanisms. The methodology uses a block-structured refinement approach that allows computational effort to be focused in regimes of the flow where it is required. The structured refinement approach provides a natural coarse-grained parallelism that has demonstrated excellent performance and scalability on distributed memory architectures.



AMR simulation of developing flame surface, as turbulence flowing in from the lower boundary begins to impinge on the initially flat premixed hydrogen flame.

ACCOMPLISHMENTS

The focus of our research in the past year was on extending our low Mach number adaptive combustion algorithm to incorporate complex kinetics mechanisms and differential diffusion effects. We used the resulting methodology to address several open questions in combustion science. One area of study concerned the interaction of vortical structures with premixed hydrogen and methane flames. Here we focused the analysis on how the time-dependent flow fields associated with a counter-rotating vortex pair in two dimensions, and isotropic turbulence in three dimensions, modify the structure of the premixed flame. A second area of research focused on quantifying the production of nitrous oxides in a diffusion flame as a function of fuel composition (methane + additives). For both cases, the computations use 30–70 species and several hundred reactions to describe methane chemistry.

SIGNIFICANCE

The modeling of turbulent fluid flow, even in the non-reacting case, remains one of the great scientific challenges. We still lack adequate predictive models for non-reacting turbulent flows in realistic engineering geometries. For realistic combustion scenarios, the picture becomes more complex because small-scale turbulent fluctuations modify the physical processes such as kinetics and multiphase behavior. These processes, in turn, couple the small scales back to the larger fluid-dynamical scales as chemical constituents react. As a result of this coupling, we must capture the structure of the subgrid fluctuations to make predictions. The use of average quantities as inputs to physical processes will generate large errors through interaction of these models. Developing techniques that accurately reflect the role of small-scale fluctuations on the overall macroscopic dynamics would represent a major scientific breakthrough.

PUBLICATIONS

M. S. Day and J. B. Bell, "Numerical simulation of laminar reacting flows with complex chemistry," Lawrence Berkeley National Laboratory report LBNL-44682 (1999).

J. B. Bell, N. J. Brown, M. S. Day, M. Frenklach, J. F. Grcar, and S. R. Tonse, "The effect of stoichiometry on vortex flame interactions," Lawrence Berkeley National Laboratory report LBNL-44730 (1999).

J. B. Bell, N. J. Brown, M. S. Day, M. Frenklach, J. F. Grcar, R. M. Propp, and S. R. Tonse, "Scaling and efficiency of PRISM in adaptive simulations of turbulent premixed flames," in *Proceedings of the 28th International Combustion Symposium* (2000); Lawrence Berkeley National Laboratory report LBNL-44732 (1999).

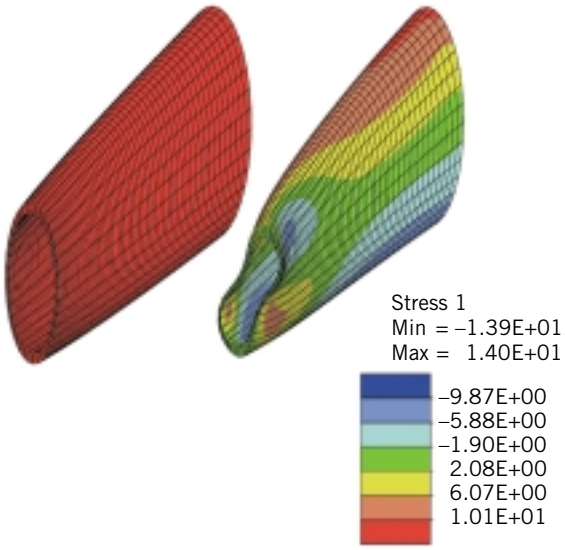
http://www.seesar.lbl.gov/ccse/

Linear Algebra Algorithms on High Performance Computers

James Demmel, University of California, Berkeley, and NERSC, Lawrence Berkeley National Laboratory
Mark Adams, Sandia National Laboratories
Xiaoye Sherry Li and David Blackston, NERSC, Lawrence Berkeley National Laboratory

RESEARCH OBJECTIVES

We provide highly optimized parallel computational kernels for DOE and other scientists. Our projects include a scalable sparse direct linear system solver (SuperLU), a scalable sparse incomplete factorization preconditioner (ILU), a scalable multigrid solver for partial differential equations on irregular meshes (Prometheus), a scalable symmetric eigensolver and singular value decomposition solver (xSTEVR), a scalable N-body code based on the fast multipole method and the Barnes-Hut algorithm (PBody), and a scalable structured matrix solver for matrices arising in astrophysical calculations.



