BROOKHAVEN
NATIONAL LABORATORY

BNL-72286-2005

Computational Science Center

in affiliation with the Departments of

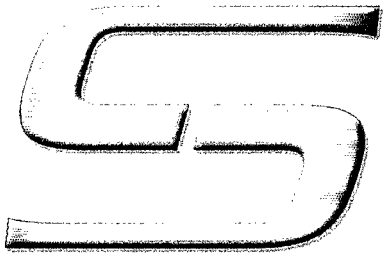
Applied Mathematics and Statistics and
Computer Science
Stony Brook University

www.bnl.gov/csc

November 2005

Brookhaven National Laboratory
P.O. Box 5000
Upton, NY 11973-5000
www.bnl.gov

DISCLAIMER



Computational Science Center

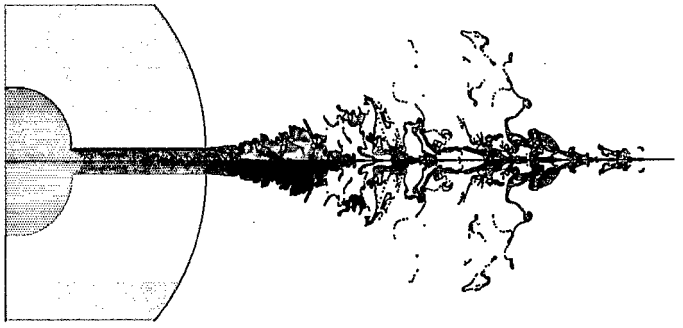
Brookhaven National Laboratory

in affiliation with the Departments of

Applied Mathematics and Statistics and Computer Science

Stony Brook University

www.bnl.gov/csc



Cover: Breakup of diesel fuel into droplets following injection in a diesel engine. See p. 24.

BROOKHAVEN COMPUTATIONAL SCIENCE CENTER

Page

1. Introduction 3

2. Overview 3

3. Personnel 4

4. Research 7

4.1 SciDAC at CSC 8
- 4.1.1 TSTT 8
- 4.1.2 TOPS 8
- 4.1.3 Advanced Computing for 21st Century Accelerator Science & Technology 9

4.2 Biological and Environmental Research 9
- 4.2.1 MD Simulations of the Activation of the Adenovirus Proteinase 10
- 4.2.2 Structural Analysis of Botulinum Neurotoxin 12
- 4.2.3 Visualization and Data Mining for the Detection of Cancer 14
- 4.2.4 Structure-based Drug Design Targeting HIV 17
- 4.2.5 Cardiac Electrophysiology 19
- 4.2.6 Atmospheric Aerosols 21

4.3 Advanced Scientific Computing 24
- 4.3.1 TSTT: Atomization and Spray 24
- 4.3.2 TSTT: Frontier-Lite 27
- 4.3.3 Magnetohydrodynamics of Free Surface Flows at Low Magnetic Reynolds Numbers 30
- 4.3.4 Molecular Dynamics on the QCDOC 31
- 4.3.5 Pellet Ablation in Tokamak Refueling Process 32
- 4.3.6 Parallel Heisenberg Spin Model on Supercomputer Architectures 34
- 4.3.7 Uncertainty Quantification 37
- 4.3.8 ViStA: Visual Statistical Analyzer 39
- 4.3.9 Parallel Visualization of Large Data Sets 43
- 4.3.10 Visualization Collaboratory 45
- 4.3.11 Cluster Computing 47
- 4.3.12 Electrodynamics on QCDOC and BG/L 48
- 4.3.13 Three Dimensional FFT on QCDOC 49

4.4 Basic Energy Sciences 50
- 4.4.1 Superconductivity 50
- 4.4.2 Scalable Localizable Density Functional Theory 52
- 4.4.3 Direct Numerical Simulation of Multiphase Flows with Phase Transitions 54
- 4.4.4 Electrodynamics Simulation: Photonic Devices and RF Cavity Design 55
- 4.4.5 DFT Study of Isocyanide Adsorption on Gold (111) Surface 57

4.5 High Energy and Nuclear Physics 58
- 4.5.1 Neutrino Factory/Muon Collider Target 59
- 4.5.2 Modeling of Wake Fields and Impedances in Accelerators 61
- 4.5.3 Unified Accelerator Library: SIMBAD 63
- 4.5.4 Spallation Neutron Source 65
- 4.5.5 Grid Computing: MonALISA 66

1. INTRODUCTION

The Brookhaven Computational Science Center brings together researchers in biology, chemistry, physics, and medicine with applied mathematicians and computer scientists to exploit the remarkable opportunities for scientific discovery which have been enabled by modern computers.

These opportunities are especially great in computational biology and nanoscience, but extend throughout science and technology and include, for example, nuclear and high energy physics, astrophysics, materials and chemical science, sustainable energy, environment, and homeland security.

To achieve our goals we have established a close alliance with applied mathematicians and computer scientists at Stony Brook and Columbia Universities.

2. OVERVIEW

Science has become increasingly dependent on computers and computing. New machines, with 100,000 processors or more have made possible the vision of simulation as the 'third tier' of science, on an equal footing with experiment and traditional theory. Many experiments produce so much data and of such complexity that they require extensive data collection, storage and analysis capabilities in order to be successful.

In recognition of these trends, DOE's Office of Science has embarked on a long-range program to upgrade the computing infrastructure within the Laboratory complex and to make the most advanced systems available to researchers throughout the scientific community.

Brookhaven currently has major efforts in data storage and analysis for the Relativistic Heavy Ion Collider (RHIC) and the ATLAS detector at CERN, and in quantum chromodynamics. The Laboratory is host for the QCDOC machines (quantum chromodynamics on a chip) 10 teraflop/s computers which boast 12,288 processors each. There are two here, one for the Riken/BNL Research Center and the other supported by DOE for the US Lattice Gauge Community and other scientific users.

Our Center has established a strong program in computational science, with an emphasis on nanoscale electronic structure and molecular dynamics, accelerator design, computational fluid dynamics, medical imaging, parallel computing and numerical algorithms. We have been an active participant in DOE's SciDAC program (Scientific Discovery through Advanced Computing).

In keeping with Laboratory initiatives we are planning a major expansion in computational biology and laying the groundwork for a large parallel machine such as IBM's BlueGene/L.

3. PERSONNEL

The Director of the Computational Science Center (CSC) is James Davenport; James Glimm is Head of Applied Mathematics, and the Administrative Assistant is Claire Lamberti. CSC is located in Building 463 (30 Bell Avenue), Room 255, at Brookhaven National Laboratory, Upton NY 11973-5000. The fax number for CSC is 631-344-5751.

CSC grew out of the Center for Data Intensive Computing, which was established in 1999. It currently has a staff of 16. As part of an active, ongoing collaboration with graduate students from Stony Brook, over thirty Ph.D. students and several undergraduates hold scientific positions in CSC.

TABLE 1. CSC STAFF

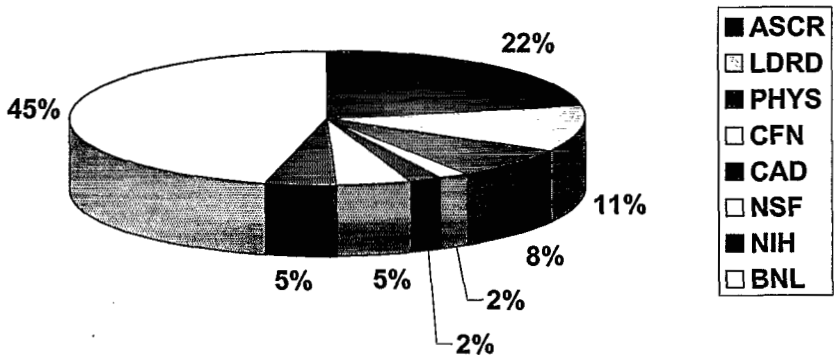
Name	Position	Phone/email
James Davenport	Director	631-344-3789 jdaven@bnl.gov
James Glimm	Head, Applied Mathematics	631-344-8155 glimm@bnl.gov
David Keyes	Visiting Scientist	631-344-5752 kd2112@columbia.edu
Nicholas D'Imperio	Physics Associate	631-344-8607/8589 dimperio@bnl.gov
Stratos Efstathiadis	Technology Architect	631-344-6080 stratos@bnl.gov
Michael McGuigan	Computer Scientist	631-344-2695 mcguigan@bnl.gov
Roman Samulyak	Associate Scientist	631-344-3304 rosamu@bnl.gov
Leonard Slatest	Advanced Technology Engineer	631-344-4102 slatest@bnl.gov
Gordon Smith	Technology Architect	631-344-3216 smith3@bnl.gov
John Spiletic	Technology Architect	631-344-4112 spiletic@bnl.gov
David Stampf	Senior Technology Architect	631-344-4148 drs@bnl.gov
Robert Bennett	Research Associate	631-344-2089 robertb@bnl.gov
Kab Seok Kang	Research Associate	631-344-3580 kskang@bnl.gov
Tianshi Lu	Research Associate	631-344-5752 tlu@bnl.gov
Zhiliang Xu	Research Associate	631-344-2089 xuzhi@bnl.gov
Claire Lamberti	Administrative Assistant	631-344-3051 lamberti@bnl.gov

TABLE 2. CSC AFFILIATES		
Name	**Affiliation**	**Expertise**
Carl Anderson	BNL/Biology	Protein Structures
Carmen Benkovitz	BNL/Environmental Sciences	Atmospheric Transport
Helene Benveniste	BNL/Directorate	Neuroscience
Kenneth Bowler	UKQCDOC	Lattice QCD
Natasa Bozovic	San Jose State University	Computational Science
Xiaodan Cai	Univ. of Science & Tech. of China	Computer Science
Michael Creutz	BNL/Physics	Lattice Gauge Theory
Yuefan Deng	SB/Applied Math. & Statistics	Molecular Dynamics
Steven Evans	Beth Israel Medical Center	Cardiac Physiology
Erwin George	SB/Applied Math & Statistics	Fluid Dynamics
David Green	SB/Applied Math & Statistics	Computational Biology
John Grove	Los Alamos National Lab.	Computational Science
Jiansheng Jiang	BNL/Biology	Protein Structures
Arie Kaufman	SB/Computer Science	Visualization
Richard Kenway	UK QCDOC	Lattice QCD
Jerome Lauret	BNL/Physics	Grid Computing
Xiaolin Li	SB/Applied Math. & Statistics	Computational Science
Brent Lindquist	SB/Applied Math. & Statistics	Porous Media; Image Analysis
Alfredo Luccio	BNL/AGS	Accelerator Design
Walter Mangel	BNL/Biology	Structural Biology
Glenn Martyna	IBM	Computational Materials Science
Edward McFadden	BNL/ITD	Parallel Hardware
James Muckerman	BNL/Chemistry	Combustion Chemistry
Klaus Mueller	SB/Computer Science	Visualization
Arnold Peskin	BNL/CSC & ITD	Visualization
John Pinezich	Advanced Acoustic Concepts	Computational Fluid Dynamics
Hong Qin	SB/Computer Science	Parallel Computing
Marwan Rasamny	Delaware State University	Computational Materials Science
John Reinitz	SB/Applied Math. & Statistics	Computational Biology
Robert Rizzo	SB/Applied Math. & Statistics	Computational Biology
Stephen Schwartz	BNL/Environmental Sciences	Aerosol and Climate Modeling
David Sharp	Los Alamos National Lab.	Computational Science
Carlos Simmerling	SB/Chemistry	Protein Structures
William Studier	BNL/Biology	Protein Structures
S. Swaminathan	BNL/Biology	Structural Biology
Michael Weinert	Univ. of Wisconsin/Milwaukee	Condensed Matter Physics
Torre Wenaus	BNL/Physics	High Energy and Nuclear Physics
Yan Yu	SB/Applied Math. & Statistics	Uncertainty Quantification
Jincheng Zheng	BNL/Materials Science	Computational Materials Science
Wei Zhu	SB/Applied Math. & Statistics	Statistics; Image Analysis