Model of a truncated cone, a mesh of linear hexahedral elements with 21,600 degrees of freedom, fixed at the base and loaded at the end with a twisting load. This model was used in an evaluation of unstructured multigrid methods for 3D finite element problems in solid mechanics.

COMPUTATIONAL APPROACH

All codes are written with performance and portability in mind. Our codes are written in C or Fortran and use standard libraries such as BLAS, MPI, BSP communications, (Par)Metis, PETSc, etc. Since our goal is high parallel efficiency, the codes use state-of-the-art algorithms, many of which we designed.

ACCOMPLISHMENTS

SuperLU was used by Rescigno et al. (see Publications below) in a breakthrough quantum mechanical computation done on the NERSC Cray T3E and featured on the cover of the December 24, 1999 issue of *Science*.

Prometheus had a second release, incorporating both aggregation and smoothed aggregation methods in its collection of restriction operators. It was used to solve a 78 million degree of freedom problem on about 1,000 processors. PBody was completed, and David Blackston began working for NERSC to incorporate PBody into the ACTS Toolkit. xSTEVR was included as part of the LAPACK 3.0 release, and is significantly faster than the previous symmetric eigensolver.

SIGNIFICANCE

We are developing computational tools for linear algebra and N-body problems that are ubiquitous in computational science and engineering. All codes will be publicly available.

PUBLICATIONS

T. N. Rescigno, M. Baertschy, W. A. Isaacs, and C. W. McCurdy, "Collisional breakup in a quantum system of three charged particles," Science **286**, 2474 (1999).

M. Adams, "Evaluation of three unstructured multigrid methods on 3D finite element problems in solid mechanics," Computing Sciences Division Technical Report CSD-00-1103, University of California, Berkeley (1999).

E. Anderson et al., *LAPACK Users' Guide*, 3rd edition (Society for Industrial and Applied Mathematics, Philadelphia, 1999).

http://www.cs.berkeley.edu/~{demmel,davidb,xiaoye,madams}

A Numerical Study of Acceleration-Driven Fluid Interface Instabilities

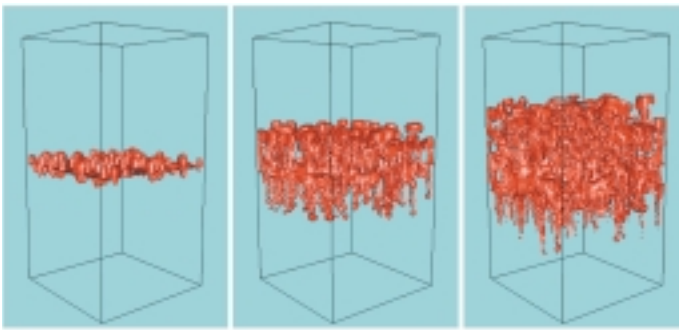
James Glimm and Xiaolin Li,
State University of New York, Stony Brook

RESEARCH OBJECTIVES

We will conduct definitive simulations of two types of acceleration-driven fluid mixing: the steady acceleration-driven Rayleigh-Taylor (RT) instability, and the shock-driven Richtmyer-Meshkov (RM) instability. We will also study the impulsive acceleration (shock)-driven RM instability.

COMPUTATIONAL APPROACH

We use the front tracking method to study the RT and RM instabilities. Recently we have implemented a robust grid-based method to handle the interface geometry. The traditional front tracking, known in contrast as grid-free tracking, also has advantages, in controlling the quality of the interface elements (triangles) and refined interface meshing. A hybrid combination of the two methods is best suited for study of both the RT and RM instabilities. The use of grid-free tracking at the initial stage of both problems gives an accurate startup of the problem. Grid-based tracking can handle the late-time chaotic stage of the fluid interface mixing without difficulty.



$Agt^2 = 4.48$ $Agt^2 = 10.08$ $Agt^2 = 15.75$

FronTier simulation of Rayleigh-Taylor instability with random initial perturbation: $\rho_H = 3$, $\rho_L = 1$, $g = 0.14$, $p = 1$, $\gamma_H = \gamma_L = 1.667$. Computational domain: $2 \times 2 \times 4$, computational grid: $112 \times 112 \times 224$, parallel partition: $8 \times 8 \times 1$. The acceleration rate of the bubble envelope: $\alpha = \{h_B/Agt^2\} = 0.075$.