TABLE 3. STUDENT AFFILIATES	
Name	**Affiliation**
Joseph Aronson	SB/Instructional Computing
Janet Braunstein	SB/Applied Mathematics & Statistics
Noel Carrascal	SB/Chemistry
Serena Chan	Cornell/Statistics
Xin Chen	SB/Applied Mathematics & Statistics
Yongzhi Chen	SB/Applied Mathematics & Statistics
Yongjun Cheng	SB/Applied Mathematics & Statistics
Jian Du	SB/Applied Mathematics & Statistics
Srabasti Dutta	SB/Applied Mathematics & Statistics
Bin Fang	SB/Applied Mathematics & Statistics
Yuxiang Gao	SB/Applied Mathematics & Statistics
Youlia Guilman	SB/Physics
Paul Guzzardo	Dowling College
Guowen Han	SB/Applied Mathematics & Statistics
Martin Horstman	SB/Applied Mathematics & Statistics
Zhuying Huang	SB/Applied Mathematics & Statistics
Xiaomei Ji	SB/Applied Mathematics & Statistics
Donghung Kim	SB/Applied Mathematics & Statistics
Noha Lim	SB/Applied Mathematics & Statistics
Xingtao Liu	SB/Applied Mathematics & Statistics
Sudipto Mukherjee	SB/Applied Mathematics & Statistics
Chamainda Kalum Palandage	University of Connecticut
Nitin Pillai	SB/Applied Mathematics & Statistics
Tatiana Polischuk	SB/Applied Mathematics & Statistics
Peter Rissland	SB/ Applied Mathematics & Statistics
Julie Stern	SB/Chemistry
Bentley Strockbine	SB/Pharmacology
Tao Sun	SB/Applied Mathematics & Statistics
Dmitri Volja	SB/Physics
Shuqiang Wang	SB/Applied Mathematics & Statistics
Xuena Wang	SB/Applied Mathematics & Statistics
Perry Woo	SB/Chemistry
Xiangfeng Wu	SB/Applied Mathematics & Statistics
Bin Xu	SB/Applied Mathematics & Statistics
Yue Zhang	SB/Applied Mathematics & Statistics

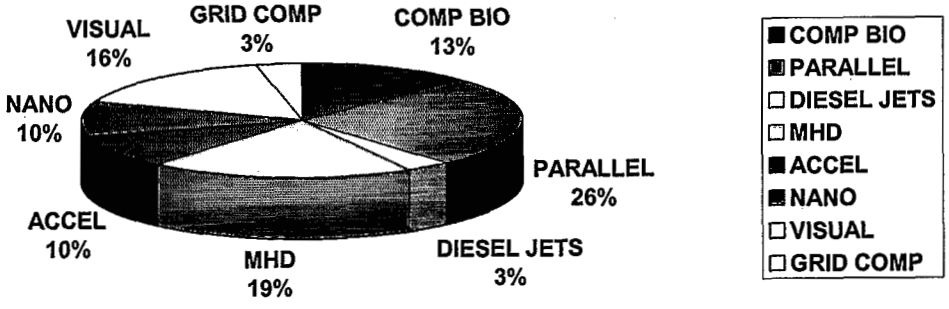
4. RESEARCH

CSC Budget FY 2005
Total: $2,520K



Funding by source. ASCR: DOE Office of Advanced Scientific Computing, LDRD: Laboratory Directed Research and Development, PHYS: BNL Physics, CFN: Center for Functional NanoMaterials, CAD: Collider-Accelerator Department, NSF: National Science Foundation, NIH: National Institutes of Health; BNL: Other BNL funding.

CSC PROJECTS



4.1 SciDAC at CSC

Scientific Discovery through Advanced Computing (SciDAC) is a five-year program to develop the Scientific Computing Software and Hardware Infrastructure to support terascale computations for DOE's research programs.

4.1.1 TSTT

J. Glimm, M.-N. Kim, and Z. Xu

At Brookhaven we have formed an alliance with researchers at Argonne, Livermore, Oak Ridge, Pacific Northwest, Rensselaer, Sandia, and Stony Brook, to develop technologies that enable scientists to use complex mesh and discretization strategies easily and interchangeably, within a single simulation on terascale computers. The Terascale Simulation Tools and Technologies (TSTT) Center provides interoperable tools to facilitate use of advanced mesh and discretization technologies. The Center has developed standardized interfaces to local mesh refinement codes. Existing codes of the Center partners and new codes are supported to create a plug and play capability, whereby an application user can experiment easily with alternative technologies. Insertion of these tools into targeted applications, including fusion, accelerator design, and climate, is part of the Center work plan.

CSC work within TSTT currently focuses on jet breakup, spray formation, and on finite element and front tracking contributions to TSTT technology.

The TSTT website is http://www.tstt-scidac.org.

4.1.2 TOPS

D. Keyes

Terascale Optimal PDE Simulations (TOPS) is a second SciDAC project connected to Brookhaven. TOPS software is being applied in CSC through a collaboration with one of its principal investigators. TOPS deploys a toolkit of open source solvers for partial differential equations, large systems of stiff ordinary differential equations, and linear and nonlinear algebraic systems, including eigenvalue problems, that arise in application areas such as accelerator design, biology, chemistry, magnetohydrodynamics, and particle physics. Scalable solution algorithms -- primarily multilevel methods -- aim to reduce computational bottlenecks by one or more orders of magnitude on terascale computers, enabling scientific simulation on a scale heretofore impossible. Along with usability, robustness, and algorithmic efficiency, an important goal is to attain high computational performance by accommodating to distributed hierarchical memory architectures.

The convergence rates of solvers traditionally employed in PDE-based codes degrade as the size of the system increases. This creates a double jeopardy for applications -- as the cost per iteration grows, so does the number of iterations. Fortunately, the physical structure of PDE problems, such as Poisson's equation for electrostatic potential, provides a natural way to generate a hierarchy of approximate models, through which the required solution may be obtained efficiently.

The efforts defined for TOPS and its collaborations incorporate existing and new optimal algorithms into scientific applications through code interoperability behind a standard interface. TOPS provides support for the software packages Hypre, PARPACK, PETSc (which has powered three Gordon Bell Prizes in recent years), ScaLAPACK, Sundials, SuperLU, and TAO. Some of these packages are in

the hands of thousands of users, who have created a valuable experience base on thousands of different computer systems.

The TOPS project webpage may be found at http://www.tops-scidac.org.

4.1.3 Advanced Computing for 21st Century Accelerator Science and Technology

R. Samulyak

The SciDAC Accelerator Modeling Project, "Advanced Computing for 21st Century Accelerator Science and Technology," was initiated in June 2001. Its primary goal is to establish a comprehensive terascale simulation environment for use by the U.S. particle accelerator community. Building upon a previous DOE Grand Challenge project as well as previous individual efforts at several national laboratories and universities, the SciDAC Accelerator Modeling Project represents the largest effort to date for the development of computer codes for accelerator design and analysis. The activities of the project are organized into three application-specific focus areas: Electromagnetics, Beam Dynamics, and Advanced Accelerators. Work in these areas is supported by collaboration with the SciDAC Integrated Software Infrastructure Centers (ISICs) and by personnel (including CSC/BNL) supported through the SciDAC Scientific Application Partnership Program (SAPP).

Research at CSC/BNL is conducted in close collaboration with the Beam Dynamics group. The primary goal of the study is the development of novel mathematical models and software modules for the computation of wake fields and their interaction with particle beams in high intensity accelerators. We have implemented our software in the MaryLie/Integrated Map and Particle Accelerator Tracking code, a parallel code that combines the magnetic optics capabilities based on the Lie algebraic technique with the 3D space charge capabilities. Currently we have been working on the implementation of the wakefield module in the Synergia framework. We have also joined the LBL/LANL effort to develop methods and high performance software tools for modeling the electron cooling and intrabeam scattering. These processes present the most important obstacles for the future upgrade of high-energy accelerators such as RHIC at BNL and Tevatron at FNAL. To provide tools for numerical studies of the electron cooling and intrabeam scattering, a code for solving the Fokker-Plank equation will be developed based on the self-consistent Langevin approach, with the diffusion and damping coefficients computed from first principles based on the Rosenbluth potentials. The problem presents significant computational challenges even for modern supercomputers.

The SciDAC website is http://www.osti.gov/scidac.

4.2 BIOLOGICAL AND ENVIRONMENTAL RESEARCH

Novel algorithms and hardware allow fine-scaled parallelism for all atom simulations to model structural biology. We project long simulated times (1-10 μsec for all atom simulations). This result will allow exploration of new time scales in structural biology. We are emphasizing conformational changes in proteins with known structures, in collaboration with BNL biologists.

Brookhaven has strong experimental programs in atmospheric chemistry and brain physiology. Both programs are data intensive. The challenge is not only data quantity, but especially data understanding. We have developed new tools for interactive visual and statistical data analysis. We are actively using these tools in diverse applications. See also Section 4.3.8.

4.2.1 MD Simulations of the Activation of the Adenovirus Proteinase

J.V. Stern, L. Slatest, J.W. Davenport, W.J. McGrath, and W.F. Mangel

The adenovirus proteinase (AVP) is a relatively inactive enzyme that can be activated by an 11-amino-acid peptide cofactor, pVIc [1,2]. The binding of pVIc increases the activity of AVP more than 3500-fold. The crystal structure of the AVP-pVIc complex has been determined experimentally using x-ray diffraction and the pVIc peptide was found to bind to AVP in an extended conformation quite far from the active site, as shown in Figure 1 [3]. The peptide traverses the surface of the protein, ranging from 14 Å to more than 30 Å away from the active site. This raises the question: how can the peptide at such a large distance from the active site have such a profound influence in the functioning of the enzyme?

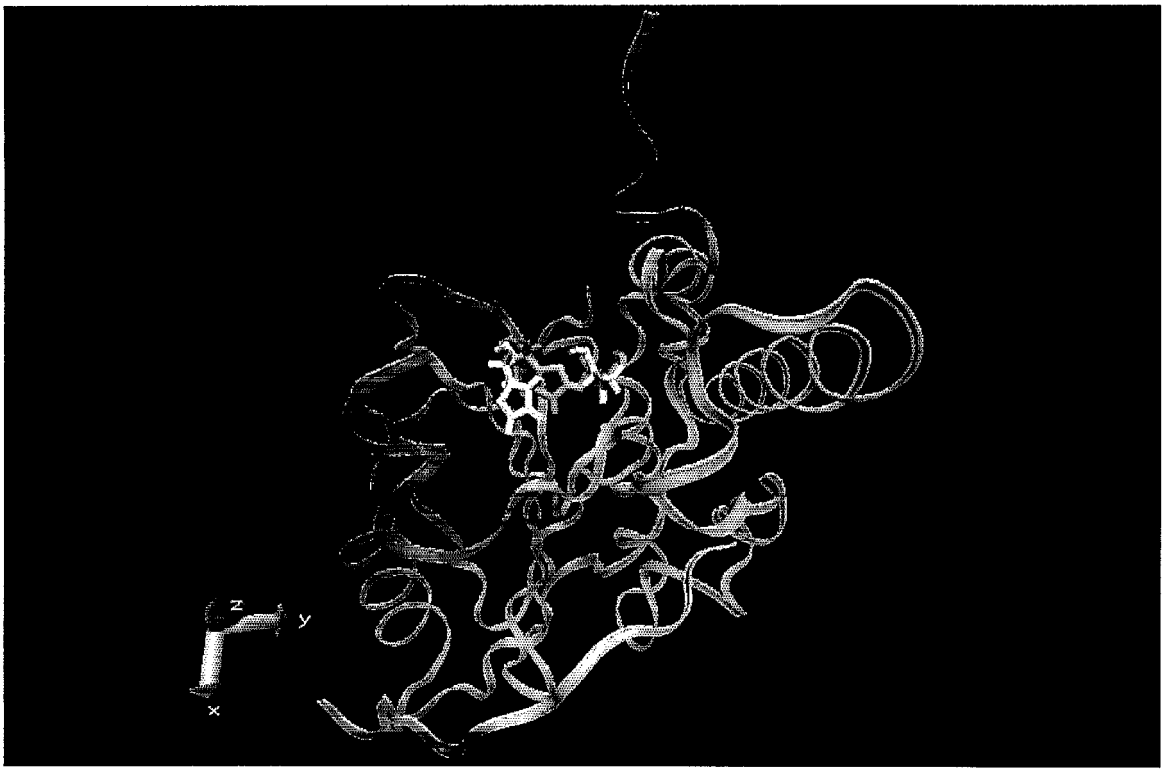


Figure 1. Crystal structure of AVP-pVIc. The pVIc peptide is overlapped in red in its approximate position. Two of the four catalytic active site residues Histidine 52 and Cysteine 120 are shown for the active structure (purple) and the result of the simulation following removal of the peptide (yellow).