We have implemented front tracking in a software package known as FronTier. This code is portable to various parallel computational platforms. Another code, the TVD level set code, uses the untracked numerical scheme for the simulation of fluid interface instabilities. This code is easily vectorized and efficient. We use it as for scientific comparison with the FronTier code.

ACCOMPLISHMENTS

The simulations in the past year study the Rayleigh-Taylor instability with randomly perturbed fluid interface. We have studied the growth rate under the variation of initial perturbation spectra, compressibility, and the growth rate in late-time chaotic mixing. In these studies, our numerical results are consistently closer to, or slightly larger than, the experimental value, in contrast to the results obtained by several other simulations. We believe that the difference is due to numerical diffusion in those simulations where fluid interface is not tracked. As a comparison, we also performed a simulation of the same problem using our own untracked TVD code. The comparison confirmed our conjecture.

SIGNIFICANCE

Acceleration-driven fluid mixing instabilities play important roles in inertially confined nuclear fusion and in stockpile stewardship. Turbulent mixing is a difficult and centrally important issue for fluid dynamics, and impacts such questions as the rate of heat transfer by the Gulf Stream, resistance of pipes to fluid flow, combustion rates in automotive engines, and the late time evolution of a supernova. Our computational study will provide a better understanding of the development of these instabilities.

PUBLICATIONS

B. Cheng, J. Glimm, X. L. Li, and D. H. Sharp, "DNS simulations and subgrid models for fluid mixing," in *Proc. 7th Int. Conf. on the Physics of Compressible Turbulent Mixing* (St. Petersburg, 1999).

J. Glimm, J. Grove, X. L. Li, W. Oh, and D. Sharp, "A critical analysis of Rayleigh-Taylor growth rates," J. Comp. Phys. (submitted).

J. Glimm, J. Grove, X. L. Li, and D. C. Tan, "Robust computational algorithms for dynamic interface tracking in three dimensions," SIAM J. Sci. Comp. (submitted).

http://www.ams.sunysb.edu/~shock/FTdoc/FTmain.html

Global Optimization Approaches to Protein Fold Refinement and Tertiary Structure Prediction for *D. Radiodurans*

Teresa Head-Gordon and Silvia Crivelli,
Lawrence Berkeley National Laboratory

RESEARCH OBJECTIVES

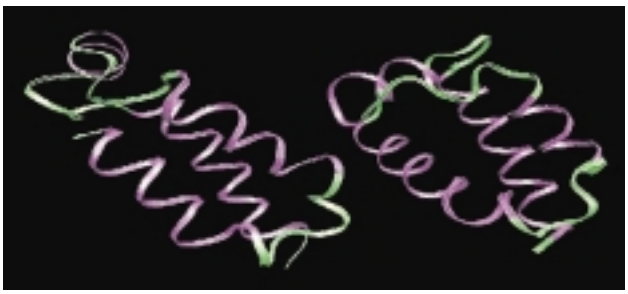
We will integrate computational techniques developed by our team of researchers to explore the basis for survival of *Deinococcus radiodurans* (DR) under extreme conditions of DNA damage. We will construct molecular models of the key components of the DNA repair system, including DNA binding proteins, which comprise one of the largest families of genes in DR; DNA repair proteins that recognize DNA damage for nucleotide excision repair, as well as recombinational repair proteins; and higher-order complexes such as the long-patch mismatch repair system (MMR). Our component of this research is structure-based tertiary structure prediction of individual proteins of DR.

COMPUTATIONAL APPROACH

We are focusing on the following three areas: large-scale global optimization algorithms, construction of accurate energy models of proteins, and large-scale parallel computation. Key features of our research include new approaches to smoothing the energy model to increase the efficiency of the global optimization; incorporation of secondary structure prediction information as non-binding constraints in the global optimization process; and a novel representation of aqueous solvent that energetically differentiates the desired structure as the lowest energy conformer relative to other misfolds.

ACCOMPLISHMENTS

This project, a collaboration between leading research groups in numerical optimization and computational biochemistry, has resulted in advances in both large-scale and parallel numerical computation, and in models and prediction methods for computational biology. Our global optimization methods require use of parallel computation, and we have explored scheduling, task migration, and load balancing to increase efficiency. We have obtained state-of-the-art results in predicting helical proteins of about 70 amino acids, and have developed robust energy models involving solvation. We have recently begun the prediction of β-sheet and mixed α/β proteins in the CASP4 competition.