Recently, the crystal structure of AVP in the absence of the cofactor was solved as well. Comparison of this structure to the structure of the AVP-pVIc complex is revealing how pVIc activates the enzyme. We hypothesize that in order for the binding of pVIc to activate the enzyme, a signal must be transduced by a sequence of structural changes beginning at the pVIc binding site on AVP and ending in the active site of AVP. Using molecular dynamics simulations, it is hoped that the sequence of structural changes that occur can be understood and exploited to provide targets for drugs. If the binding of a drug to a target within the sequence of structural changes can prevent the signal from being transduced, therefore preventing activation, the drug could be an effective anti-viral agent.

We have performed several sets of molecular dynamics simulations on this protein. In the first, using a generalized Born implicit water model, the peptide was removed from the known structure of the AVP-pVIc complex and the structure was equilibrated at 310K using a series of molecular dynamics and energy minimization steps. Following that, a 1 nanosecond simulation led to a new structure in

which the geometry of the 4 residue active site was significantly expanded. In fact the distance from HIS 52 to CYS 120 increased by 1.4 Å. We hypothesize that this increase is the immediate cause of the deactivation of the protease.

In the second set, an explicit water model was used [4]. Simulations were started from the experimental AVP structure (without the pVIc) to which a pVIc peptide had been docked. These simulations were run for ~ 3.7 nsec. They illustrate further changes in the backbone structure including a most interesting helix-coil transition. Using this approach, one possible step in the activation mechanism has been observed. A third set of simulations was run starting with the known AVP-pVIc complex but with the disulfide bond that binds pVIc to AVP removed. The objective was to test a model describing how pVIc binds to AVP by inducing a pocket in AVP, another possible step in the activation mechanism. A simulation proceeding in the reverse direction, unbinding, should therefore show some loss of the induced pocket. After ~ 100 picoseconds the last three amino acids of PVIc were indeed seen to leave the induced pocket.

The explicit water model molecular dynamics (MD) simulations were performed using NAMD [5], a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems. AMBER force fields [6] were used throughout. Our protein contains 204 amino acids plus the 11 amino acid peptide cofactor. After protonation this constitutes about 3500 atoms. Adding water yields a total of 30,000 atoms. The implicit water model MD simulations used a parameterization of the "OBC" model of generalized Born [7], the AMBER force fields, and the AMBER parallel molecular dynamics code [8]. Langevin dynamics with a collision frequency of one inverse picosecond was employed to maintain a constant temperature of 310 K throughout the MD run. The fastest motion in the system, i.e. the bond stretching freedom, was removed for bonds involving Hydrogen using the SHAKE algorithm, thereby allowing a larger timestep (1 femtosecond) to be used.

We plan to continue these simulations to elucidate the transformations involved in activating the enzyme and to discover sites along the pathway that may serve as targets for drugs that will act as antiviral agents.

References

[1] Mangel, W.F., McGrath, W.J. et al. Viral DNA and a viral peptide can act as cofactors of adenovirus virion proteinase activity. Nature **361**: 274-275 (1993).

[2] Mangel, W.F., Baniecki, M.L., and McGrath, W. J. Specific interactions of the adenovirus proteinase with the viral DNA, an 11-amino-acid viral peptide, and the cellular protein actin, Cell. Mol. Life Sci. **60:** 2347-2355 (2003).

[3] Ding, J., McGrath, W.J., et al.. Crystal structure of the human adenovirus proteinase with its 11 amino-acid cofactor. EMBO J. **15**: 1778-1783 (1996).

[4] Jorgensen, W.L., Chandrasekhar, J. et al. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. **79**: 926-935 (1983).

[5] Laxmikant, K.R., et al. NAMD2: Greater scalability for parallel molecular dynamics. J. Comp. Physics **151**: 283-312 (1999).

[6] Cornell, W.D., Cieplak, P. et. al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J. Am. Chem. Soc. **117**: 5179-5197 (1995).

[7] Onufriev, A., Bashford, D., Case, D.A. Exploring protein native states and large-scale conformational changes with a modified generalized Born model. Proteins (2004).

[8] Pearlman, D.A., Case, D.A. et. al. AMBER, a computer program for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to elucidate the structures and energies of molecules. Comp. Physics Commun. **91:** 1-41 (1995).

4.2.2 Structural Analysis of Botulinum Neurotoxin

Y. Deng, X. Chen, Y. Chen, B. Fang, G. Han, and P. Rissland

The toxin produced by the bacterium Clostridium botulinum is one of the deadliest known [1]. Recently a great deal of information about its mode of action has become available through structural and other biophysical studies. The toxin itself (BoNT) is a protein of approximately 1300 residues. It is made up of two chains, a light chain (LC) of approximately 50 kDa and a heavy chain (HC) of approximately 100 kDa. The two are linked by a disulfide bond. The light chain contains a Zn^{2+} ion which acts within the cell to cleave a protein necessary for neurotransmission leading to paralysis and death. This process is believed to involve several steps: binding of the toxin to the endosomal membrane, translocation through the lipid bilayer, and proteolysis of specific neurotransmitters within the cell. The structural studies show that the heavy chain contains two domains responsible for binding and translocation while the light chain contains the catalytic domain. Both the structure of the toxin outside the cell (but not at endosomal pH) and the mechanism of action of the zinc protease are reasonably well understood. What is not known are the structures at low pH (5.1-5.4) and the mechanism of translocation.

We have begun to study these two aspects in detail using highly parallel molecular dynamics and energy minimization codes. The force field used is given by [2]

$$E = \sum_{bonds} \frac{a_i}{2}(l_i - l_{i0})^2 + \sum_{angles} \frac{b_i}{2}(\theta_i - \theta_{i0})^2 + \sum_{torsions} \frac{V_n}{2}(1+\cos(n\omega - \gamma))$$
$$+\frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{q_i q_j}{r_{ij}}$$

To study the pH dependence we have performed molecular dynamics simulations on the catalytic domain (light chain) portion alone. We followed the traditional method of varying the protonation state of polarizable residues - those shown in red in Figure 2 (see for example [3]).

The results of 190 nanoseconds simulated time for two different temperatures and pH values are shown in Figure 3. We are currently analyzing these results to understand the physiologic importance of these changes.

Figure 2. Botulinum neurotoxin. Residues shown in red are changed as a function of pH.

We have developed a parallel molecular dynamics code MDoC (Molecular Dynamics on a Chip), a package designed by our group for ultrascalable architectures, such as IBM's BG/L and QCDOC. It uses the PDB and Amber [2] parameter files that are commonly employed in popular MD simulation packages. These files are generated to describe the molecular systems to be simulated, such as atom positions, force constants, charges, mass, etc. This enables the direct use of pre-existing files in MDoC without additional effort. MDoC uses Ewald summation to compute the Coulomb forces with complexity $O(N^{3/2})$ for N particles. Our research shows pure Ewald to perform better than the more conventional P3ME algorithm for large processor numbers. Our MDoC for ultrascalable architectures is designed to exploit multi-dimensional torus network topologies by tailoring MPI functions to achieve the optimal network load balance and maximal wire utilization during the electrostatic force calculation for which positions and charges of all particles must be passed to all processors. This was not incorporated in any earlier MD package. The code also integrates new algorithms for the FFT and for multiple time stepping.

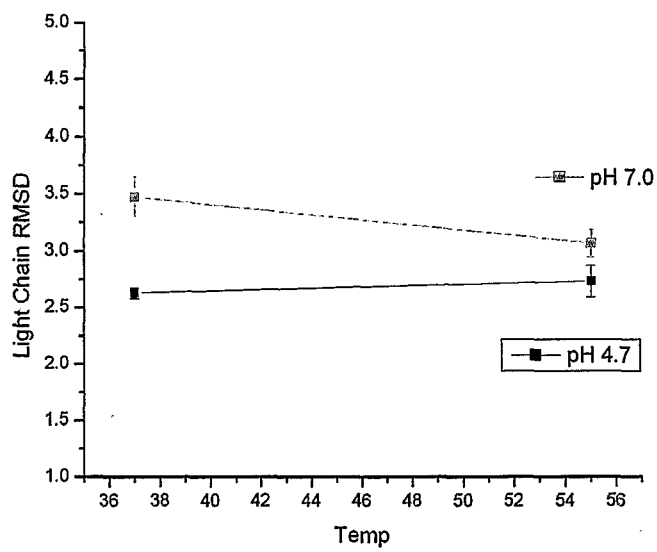


Figure 3: RMSD of light chain versus temperature and pH.

It is well known that the pH dependence of these structures is not fully described by the traditional method [3]. The protonation state is a function of pH, but the change with pH depends on the protonation state, requiring a self-consistent solution. Recently a new approach combining MD and Monte Carlo simulation has been developed [4]. We plan to implement this method into our own code and to use it to validate our current results.

References

[1] Turton, K., Chaddock, J.A., and Acharya, K.R. Botulinum and tetanus neurotoxins: structure, function and therapeutic utility. Trends in Biochemical Sciences **27:** 552-558 (2002).

[2] AMBER: http://www.amber.scripps.edu

[3] Smith, L.J., Dobson, C.M., and van Gunsteren, W.F. Molecular dynamics simulations of human α-Lactalbumin: Changes to the structural and dynamical properties of the protein at low pH. Proteins **36:** 77 (1999).

[4] Mongan, J., Case, D.A., and McCammon, J.A. Constant pH molecular dynamics in generalized Born implicit solvent. J. Comp. Chem. **25:** 2038 (2004).

4.2.3 Visualization and Data Mining for Detection of Cancer

W. Zhu, J. Kovach, and M. McGuigan

At an early stage, ovarian cancer has a high (90%) cure rate, but cannot be readily detected. At late stages, when detection usually occurs, the cure rate is 35%. Annually, 14,500 women die from this disease in the U.S. Mass spectrum analysis of protein fragments in blood samples has been proposed by Petricoin et al. (reported in Lancet [1]) as a diagnostic tool, specifically for diagnosis of cancer. The idea has two steps. The hardware, or physical measurement step, is by a mass spectrometer. This machine looks at blood fractions and trace proteins, and records mass (in relation to charge). Of course there are many hundreds of thousands of these signals ("markers"). The second step is to find the needle in the haystack. Which markers distinguish a normal patient from one with ovarian cancer? This is a problem in statistics and its solution enables the diagnostic procedure. Professor Wei Zhu, CSC staff member and member of the Department of Applied Mathematics and Statistics, John Kovach, MD, Director of the Long Island Cancer Center and a team of graduate students have made significant process with the pattern recognition problem [2].

The statistical problem is solved using diagnostic data, that is, blood samples from patients with known diagnosis, having ovarian cancer or normal. The data is divided in half, one half to train the statistics and the other half to test the conclusions. All markers are assessed for their ability to distinguish between normal and afflicted patients in the training set. About 150 markers show a statistically unusual pattern of difference between these populations. This number depends on the resolution power of the mass spectrum equipment. A subject with a positive reading on selected markers is predicted as possibly afflicted with ovarian cancer. When the test is applied to the testing set, accuracy in the range of 80-95% results. More recently, we have applied the same technology for the detection of head and neck cancer. Again, accuracy in the range of 80-95% results [3].

We are developing an open source toolkit for NIH/FDA/NCI to assist with the processing and analysis of datasets derived from proteomic experiments [4,5,6]. Open source software provides a mechanism for leveraging existing toolkits, sharing expertise and accelerating development. Our toolkit utilizes components from several existing open source projects. These include the visualization toolkit (VTK) and the machine learning in C++ library (MLC++). It will initially be

customized for serum proteomic pattern diagnostics involving the monitoring for ovarian cancer recurrence.

Serum proteomic tests involve a new form of biomarker, specifically, one that is multiplexed and pattern based. Following processing by mass spectroscopy, experimentalists have used serum proteomic information content to detect the onset of disease at a very early stage, thus maximizing the likelihood of successful outcomes.