A comparison between the NMR structure of a four helix bundle DNA binding protein, 1pou (right), and the outcome from our global optimization algorithm (left). Purple indicates helical regions, while green indicates residues that are coil.

SIGNIFICANCE

For the first time, experimental data acquisition arising from high-throughput methods of genome sequencing, structural genomics, and gene-expression assays provides the information needed to understand the molecular and cellular decision-making of processes such as DNA repair. The analysis and synthesis of rapidly accumulating data requires computational and theoretical tools that remain to be integrated or further developed, which is one goal of this research.

PUBLICATIONS

A. Azmi, R. H. Byrd, E. Eskow, R. Schnabel, S. Crivelli, T. M. Philip, and T. Head-Gordon, "Predicting protein tertiary structure using a global optimization algorithm with smoothing," in *Optimization in Computational Chemistry and Molecular Biology: Local and Global Approaches*, edited by C. A. Floudas and P. M. Pardalos (Kluwer Academic Publishers, Netherlands, 2000).

J. M. Sorenson, G. Hura, A. K. Soper, A. Pertsemlidis, and T. Head-Gordon, "Determining the role of hydration forces in protein folding," J. Phys. Chem. B **103**, 5413 (1999).

J. M. Sorenson and T. Head-Gordon, "Matching simulation and experiment: A new simplified model for simulating protein folding," J. Comp. Bio. (in press).

http://www.lbl.gov/~thg

Electron-Atom and Electron-Molecule Collision Processes

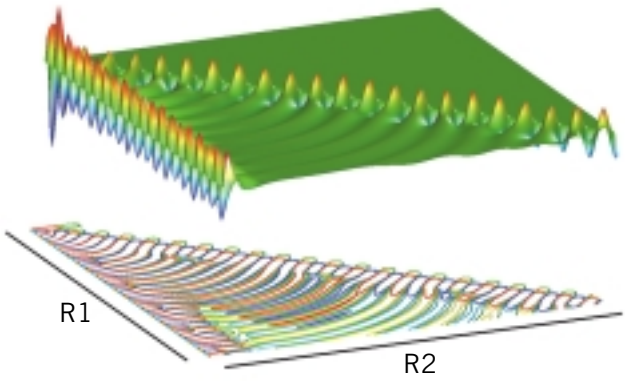
C. William McCurdy, Thomas N. Rescigno, and Daniel A. Horner,
Lawrence Berkeley National Laboratory
William A. Isaacs, Lawrence Livermore National Laboratory
Mark Baertschy, JILA, University of Colorado

RESEARCH OBJECTIVES

This project seeks to develop theoretical and computational methods for treating electron collision processes that are currently beyond the grasp of first-principles methods. We are developing methods for studying electron-atom and electron-molecule collisions at energies above that required to ionize the target, for calculating detailed electron impact ionization probabilities for simple atoms and molecules, and for treating low-energy electron collisions with polyatomic molecules, complex molecular clusters, and molecules bound to surfaces and interfaces.

COMPUTATIONAL APPROACH

We are developing a new generation of electron-atom and electron-molecule scattering codes. These codes now include complex optical potential interactions, a scattered-wave/flux operator formalism, and a variety of techniques based on analyticity. Our new FEM/DVR discretization method combines finite elements with a discrete variable representation based on the use of Gauss-Lobatto quadrature, thus offering the best features of both methods (sparse matrices and simplicity of representation). We have also been examining more efficient new integral methods for computing breakup cross sections. These developments will allow us to extend our ionization studies to two-electron targets and to treat a broad range of problems, from low-energy electron-molecule collisions using elaborate variational wave functions, to direct solutions of the Schrödinger equation for simple atomic targets that provide detailed ionization cross sections.



Wave function for a model 2D simulation of positron impact on atomic hydrogen. A 20 eV positron is incident along the r1 axis. The radial waves in the sector r1 > r2 correspond to ionization, leaving a bare proton plus a free electron and a free positron, which carries off most of the energy. The concentration of density along the r1 = r2 axis corresponds to rearrangement events that produce bound positronium.

ACCOMPLISHMENTS

The methods we developed have been shown capable of treating all details of electron impact ionization of atoms, including energy sharing differential cross sections and triply differential cross sections (energy and solid angles for both ejected electrons).

We have finished the first computational study of low-energy electron-CO_2 scattering that successfully reproduced the two features that dominate the low-energy cross sections, namely, the dramatic rise in the elastic cross section below 2.0 eV and the resonance peak centered near 3.8 eV. We initiated the second phase of work on this system, which is to further explore the effects of nuclear motion on the low-energy scattering cross sections. We have found that both symmetric stretch and bending motion are crucial in determining accurate resonant vibrational excitation cross sections and in understanding the nature of the low-energy virtual state enhancement of the elastic and momentum transfer cross section. We have begun to carry out multi-dimensional time-dependent wavepacket calculations using the results of our electronic fixed-nuclei electron scattering calculations to quantify the nuclear dynamics and to compute the vibrational excitation cross sections.

SIGNIFICANCE

Electron collision processes play a key role in such diverse areas as fusion plasmas, plasma etching of silicon chips, and mixed radioactive waste remediation. Understanding of these processes is severely hampered by the lack of a database of electron-molecule collision cross sections. This project will significantly add to that base of knowledge.

PUBLICATIONS

T. N. Rescigno, M. Baertschy, W. A. Isaacs, and C. W. McCurdy, "Collisional breakup in a quantum system of three charged particles," Science **286**, 2474 (1999).

T. N. Rescigno, D. A. Byrum, W. A. Isaacs, and C. W. McCurdy, "Theoretical studies of low-energy electron-CO_2 scattering: Total, elastic and differential cross sections," Phys. Rev. A **60**, 2186 (1999).

C. W. McCurdy and T. N. Rescigno, "Practical calculations of quantum breakup cross sections," Phys. Rev. A **62**, 32712 (2000).

http://jolt.lbl.gov

Metastability in Materials Science and Scalability of Parallel Discrete Event Simulations

Mark Novotny, Per Arne Rikvold, Gregory Brown, and Kyungwha Park, Supercomputer Computations Research Institute, Florida State University
Gyorgy Korniss, Rensselaer Polytechnic Institute
Vladimir Antropov, Ames Laboratory, Iowa State University

RESEARCH OBJECTIVES

(1) Test the scalability of parallel discrete event simulations (PDES) on massively parallel computers. (2) Perform large-scale dynamic Monte Carlo simulations related to hysteresis, metastbility, and dynamics of domain-wall motion in nanoscale magnetic systems.

COMPUTATIONAL APPROACH

We have implemented Lubachevsky's partially rejection-free method for PDES and used this code to obtain large simulations of hysteresis in Ising models to allow us to obtain nonequilibrium dynamic exponents using finite size scaling techniques. For the study of metastable systems with discrete spins, we utilize the Monte Carlo with absorbing Markov chains algorithm and projective dynamics methods. We include the broad histogram method to enable us to apply these types of algorithms to systems with continuous spins, such as the classical Heisenberg model. We use these methods in stochastic differential equations, in particular Langevin micromagnetic calculations.

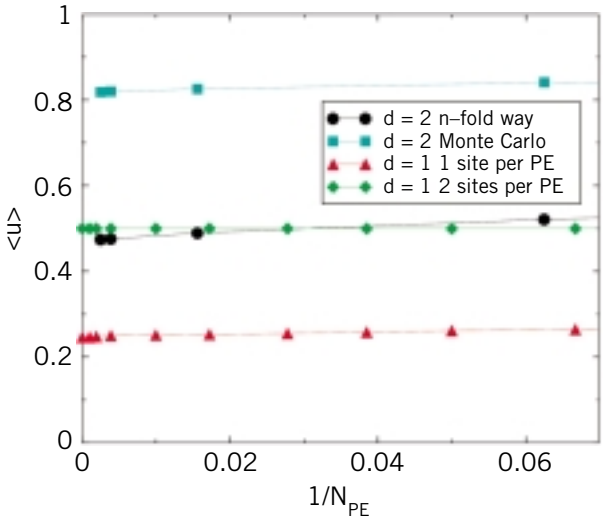
ACCOMPLISHMENTS

We proved that all conservative implementations of PDES in one dimension ($d = 1$) are scalable in the calculation phase. We numerically tested our proof of scaling in $d = 1$, and provided numerical evidence for scaling in $d = 2$ and $d = 3$. We used our PDES to perform dynamic Monte Carlo simulations for thermal switching of nanoscale magnets and to perform finite-size scaling for a dynamic phase transition for magnetic thin films. Our dynamic Monte Carlo algorithms allow simulations that are true to the underlying physical dynamic and span very large time scales, from atomistic times to engineering times.