Novel approaches to these issues are an active area of research. New pattern-based diagnostic methods and innovative software tools are needed for the management and analysis of this challenging data.

A concrete example, composed of an imaging study integrated with a serum clinical proteomic test, is now described. Figure 4 shows a high resolution Doppler ultrasound image of a woman who presented to her physician complaining of progressive pelvic pain. The image shows a complex right mass centered on the right ovary. The Doppler signal indicates no significant blood flow within the mass. Given the lack of specificity of the image finding, significant benefit could be gained by obtaining a proteomic pattern analysis of the patient's blood. The analysis would compare the patient's unique proteomic profile with well classified reference standards in order to obtain a score indicating the likelihood of cancer. Subsequent figures will illustrate dissection of large and complex proteomic spectra via visual data mining techniques.

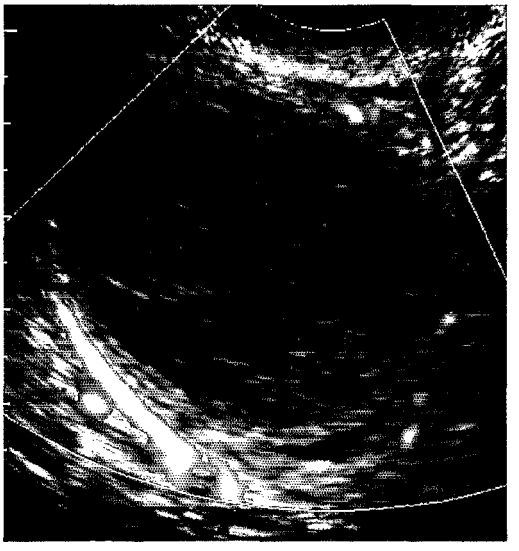


Figure 4. Color flow Doppler ulstraound of an ovarian mass.

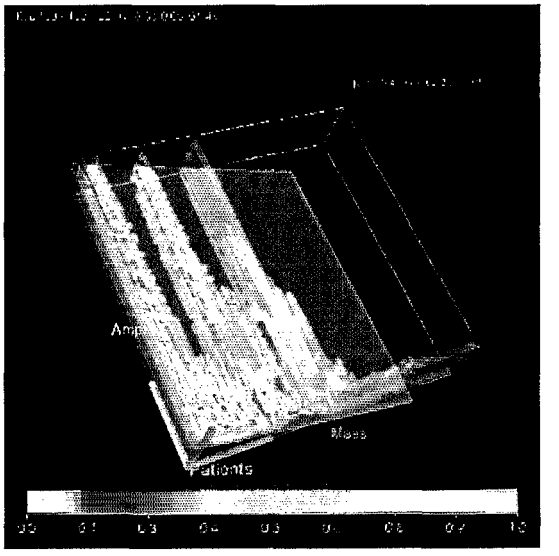


Figure 5. Splat visualization displaying Serum proteomic patterns from 22 patients with ovarian cancer.

Figure 5 shows a Splat Visualization rendering of serum proteomic data from 22 patients following processing by a high-resolution mass spectrometer. The majority of the spectra are from patients with known ovarian cancer. The Splat Visualizer tool is an aggregate 3-D plotter that allows rapid manipulation and viewing of large datasets. The tilted x-axis (labeled *Mass*) represents the mass values of the peptides and protein fragments. The y-axis (labeled *Amp*) shows normalized amplitude values (relative measure of abundance). The z-axis (labeled *Patients*) has been used to stack individual patient spectra, thus allowing comparison. The graphic has also been automatically colored as a function of support (number of data points behind the voxel) white and blue having the

least amount of support, and red the most. Therefore, the patient's unique proteomic signature can be compared and contrasted with known controls and calibrators.

Figure 6 displays output of the Multi-view visualization tool. This tool can be thought of as a superset of the Splat Visualizer, since it allows for either synchronous or asynchronous coordination of multiple windows, all containing different proteomic datasets. In all four views the x and y-axes are defined as in Figure 5. However, each view employs a different clinical parameter in the z-axis, allowing simultaneous proteomic pattern viewing, but in that window's particular context.

In each window, serum proteomic spectra from approximately 100 patients with known ovarian cancer have been aggregated and grouped according to a clinical parameter of interest. Beginning top left and moving clockwise: spectra are rendered as a function of the patient's age, next spectra are grouped and displayed according tumor grade, next by a lab value (Cancer Antigen 125 (CA 125)), and finally by stage of disease. Any window can be maximized and subsequent operations performed, such as: zoom, pan, rotate, drill, or execution of a defined command.

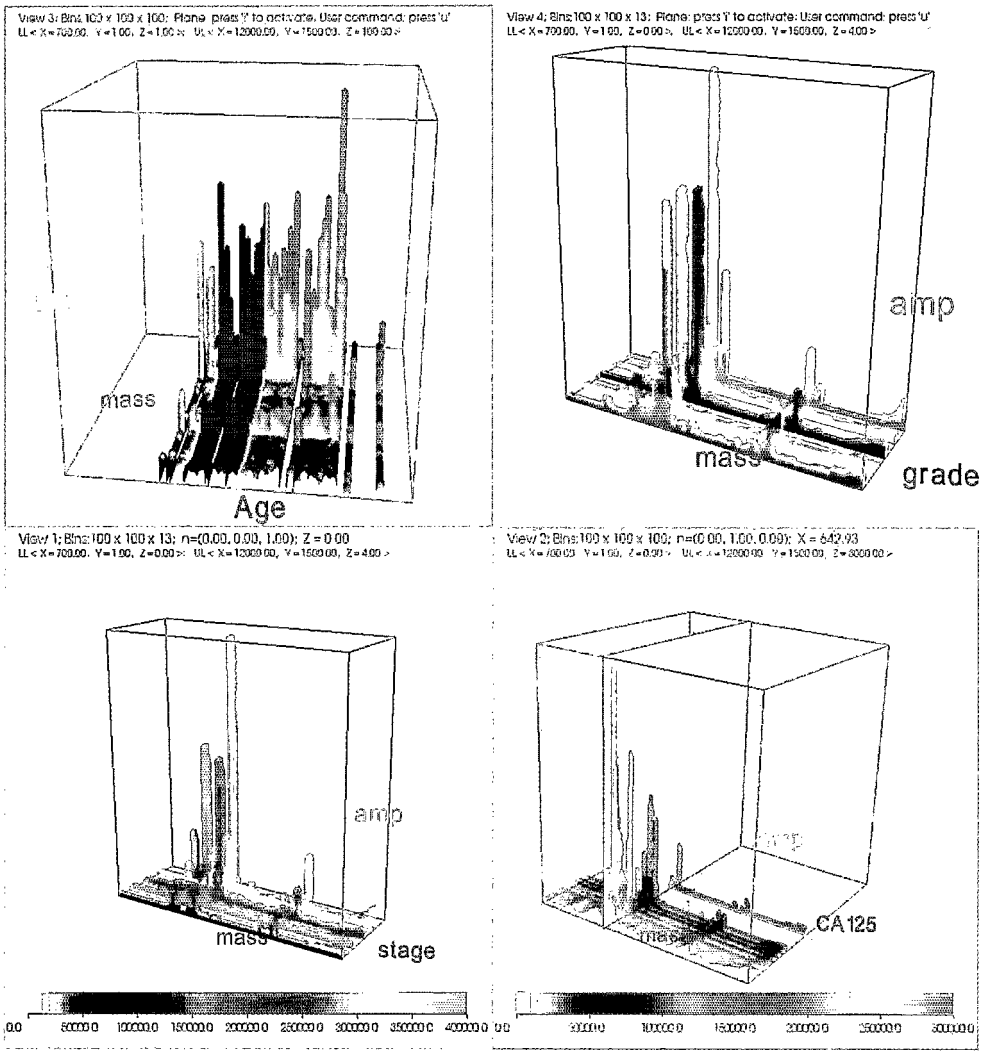


Figure 6. Multi-view visualization of aggregate proteomic patterns (approximately 100 patients, most with ovarian cancer) viewed as a function of a clinical parameter.

The information archive present in the low molecular weight serum proteome is complex. We have found visual data mining techniques coupled with traditional and novel pattern recognition techniques to be beneficial in the analysis of complex proteomic datasets. Open source software toolkits can be an essential component for advancements in the early detection of disease and its successful medical management.

References

[1] Petricoin III, E.F., Ardekani, A.M., Hitt, B.A., Levine, P.J., Rusaro, V.A., Steinberg, S.M, Mills, G.B., Simone, C., Fishman, D.A., Kohn, E.C., and Liotta, L.A. Lancet **359:** 572-577 (2000) and Lancet **360:** 169-171 (2002).

[2] Zhu, W., Wang, X., Ma, Y., Rao, M., Glimm, J., and Kovach, J. Detection of cancer specific markers amidst massive mass spectral data. Proc. Nat. Acad. Sci. **100:** 14666-14671 (2003).

[3] Wang, X., Zhu, W., Pradhan, P., Ji, C., Ma, Y., Semmes, J., Glimm, J., and Mitchell, J. Feature extraction in the analysis of proteomic mass spectra. Proteomics, to appear.

[4] Johann, D.J., McGuigan, M.D., Tomov, S., Fusaro, V.A., Ross, S., Conrads, T.P., Veenstra, T.D., Fishman, D.A., Whiteley, G.R., Petricoin, E.F., and Liotta, L.A. Novel approaches to visualization and data mining reveal diagnostic information in the low amplitude region of serum mass spectra from ovarian cancer patients. Disease Markers **19:** 197-207 (2004).

[5] Johann, D.J., McGuigan, M.D., Tomov, S., Blum, E., Whiteley, G.R., Petricoin, E.F., and Liotta, L.A. Toward a Systems Biology Software Toolkit, 17th IEEE Symposium on Computer-Based Medical Systems, June 2004.

[6] Johann, D.J., McGuigan, M.D., Patel, A.R., Tomov, S., Ross, S., Conrads, T.P., Veenstra, T.D., Fishman, D.A., Whiteley, G.R., Petricoin, E.F., and Liotta, L.A. Clinical proteomics and biomarker discovery. Annals of the New York Academy of Sciences **1022:** 295-306 (June 2004).

4.2.4 Structure-based Drug Design Targeting HIV

B. Strockbine, S. Mukherjee, N. Carrascal, and R.C. Rizzo

A promising new avenue in the fight against HIV/AIDS is to target proteins involved in viral-host membrane fusion which lead to infection [1]. Currently, the only marketed inhibitor for membrane fusion is a large 36 amino acid peptide called Fuzeon that targets the viral protein called gp41 [2]. The goal of our research is to develop and use computational methods to discover low molecular weight compounds that target HIVgp41 and block membrane fusion with host cells. As a class, small molecules may have fewer side effects, be easier to take, and cheaper to produce than Fuzeon (ca. $20,000 per patient/per year). Affordable drugs are desperately needed; over 95% of all new HIV infections are in developing countries. And, as HIV continues to mutate, drug resistance develops and the discovery of new drug classes with novel modes of inhibition is increasingly important.

Experimental studies of the core domain of gp41 have revealed a hairpin structure (fusion-active) in which three outer C-helices (Figure 7, green) loop and wrap around three inner N-helices (Figure 7, red) [1]. During infection, gp41 undergoes a transient intermediate state (pre-hairpin) in which the inner trimer core (N-helices) is exposed prior to the membrane fusion. Binding of peptides to the inner core during this intermediate stage prevents the reformation of the trimer of hairpins, thereby inhibiting virus-host cell membrane fusion [1,3].

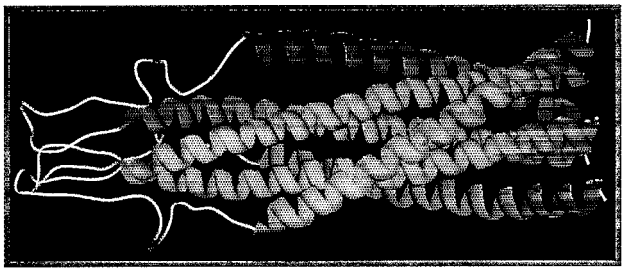


Figure 7. HIVgp41 coiled-coil involved in viral-host cell-membrane fusion. Outer C-helices in green, inner N-helices in red. Structure coordinates from pdb entry 1IF3.

To discover small molecule lead compounds capable of disrupting fusion events in HIV we are pursuing computational high-throughput screening (docking) of large ligand libraries to a specific hydrophobic pocket on the inner trimer core of gp41. Docking will be used to predict how compounds interact with gp41 by considering thousands of different conformations, orientations, and the electrostatic and steric complementarity of each protein-ligand complex. Our previous work employed docking to provide detailed atomic-level structures for subsequent drug design efforts targeting HIV reverse transcriptase [4]. As co-developers of the DOCK [5] program, we have recently optimized procedures [6] for computing important desolvation terms associated with molecular recognition. Specifically, we have added new algorithms to DOCK, based on Molecular Mechanics Generalized Born Surface Area (MM-GBSA) methods, to more accurately estimate free energies of binding [7,8]. We have recently used MM-GBSA methods to simulate physiologically relevant domain motions in HIV protease which involve large-range flap movement required for substrate access [9].

Here, for each ligand successfully docked to the gp41 hydrophobic pocket, the lowest energy binding modes are saved, ranked for affinity with the receptor, and then clustered into groups according to heavy atom root-mean-square-deviation (rmsd) similarity. Top-scoring cluster heads are then post-processed with MM-GBSA scoring. The best re-scored ligands are visually inspected for complementarity with the gp41 hydrophobic pocket using stereo 3D graphics in the context of the hydrophobic pocket. Additional computational filtering criteria are applied to maximize the diversity and drug likeness of any compounds prior to purchase.

Experimental testing for compounds suggested to bind tightly to gp41 are being done in collaboration with Dr. Miriam Gochin at the University of California at San Francisco (San Francisco, California). Compounds suggested by computational docking and determined experimentally to bind will be analyzed with *in vitro* syncytium and HIV viral infectivity assays. The long-term goal of Dr. Rizzo's research is to develop promising non-peptide drug candidates for HIV-1 fusion inhibition. We gratefully acknowledge support from the Department of Applied Mathematics at Stony Brook, the NYSTAR James D. Watson Investigator Program, and the Computational Science Center at Brookhaven National Laboratory.

References

[1] Eckert, D.M. and Kim, P.S. Mechanisms of viral membrane fusion and its inhibition. Annu. Rev. Biochem. **70:** 777-810 (2001).

[2] FDA Approved Anti-HIV Drugs. http://www.niaid.nih.gov/daids/dtpdb/fdadrug.htm.

[3] Jiang, S., Lin, K., Strick, N., and Neurath, A.R. HIV-1 inhibition by a peptide. Nature **365:** 113 (1993).

[4] Rizzo, R.C.; Wang, D., Tirado-Rives, J., and Jorgensen, W.L. Validation of a model for the complex of HIV-1 reverse transcriptase with sustiva through computation of resistance profiles. J. Am. Chem. Soc. **122:** 12898-12900 (2000).

[5] *DOCK*; University of California at San Francisco: San Francisco, CA.

[6] Rizzo, R.C., Aynechi, T., Case, D.A., and Kuntz, I.D. Estimation of absolute free energies of hydration using continuum methods: Accuracy of partial charge models and optimization of nonpolar contributions. J. Chem. Theory Comp., in press, 2005.

[7] Srinivasan, J., Cheatham, T.E., Cieplak, P., Kollman, P.A., and Case, D.A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate - DNA helices. J. Am. Chem. Soc. **120:** 9401-9409 (1998).

[8] Massova, I. and Kollman, P.A. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. Perspect. Drug Discov. Design **18:** 113-135 (2000).

[9] Hornak, V, Okur, A., Rizzo, R.C., and Simmerling, C. Flap structure and dynamics in HIV-1 protease simulations. Proc. Natl. Acad. Sci. USA, submitted, 2005.

4.2.5 Cardiac Electrophysiology

S. Evans, F. Fenton, J. Glimm, H. Hastings, and W. Oh

Sudden cardiac failure is a major cause of death. The causes and conditions leading to sudden cardiac death are complex and varied, and have been the subject of extensive studies. In collaboration with researchers from Hofstra University and Beth Israel Medical Center, we have been studying the dynamics of ventricular fibrillation, an irregular and chaotic heart rhythm that almost always results in sudden cardiac death.

We model cardiac electrophysiology through computer simulation. For this purpose it is common to use the monodomain equation that limits the electrical activity within cells. In reality, the ion-mediated current travels into and out of cells, requiring a bidomain model to more accurately simulate complex cardiac electrophysiology, using the equations

$$\begin{aligned} \frac{\partial_u}{\partial_t} &= \frac{1}{\chi C_m}\nabla\cdot\sigma_e\nabla u_e + \frac{f(u)}{R_m C_m}, \\ \nabla\cdot\sigma_i\nabla u + \nabla\cdot\left(\sigma_i+\sigma_e\right)\nabla u_e &= 0. \end{aligned} \tag{1}$$

Here u_i and u_e are the intra and extra cellular potentials, $u = u_i - u_e$ is the transmembrane potential, σ_i and σ_e are conductivity tensors, and C_m, R_m, and χ are the membrane capacitance, membrane resistance and surface-to-volume ratio. The reaction term f(u) represents the ionic current.

We developed a parallel code to solve the bidomain equations in 2d and 3d. With this code, we will be able to study the onset and the dynamics of ventricular fibrillation.

In Figure 8, we present a 3D simulation using the bidomain code. The transmembrane potential is plotted at 100 msec intervals. An initial stimulus at the center of the domain has propagated radially. The first frame is at 100 msec. Since we used anisotropic conductance, the wave propagates with

slightly different speeds in the x, y and z directions. At t=215 msec (before the third frame), we stimulated the center again and observed scroll waves. One practical way to identify the scroll wave is to arbitrarily choose a single isopotential surface of constant potential u(x,t)=c. The wave tip, or vortex filament, can be defined as the intersection of u(x,t)=c and \partial_t u = 0. In Figure 9, we present the dynamics of the vortex filament starting at 300 msec at 100 msec interval.

The simulations will be validated by comparison with experimental data, and the code will be used in a large-scale study of wave stability and scroll wave dynamics in realistic 3D anatomy. Large-scale simulations will provide critical insights into basic cardiac physics and dynamics. Moreover, the use of these models will allow experiments that cannot be performed with animal models because variability among individuals within a species makes it difficult to isolate causal factors.

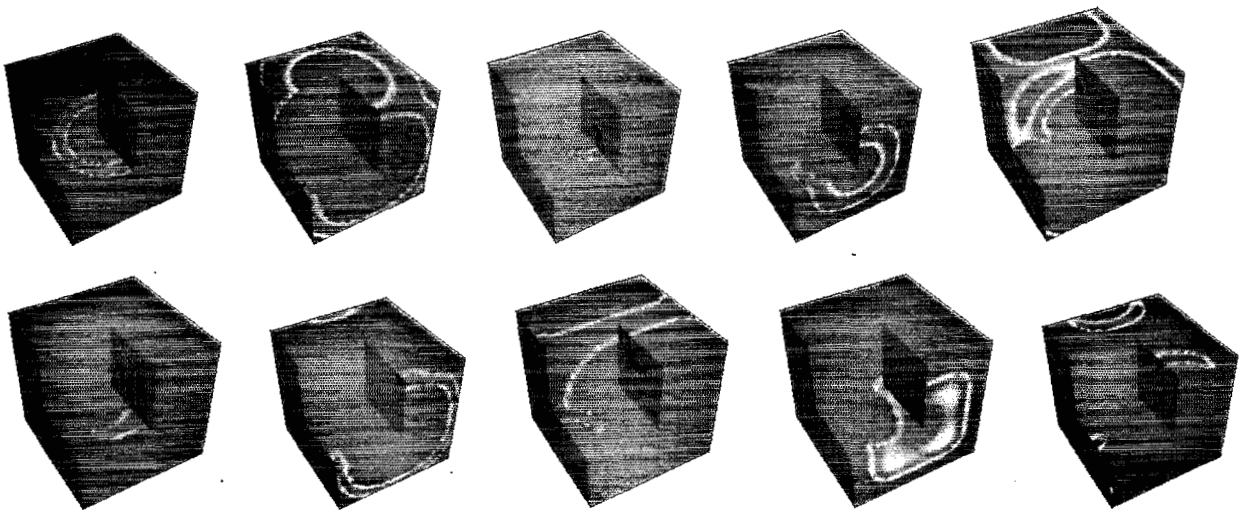


Figure 8. Successive time frames showing the propagation of the transmembrane potential and development of scroll waves, based on the bidomain equation.

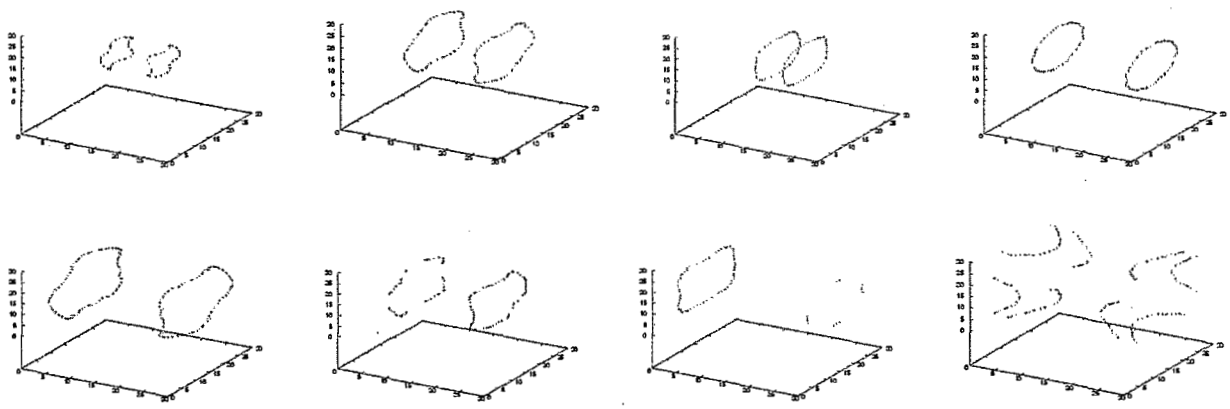


Figure 9. Successive time frames showing the vortex filament of the scroll waves shown in Figure 8.

References

[1] Fenton, F. and Karma, A. Vortex dynamics in 3D continuous myocardium with fiber rotation: fibrillation. CHAOS **8:** 20 (1998).

[2] Fenton, F., Evans, S., and Hastings, H.M. Memory in an excitable medium: A mechanism for spiral wave breakup in the low excitability limit. Phys. Rev. Lett. **83:** 3964-3967 (1999).

[3] Hastings, H.M., Fenton, F., Evans, S., Hotomaroglu, O., Geetha, J., Gittelson, K., Nilson, J., and Garfinkel, A. Alternans and the onset of ventricular fibrillation. Phys Rev E **62:** 4043-4048 (2000).

[4] Keener, J.P. and Bogar, K. A Numerical method for the solution of the bidomain equations in cardiac tissue. Chaos **8(1):** 234-241 (1998).

[5] Fenton, F.H., Cherry, E.M., Hastings, H.M, and Evans, S.J. Multiple mechanisms of spiral wave breakup in a model of cardiac electrical activity. CHAOS **12 (3):** 852-892 (2002).

4.2.6 Atmospheric Aerosols

W. Zhu, K. Mueller, and S. Schwartz

The purpose of this project is to develop data processing tools for the classification and time series analysis of atmospheric aerosol particles. Atmospheric aerosols play an important role in both climate and public health. Current research is focused on determining the chemical composition of aerosols, and thus the composition of air and its evolution over time. Massive mass spectra data, on the scale of terabytes per week, have been collected continuously over time by researchers in the Atmospheric Sciences Division at Brookhaven, using their cutting-edge, field deployable, high-precision mass spectrometer. The first task is to determine the chemical composition of each aerosol based on its mass spectrum. To achieve this, we classify aerosols by their chemical compositions and thus determine the air composition at each time point. Finally, we study the time series evolution of atmospheric compositions identifying normal as well as abnormal patterns in real time.