SIGNIFICANCE

We proved for the first time that a nontrivial parallelization method is scalable. Since DES are used in a wide variety of fields (from switching of cellular communications networks to war-game simulations), this general proof is of broad interest to researchers in various fields.

Metastability and hysteresis are ubiquitous in materials systems, ranging from ferromagnets to aging and failure of materials. We study both simple models where the physics can be fully explored and understood, and realistic calculations for models of actual systems (nanoscale ferromagnets). The time scales range from microscopic to geologic. To span these disparate time scales, we have introduced novel algorithms that can gain many orders of magnitude in simulation speed, and we have implemented an efficient code for stochastic simulation on a distributed-memory machine, where the pattern of communication between the underlying PEs is completely unpredictable.



The average utilization, $\langle u \rangle$, in a nearest neighbor PDES simulation as a function of N_{PE}^{-1}. One-dimensional simulations with one lattice site per PE (red triangles) and two lattice sites per PE (green diamonds). Two-dimensional Monte Carlo (cyan squares) and *n*-fold way (black circles) simulations with each PE having a block of lattice sites 128 on a side.

PUBLICATIONS

G. Korniss, Z. Toroczkai, M. A. Novotny, and P. A. Rikvold, "From massively parallel algorithms and fluctuating time horizons to non-equilibrium surface growth," Phys. Rev. Lett. **84**, 1351 (2000).

G. Korniss, M. A. Novotny, and P. A. Rikvold, "Parallelization of a dynamic Monte Carlo algorithm: A partially rejection-free conservative approach," J. Comp. Phys. **153**, 488 (1999).

M. A. Novotny, "A tutorial on advanced dynamic Monte Carlo methods for systems with discrete state spaces," in *Annual Reviews of Computational Physics IX*, edited by D. Stauffer (World Scientific, Singapore, in press).

http://www.csit.fsu.edu/~novotny

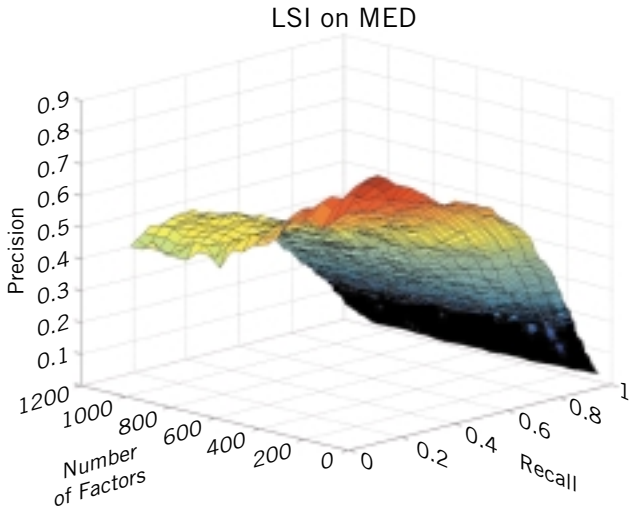
Sparse Linear Algebra Algorithms and Applications for MPPs

Horst Simon, Chris Ding, and Parry Husbands, NERSC, Lawrence Berkeley National Laboratory
Hongyuan Zha, Pennsylvania State University
Inderjit Dhillon, University of Texas, Austin

RESEARCH OBJECTIVES

The goal of this research is to use high performance computing to help understand and improve subspace-based techniques for information retrieval, such as LSI (latent semantic indexing). In the past these techniques have been used only on very small data sets. We are developing an environment and scalable linear algebra algorithms with which LSI-based information retrieval, for example, can be applied to matrices representing millions of documents and hundreds of thousands of key words.

We are studying both retrieval accuracy and algorithmic efficiency. We hope to: (1) Further develop our subspace-based model and gain deeper understanding of the effectiveness of LSI. We will also explore the possibility of extending the subspace-based model for handling image and multimedia retrieval problems. (2) Develop more robust and computationally efficient statistical tests for determining the optimal latent-concept subspace dimensions. We will especially investigate methods based on cross-validation. (3) Further explore the low-rank-plus-shift structures of the term-document matrices and develop fast and memory-efficient numerical algorithms for the computation of low-rank matrix approximations. We will also investigate linear-time partial singular value decomposition (SVD) algorithms based on term and document sampling. (4) Investigate a variety of statistical methods such as canonical correlation analysis and generalized linear model for translingual text retrieval. (5) Investigate how the evolution of the link structure of the World Wide Web can augment retrieval schemes.