To convey the analysis results to the scientists in an intuitive way and to incorporate their expert knowledge into the data analysis phase, an interactive graphical user interface utilizing modern scientific visualization techniques becomes a necessity. Large-scale BNL data collections have reached sizes for which straightforward visualization techniques are starting to fail. In this spirit, our proposed approach couples a powerful Bayesian classification and multivariate time series analysis engine with an intuitive and responsive graphical user interface to fine-tune the underlying models. We call this application *SpectrumMiner*, a significant component of a more general visual data mining framework and application, called ViStA, which we are currently developing at BNL [1,2,3] (for more detail see Section 4.3.8).

The statistics engine. We first classify aerosols based on their mass spectra through an iteration of expert-machine interaction using a Bayesian classifier. The Bayesian classifier can incorporate not only explicit classification rules but also prior knowledge in terms of a partial training data set. We have also been implementing automatic procedures to elucidate the structure of the compounds. To study the evolution of aerosols and the changes in atmospheric composition versus time, both the univariate and the multivariate time series analyses are employed to unravel trends and patterns.

The traditional statistical classifier would take only explicit classification rules. Machine learning techniques such as the neural network would take only a training set. However, the semi-supervised clustering approach we are taking requires the classifier to learn from both the explicit rules and the

implicit rules embedded in a partial training set established by the experts as they survey the current clustering results and make adjustments to the clusters. For this we have implemented an interactive Bayesian classification framework that could absorb the up-to-date prior information in both the explicit and implicit formats, and produce the updated posterior classification results.

For efficient classification and automatic structure elucidation, we have been constructing a molecule library where the signature profile and class membership of each molecule is established. The molecules are classified along a natural chemical classification tree with two categories--organic and inorganic--at the initial node. Subsequently, the organics are further divided into classes of carboxylic acids, aldehydes, ketones, alkenes, alkanes, aromatics, etc. We begin with the spectra of known molecules (NIST library or lab-generated).

Visualization. The statistical analysis engine is combined with a highly visual interface to facilitate interactive exploration, mining, classification and survey of these large, high-dimensional data collections. In order to empower scientists to control and fine-tune the mining and classification process in an intuitive and interactive way, SpectrumMiner's hierarchical classification algorithm is user-steerable via a novel multi-modal visual interface. An important component of this interface is the *interactive dendrogram* [1], where hierarchy nodes are placed on concentric circles whose radii are determined by the dissimilarity of the node's sub-tree. We chose a circular layout of the dendrogram since it makes better use of space than its linear counterpart. It inherently dedicates less drawing space to the higher-level, less numerous nodes, and distributes more space to the many leaf nodes along the circumference of the circle. See Figure 10, where we show a screen capture of SpectrumMiner, with the interactive dendrogram located on the bottom right. Edges are colored using a rainbow colormap to indicate the number of data items they carry. To the left of the dendrogram is the node viewer. Selecting a particular node will display the average spectrum of all data items classified into the node (window with white background), as well as the node's data composition (window with blue background just below). We are currently incorporating the classification steering capabilities into our system. By inspecting the present classification in the node viewer and the dendrogram, scientists may decide that some particles, or the entire node, have been misclassified. To correct this error, they then simply drag the concerned particles or node into the proper location within the hierarchy, which subsequently triggers a refinement of the classification rules.

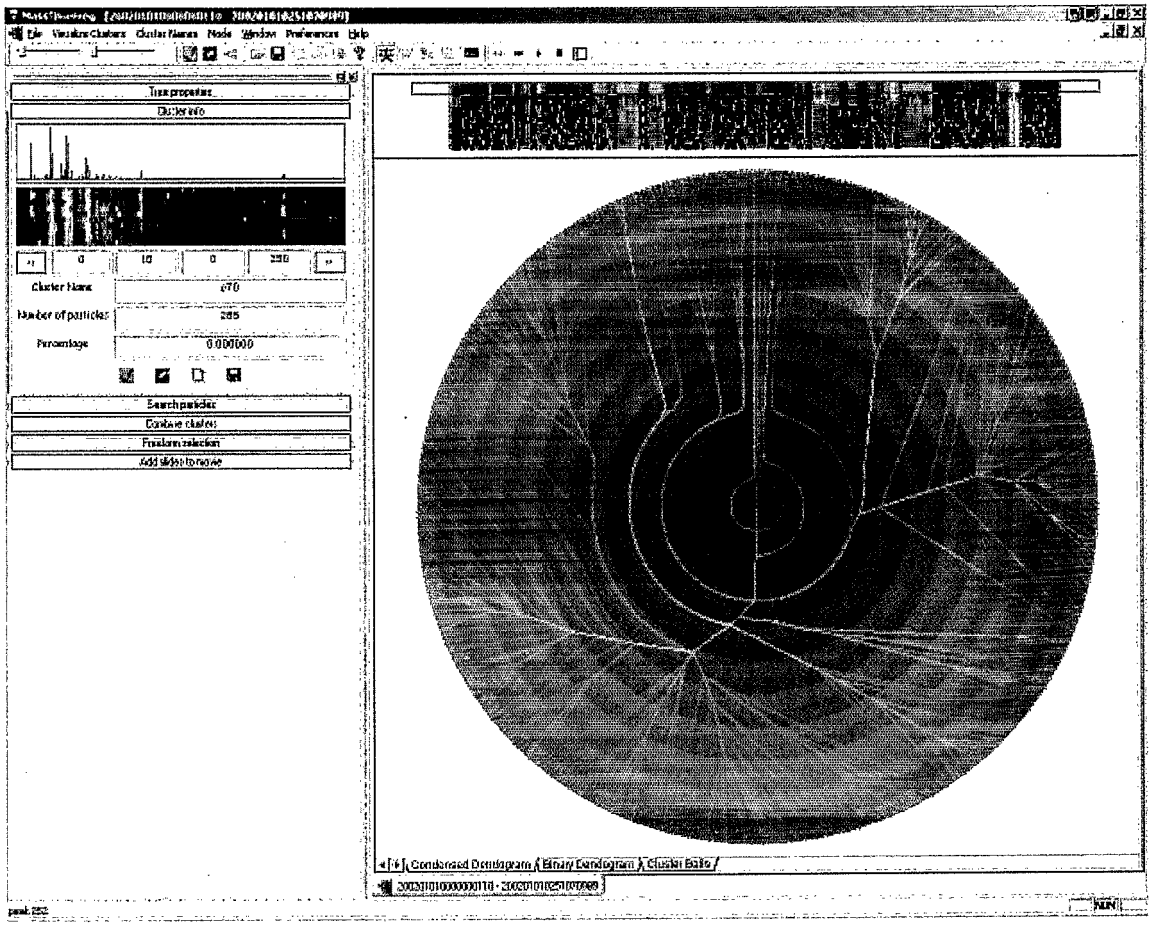


Figure 10. SpectrumMiner's interactive dendrogram interface.

References

[1] Imrich, P., Mugno, R., Mueller, K., Imre, D., Zelenyuk, A., and Zhu, W. Interactive poster: Visual data mining with the interactive dendrogram. IEEE Information Visualization Symposium'02; Boston, October 2002.

[2] Imrich, P., Mueller, K., Imre, D., Zelenyuk, A., and Zhu, W. A Hardware- Accelerated Rubbersheet Focus + Context Technique for Radial Dendrograms. IEEE Information Visualization Symposium'03; Seattle, October 2003.

[3] Imrich, P., Mueller, K., Imre, D., Zelenyuk, A., and Zhu, W. 3D ThemeRiver. IEEE Information Visualization Symposium'03; Seattle, October 2003.

4.3 ADVANCED SCIENTIFIC COMPUTING

4.3.1 TSTT: Atomization and Spray

Z. Xu, J. Glimm, M.N. Kim, X. Li, A. Marchese, R. Samulyak, and C. Tzanos

TSTT: Jet Breakup and Spray Formation in a Diesel Engine

In the past year, we have studied the mechanisms leading to jet breakup and spray formation for a high-speed diesel jet injected through a circular nozzle through the first principles simulations. The simulations are conducted using the Front Tracking method within a 2D axis-symmetric geometry. Our goal is to model the spray at a microphysical level, with the creation of individual droplets.

Through our study, we have found that the formation of cavitation vapor bubbles is the key phenomenon contributing to the breakup. During the inception stage of the cavitation, the cavitation is of traveling bubble type. It grows intensively and coalesces into big bubbles. These bubbles could create an intense turbulence and produce an atomizing spray.

We have developed a heterogeneous EOS model consisting of explicitly tracked cavitation vapor bubbles within the liquid diesel fuel for the mixed phase (diesel vapor/liquid) regime. Based on the homogeneous nucleation theory, we have formulated a dynamic bubble insertion algorithm to simulate the incipient creation of cavitation bubbles.

The interface which separates the liquid and the vapor is modeled as a phase boundary. A new phase transition model to calculate the solutions for the dynamic phase transitions has also been developed. In this model, viscosity and the surface tension on the phase boundary are neglected, because the thermal effect is normally dominant over the viscous effect in phase transitions. We use the kinetic theory of evaporation to give the evaporation rate with a coefficient determined experimentally. The net mass flux of evaporation then is $M_{ev}=\alpha(P_{sat}(T)-p_v)/(\sqrt{(2\pi RT)})$. On the phase boundary, the temperature is continuous, while the vapor pressure is allowed to have a deviation from the Clausius-Clapeyron equation. Thus the interface motion depends on the phase change under nonequilibrium thermodynamic conditions along with hydrodynamic conditions.

In the simulations, we used n-heptane as a replacement for No. 2 diesel fuel because the thermal data of No. 2 diesel fuel is not available, and N-heptane is the major component of No. 2 diesel fuel. We have compared our simulation results with the experimental data. These experimental data present mass vs. time in a 0.55 mm wide observation window which is centered 1 mm from the nozzle exit. Like the experiments, the simulations also predicted a peak mass. After the peak, the predicted mass exhibits a small variability (Figure 11). The opening angle of the jet (Figure 12), which varies as a function of time, is about 15 to 30 degrees and is in agreement with its experimental value. Figure 13 shows a comparison of jet tip velocity computed from simulations with experimental data. Although the value computed from simulation predictions exhibits a wide variability, on the average, it is in agreement with the value computed from the experiment. It should be noted that the experimental data has been averaged over 100 injection cycles to remove fluctuations.

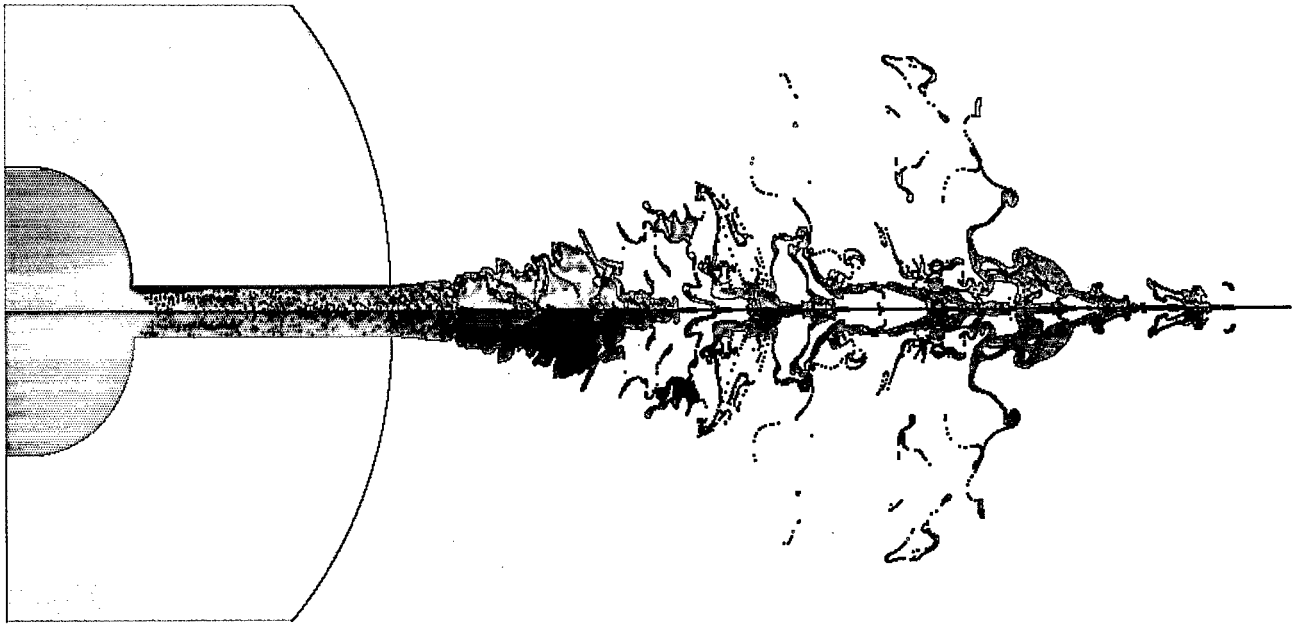


Figure 11. The vorticity (above) vs. density (below). The dark spots are the diesel vapor bubbles. The bubbles appear and disappear as rarefaction waves progress up and down the nozzle. The large blue region in the lower frame is a vapor region.