A graph showing the effects of changing the number of dimensions on the retrieval accuracy of LSI.

COMPUTATIONAL APPROACH

We plan to continue to explore the use of SVD and other low-rank approximations in information retrieval/data mining, not only of pure text collections, but of those augmented with hyperlinks (such as in the World Wide Web).

ACCOMPLISHMENTS

The major accomplishment of the past year was a deeper understanding of the representation produced by the SVD. We discovered that not all of the properties of this decomposition are desirable for accurate retrieval. This discovery came about through the analysis and visualization of a very large text collection using the T3E. We have also acquired a large data set of Web link data and have completed a preliminary analysis.

SIGNIFICANCE

This project is one of the first to compute decompositions of complete, large term-document matrices. Other efforts have had to resort to sampling to keep things computationally tractable. We are in a unique position to study how scale affects subspace-based retrieval techniques. We are also well poised to discover new and interesting features about the structure of the Web. The algorithms developed here will be of use not only in text retrieval, but in more complicated settings such as classification of image collections and extraction of images with desired features from large collections. Combining effective search and classification algorithms for image problems with the compute and storage capabilities of future NERSC systems will position Berkeley Lab in a leadership position when it comes to the development of algorithmic techniques for the data and visualization corridors of the next decade.

PUBLICATIONS

Parry Husbands, Charles Isbell, and Alan Edelman, "MATLAB*P: A tool for interactive supercomputing," in *Proceedings of the Ninth SIAM Conference on Parallel Processing for Scientific Computing*, edited by Bruce Hendrickson et al. (Society for Industrial and Applied Mathematics, Philadelphia, 1999).

Parry Husbands, Horst Simon, and Chris Ding, "On the use of the singular value decomposition for text retrieval" (in preparation).

http://www.nersc.gov/research/SIMON

APPENDIX A

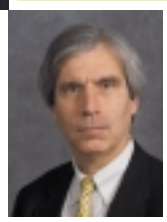
NERSC Policy Board

Bastiaan J. Braams
(ex officio, NERSC Users Group Chair 2000)
Courant Institute of Mathematical Sciences
New York University

Robert J. Goldston
Princeton Plasma Physics Laboratory

Fred E. Cohen
Department of Medicine
University of California, San Francisco

Paul C. Messina
Defense Programs
U.S. Department of Energy
Washington, DC



Stephen Jardin
(ex officio, Chair of Program
Advisory Committee)
Princeton Plasma Physics Laboratory

Albert Narath
Lockheed-Martin
Corporation (retired)

Robert D. Ryne
(ex officio, NERSC Users Group Chair 2001)
Los Alamos National Laboratory

Larry L. Smarr
University of California, San Diego

Michael Witherell
Fermi National Accelerator
Laboratory

APPENDIX B

NERSC Program Advisory Committee (PAC)

Stephen Jardin
PAC Chair
Princeton Plasma Physics Laboratory

David Bailey
NERSC Liaison

Ian Foster
Computer Science
Argonne National Laboratory

Steven Hammond
Environmental Sciences
National Center for Atmospheric Research

Bruce Harmon
Material Sciences
Ames Laboratory

Robert Harrison
Chemistry
Pacific Northwest National Laboratory

Dale Koelling
Material Sciences
Argonne National Laboratory

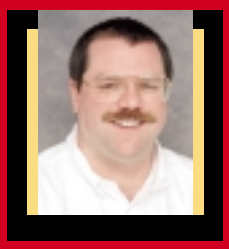
Jean-Noel Leboeuf
Fusion Energy
University of California, Los Angeles

Gregory Newman
Earth and Engineering Sciences
Sandia National Laboratories

Robert Ryne
Accelerator Physics
Los Alamos National Laboratory

Shankar Subramaniam
Life Sciences
University of California, San Diego

Robert Sugar
High Energy Physics
University of California, Santa Barbara



Douglas Swesty
Nuclear and Astrophysics
State University of New York, Stony Brook

Mary Wheeler
Applied Mathematics
University of Texas

APPENDIX C

NERSC Users Group Executive Committee

MEMBERS AT LARGE

Bastiaan J. Braams (NUG Chair, 2000), Courant Institute of Mathematical Sciences, New York University
Brian E. Hingerty (NUG Vice Chair, 2000), Oak Ridge National Laboratory
Ahmet Y. Aydemir, University of Texas at Austin
Ricky A. Kendall, Ames Laboratory, Iowa State University
Gregory Kilcup, The Ohio State University
Douglas Olson, Lawrence Berkeley National Laboratory