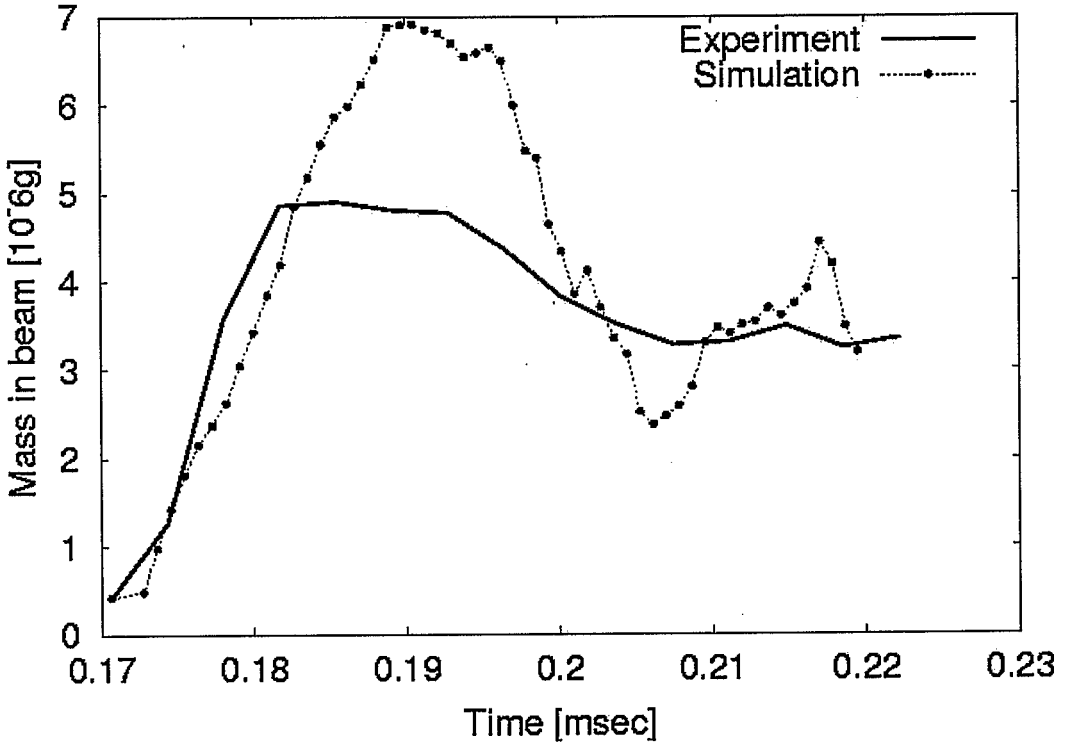


Figure 12. The plot of mass at 1mm from the nozzle exit.

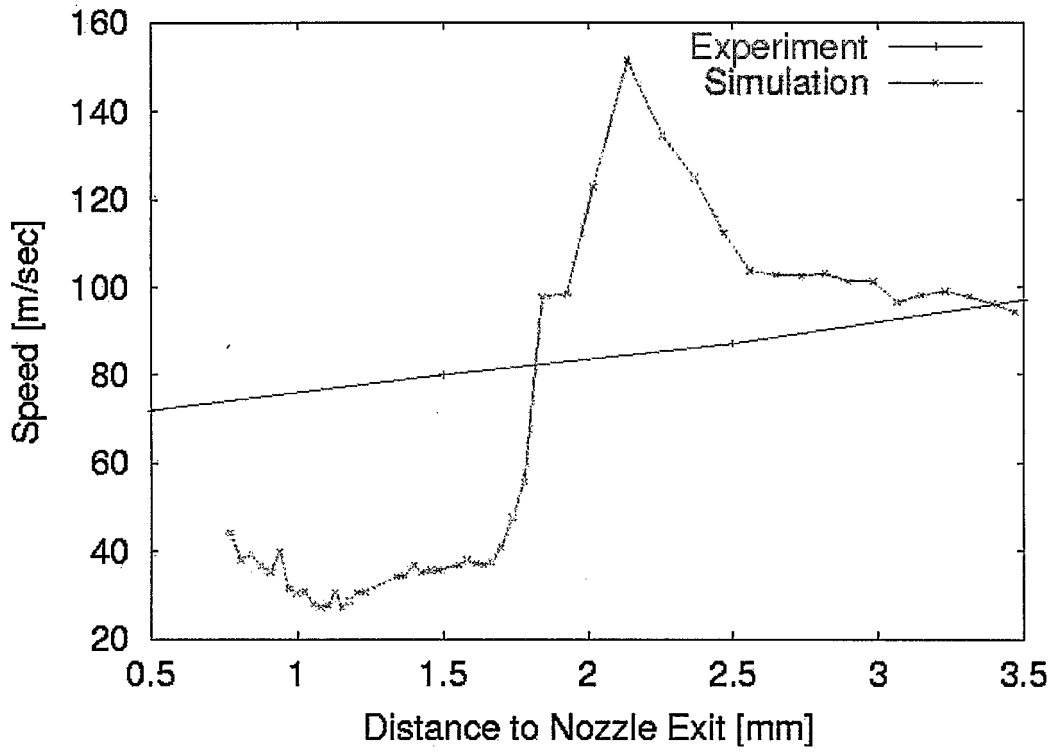


Figure 13. Comparison of jet tip velocity computed from simulations with experimental data.

The current work is focused mainly on the simulation of the two-phase mixture resulting from cavitation, and the influence of other parameters on spray formation is a subject of further research.

References

[1] MacPhee, A.G., Tate, M.W., and Powell, C.F. X-ray imaging of shock waves generated by high pressure fuel sprays. Science **295:** 1261-1263 (2002).

[2] Poola, R.B. et al. Development of a quantitative measurement of a diesel spray core by using synchrotron X-rays. Eighth International Conference on Liquid Atomization and Spray Systems, Pasadena, CA, July 16-20, 2000.

[3] Powel, C.F., Yue, Y., Poola, R., and Wang, J. Time resolved measurements of supersonic fuel sprays using synchrotron x-rays. J. Synchrotron Rad. **7:** 356-360 (2000).

[4] Powel, C.F., Yue, Y., Poola, R., Wang, J., Lai, M., and Schaller, J. Quantitative x-ray measurements of a diesel spray core. In *Proc. 14th Annual Conference on Liquid Atomization and Spray Systems (ILASS), Dearborn, MI*, 2001.

[5] Glimm, J., Kim, M.-N., Li, X.-L., Samulyak, R., and Xu, Z.-L. Jet simulation in a diesel engine. *Proc. Third MIT Conference on Computational Fluid and Solid Mechanics,* 2004, Computational Fluid and Solid Mechanics 2005, ISBN: 0-08-044476-8.

[6] Xu, Z., Kim., M.-N., Oh, W., Glimm, J., Samulyak, R., Li, X., Tzanos, C., and Lu, T. Atomization of a high speed jet. Physics of Fluids, submitted.

4.3.2 TSTT: Frontier-Lite

X. Li, Z. Xu, and B. Fix

A general purpose software package for geometry and interface dynamics has been extracted from the FronTier code and is now publicly available.

This package, called FronTier-Lite [1], is designed for users with little training in the algorithms used in the front tracking method and yet deciding to apply the high quality front tracking method to various scientific problems with dynamic front propagation. This code is downloadable from the web, and is accompanied by a web-based testing and evaluation site and extensive web-based documentation.

The software is organized in three levels. The first level deals with the static manifolds and geometry. This includes functions to initialize the interface, optimize an existing interface, and calculate geometry-dependent variables.

The dynamic front package provides a set of functions for the propagation of the interface in a given velocity field and with geometry-dependent velocity functions. The driver function advance_front() contains preset pointers to functions based on dimension and algorithms selected by the user. This function will also detect the geometrical and topological correctness of the interface after the propagation. When necessary, it will perform a sequence of operations to guarantee the soundness and quality of the interface before the next propagation step.

The front package allows users to provide their own velocity function for the propagation of the interface. To do so, the user needs to create a data structure containing all necessary parameters for the velocity function. This data structure is cast into a pointer after it is initialized. It will be passed to the velocity function, which is also a pre-assigned anonymous pointer. Inside the velocity function, the pointer of velocity parameters will be cast back to point to the original data structure, and the velocity function retrieves its parameters for the computation of velocity of each marker point.

This front tracking method has been compared with other interface methods on some benchmark problems. In Figure 14, we compare with the level set method. We used the fifth order WENO scheme for the convection of the level set function, while for the front tracking code, we used the fourth order Runge-Kutta method for the point propagation. After 13 revolutions, the fourth order Runge-Kutta method appears to be extremely accurate in the front tracking simulation, while the level set computation begins to show edge smoothing after the second rotation. At the end of the 13^{th} circulation, the slot is closed at the top, resulting in a topologically incorrect bifurcation.

Comparison with the volume of fluid method also showed high quality for the front tracking code. One of the benchmark tests is the three-dimensional deformation velocity field described by the velocity functions

$$u(x,y,z) = 2\sin^2(\pi x)\sin(2\pi y)\sin(2\pi z)\cos(\pi t/T) \quad (1)$$

$$v(x,y,z) = -\sin(2\pi x)\sin^2(\pi y)\sin(2\pi z)\cos(\pi t/T) \quad (2)$$

$$w(x,y,z) = -\sin(2\pi x)\sin(2\pi y)\sin^2(\pi z)\cos(\pi t/T). \quad (3)$$

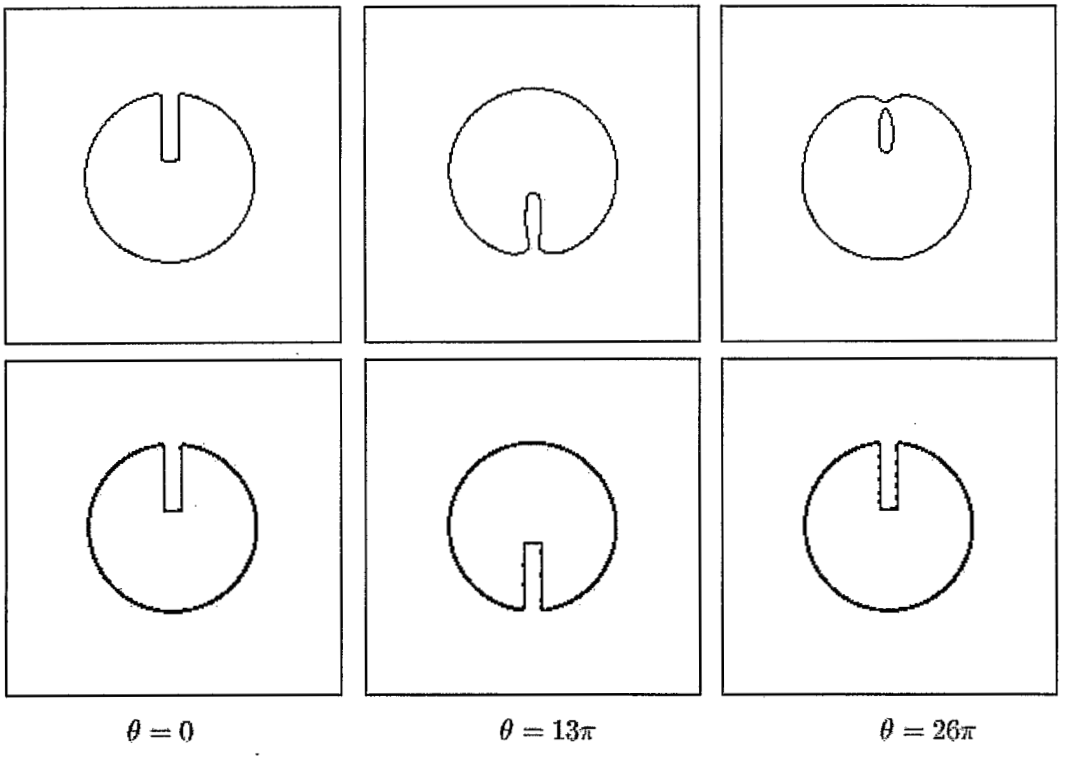


Figure 14. Comparison of slotted disk simulation using high order methods. The upper sequence shows the result of the level set method using the fifth order WENO scheme, and the lower sequence shows the result of front tracking using the fourth order Runge-Kutta method.