OFFICE OF ADVANCED SCIENTIFIC COMPUTING RESEARCH

Ahmed Ghoniem, Massachusetts Institute of Technology
Xiaolin Li, State University of New York, Stony Brook
Michael Minkoff, Argonne National Laboratory

OFFICE OF BASIC ENERGY SCIENCES

David Seidman, Northwestern University
G. Malcolm Stocks, Oak Ridge National Laboratory
Theresa L. Windus, Pacific Northwest National Laboratory

OFFICE OF BIOLOGICAL AND ENVIRONMENTAL RESEARCH

Carmen M. Benkovitz, Brookhaven National Laboratory
Douglas Rotman (NUG Vice Chair, 2001), Lawrence Livermore National Laboratory
Vincent Wayland, National Center for Atmospheric Research

OFFICE OF FUSION ENERGY SCIENCES

Stephen C. Jardin, Princeton Plasma Physics Laboratory
Jean-Noel Leboeuf, University of California, Los Angeles
Donald A. Spong, Oak Ridge National Laboratory

OFFICE OF HIGH ENERGY AND NUCLEAR PHYSICS

David Dean, Oak Ridge National Laboratory
Kwok Ko, Stanford Linear Accelerator Center
Robert D. Ryne (NUG Chair, 2001), Los Alamos National Laboratory

APPENDIX D

Scientific Computing Applications Council

The Scientific Computing Applications Council (SCAC) is responsible for setting the policies associated with the utilization of computing, data storage, and other auxiliary services available to DOE Office of Science (SC) researchers and otherwise coordinating SC's computational projects. The Council sets the distribution of NERSC and other available Office of Advanced Scientific Computing Research resources for scientific computing applications every year and reviews the operation of NERSC every two years.

COUNCIL CHAIRMAN	David Goodwin
EXECUTIVE CHAIR	Thomas Kitchens
EXECUTIVE ASSISTANT	Linda Twenty
OFFICE OF BASIC ENERGY SCIENCES	Robert Astheimer • William Kirchhoff
OFFICE OF HIGH ENERGY AND NUCLEAR PHYSICS	David Goodwin • Victoria White
OFFICE OF ADVANCED SCIENTIFIC COMPUTING RESEARCH	Bruce Hendrickson • Thomas Kitchens Linda Twenty
OFFICE OF FUSION ENERGY SCIENCES	Michael Crisp • John Willis
OFFICE OF BIOLOGICAL AND ENVIRONMENTAL RESEARCH	Michael Riches • David Bader

APPENDIX E

Office of Advanced Scientific Computing Research

The primary mission of the Advanced Scientific Computing Research (ASCR) program, which is carried out by the Mathematical, Information, and Computational Sciences (MICS) subprogram, is to discover, develop, and deploy the computational and networking tools that enable researchers in the scientific disciplines to analyze, model, simulate, and predict complex phenomena important to the Department of Energy. To accomplish this mission, the program fosters and supports fundamental research in advanced scientific computing — applied mathematics, computer science, and networking — and operates supercomputer, networking, and related facilities. In fulfilling this primary mission, the ASCR program supports the Office of Science Strategic Plan's goal of providing extraordinary tools for extraordinary science as well as building the foundation for the research in support of the other goals of the strategic plan. In the course of accomplishing this mission, the research programs of ASCR have played a critical role in the evolution of high performance computing and networks.

As the flagship facility of the ASCR program, NERSC thanks the program managers with direct responsibility over the NERSC Division:

C. Edward Oliver
Associate Director, OASCR, and Acting MICS Division Director

Thomas Kitchens
NERSC and Applications Program Manager (2000)

Walt Polansky
NERSC and Applications Program Manager (2001)

Bruce Hendrickson
Applied Mathematics

Daniel Hitchcock
Computer Science Research

Frederick Johnson
Computer Science Research

Mary Anne Scott
Distributed Computing and Grid Technologies

Linda Twenty
OASCR Program Analyst

Projects in the NERSC Division also received support from the Office of Biological and Environmental Research. We acknowledge the support of David Bader, Climate Modeling, and Daniel Drell, Bioinformatics.

Index of Researchers

The researchers listed below produced the scientific results highlighted in this report.

Published by the Berkeley Lab
Technical and Electronic Information Department
in collaboration with NERSC researchers and staff.
J0#2259