The interface evolves dynamically from an initial sphere of radius 0.15 centered at (0.35, 0.35, 0.35) to t=1.5. The velocity field will then reverse its direction. At t=3.0, the interface comes back to its initial state. The error comparison with the two PLIC methods in [8] is given in Table 1, and shows superior performance for LGB Front Tracking.

Mesh	LGB	Order	CVTNA	Youngs
32^3	5.72×10^{-3}	3.72	7.41×10^{-3}	7.71×10^{-3}
64^3	4.33×10^{-4}	1.82	1.99×10^{-3}	2.78×10^{-3}
128^3	1.23×10^{-4}	N/A	3.09×10^{-4}	7.58×10^{-4}

Table 1. L_1 norms at t = .3 for the LGB method in the three-dimensional deformation simulation compared to the two interface methods used in [8] with *CFL* = 0.5.

The third level of the front tracking code includes applications to the physical problems, especially the CFD code. The code has recently been used to achieve agreement to experimental value of α, the mixing rates for the Rayleigh-Taylor instability [2,3]. The code can interoperate with the LLNL code Overture, which provides adaptive mesh refinement (AMR).

The locally-grid based tracking (LGB) provides robust and accurate resolution for the interface geometry. Conservative tracking [6,5,4,7,9] has been developed in research code and will be included in a future release. It gives higher accuracy for both the interior and the front.

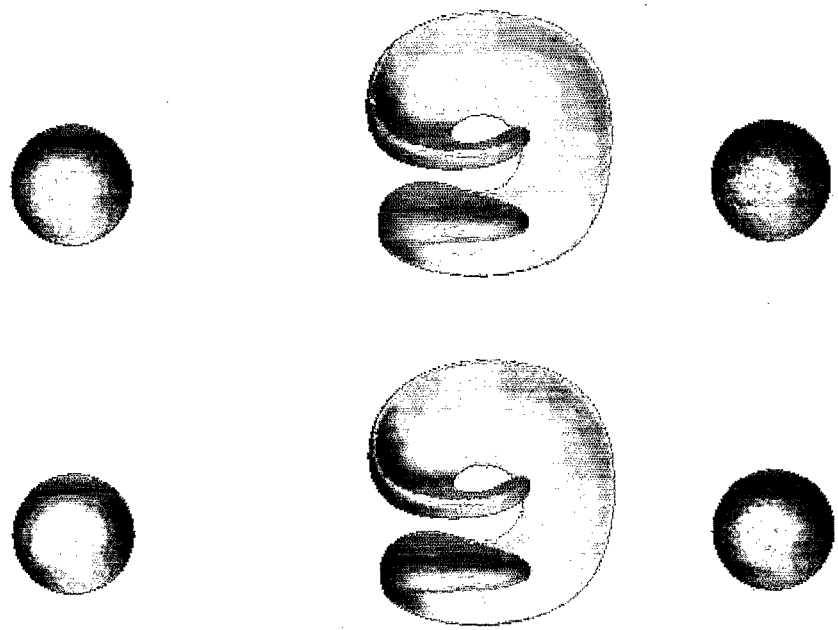


Figure 15. Reversal test of a 3D interface in deformation velocity field with CFL = 0.5. The sequence above has the mesh of 64^3, and the sequence below has the mesh of 128^3. From left to right are t=0, 1.5, 3 respectively.

References

[1] Du, J., Fix, B., Glimm, J., Li, X., Li, Y., and Wu, L. A simple package for front tracking. J. Comp. Phys, in press, 2005. Stony Brook University preprint SUNYSB-AMS-05-02.

[2] George, E. and Glimm, J. Self similarity of Rayleigh-Taylor mixing rates. Phys. Fluids **7:** 054101-1 – 054101-13 (2005). Stony Brook University preprint SUNYSB-AMS-04-05.

[3] George, E. Glimm, J., Li, X.L., Li, Y.H., and Liu, X.F. The influence of surface tension on turbulent mixing rates of immiscible fluids. Phys. Rev. Lett. Submitted, 2005. Stony Brook University preprint SUNYSB-AMS-05-11.

[4] Glimm, J., Li, X.L., Liu, Y.J., Xu, Z.L., and Zhao, N. Conservative front tracking with improved accuracy. SIAM J. Numerical Analysis **41:** 1926-1947 (2003).

[5] Liovic, P., Rudman, M., Liow, J.-L., Lakehal, D., and Kothe, D. A 3d unsplit-advection volume tracking algorithm with planarity-preserving interface reconstruction. Computers and Fluids, submitted, 2005.

[6] Liu, J.-J., Glimm, J., and Li, X.L. A conservative front-tracking method. In *Proceedings of the Tenth International Conference on Hyperbolic Problems: Theory, Numerics, and Applications*, Yokohama Publishers, Osaka, Japan, in press, 2005.

4.3.3 Magnetohydrodynamics of Free Surface Flows at Low Magnetic Reynolds Numbers

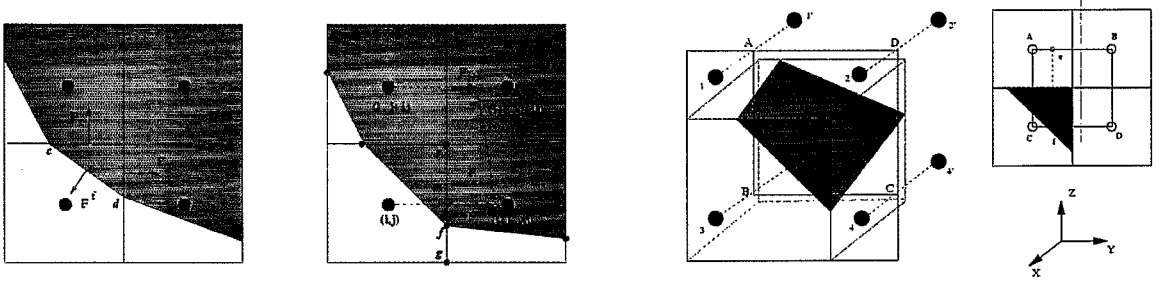
R. Samulyak, J. Du, and J. Glimm

The purpose of this project is to develop mathematical models, numerical algorithms, and computational software for the numerical simulation of fundamental and applied problems of free surface magnetohydrodynamic (MHD) flows of conducting liquids and flows of partially ionized plasmas in the presence of phase transitions and high power particle and laser beams, and to perform numerical simulations of MHD processes in several challenging DOE projects. The simulation effort focuses on the refueling of tokamaks through the injection of frozen deuterium pellets and high energy density gaseous deuterium jets, and the laser driven acceleration of pellets by ablation recoil (rocket) effect as well as liquid metal targets for future particle accelerators.

The system of free surface MHD equations in the low magnetic Reynolds number approximation is a coupled hyperbolic-elliptic system in a geometrically complex moving domain. We have developed a numerical method for solving such equations along with the corresponding parallel software. The numerical method treats the MHD system in an operator split manner. We use the front tracking hydro code FronTier with free interface support for solving the hyperbolic subsystem. The Poisson equation for the electric potential can be solved using techniques for irregular domains described below. FronTier represents interfaces as lower dimensional meshes moving through a volume filling grid. The traditional volume filling finite difference grid supports smooth solutions located in the region between interfaces. The location of the discontinuity and the jump in the solution variables are defined on the lower dimensional grid or interface. The dynamics of the interface comes from the mathematical theory of Riemann solutions, which are idealized solutions of single jump discontinuities for a conservation law. FronTier is capable of tracking 3D interfaces and resolving their topological changes. Some features of the FronTier hyperbolic solvers include the use of high-resolution methods such as Monotonic Upstream-centered Scheme for Conservation Laws (MUSCL), exact and approximate Riemann solvers, and equation of state models for real materials. The existence of a tracked surface, across which physical parameters and the solution change discontinuously, has important implications for the solution of an elliptic or parabolic system. We have developed an algorithm and software for the dynamic generation of finite element meshes conforming to the interface based on the point shift technique. We enhanced recently capabilities of FronTier-MHD to work with complex interfaces in 2D by developing an elliptic solver based on the embedded boundary method [1]. The method uses a finite volume discretization with control volumes being rectangular grid cells away from interfaces, and partial cells near the interface (Figure 16a). To achieve second order accuracy of the solution, fluxes through irregular cell boundaries are calculated using an interpolation of fluxes in neighboring cells (Figure 16b). The corresponding parallel iterative solvers for linear systems of equations are based on PETSc libraries and the algebraic multigrid method implemented in the Hypre software package. The implementation of 3D imbedded boundary elliptic solvers is in progress. We have validated the new MHD code using analytical solutions to idealized problems and applied to the numerical simulation of the Neutrino Factory/Muon Collider target.

The most important modeling problems related to our targeted applications are the description of thermodynamic and mechanic properties of materials interacting with intense sources of external energies. These include nonlinear wave phenomena, cavitation and vaporization of fluids, condensation of gases, the ablation of solids by intense laser or electron beams, laser-plasma interaction, and atomic processes (ionization, dissociation and recombination) in gases. Such models have been developed and implemented in the FronTier code. Namely, homo- and heterogeneous (direct numerical simulation) methods for phase transitions such as cavitation, a technique that would allow numerical simulation of the cavitation in liquid hydrogen jet proposed for the mitigation of plasma disruptions, surface ablation models, and models for the interaction of gas/plasma with hot electrons through a simplified solution of the Fokker-Planck equation [8] have been implemented.

Some of the current applications of the FronTier-MHD code are described in Sections 4.3.5, 4.4.4, and 4.5.1.



a) b) c)

Figure 16. a) Control volumes near the interface and fluxes through irregular cell boundaries. b) Flux interpolation. c) Control volume in 3D and bilinear flux interpolation.

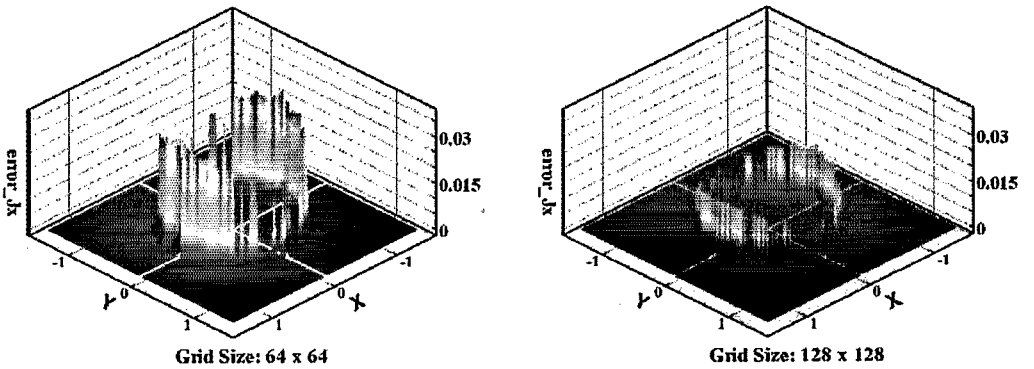


Figure 17. Convergence of the solution of the embedded boundary elliptic problem in a geometrically complex domain with grid refinement. L_2 error of the gradient of the electric potential is shown.

Reference

[1] Samulyak, R., Du, J., Glimm, J., and Xu, Z. A numerical algorithm for MHD of free surface flows at low magnetic Reynolds numbers. J. Comp. Phys., 2005 (submitted).

4.3.4 Molecular Dynamics on the QCDOC

J. Davenport, Y. Deng, and J. Glimm

Molecular dynamics simulations [1,2] are ideal for massively parallel computing because the time spent on communication can be a small fraction of the time spent on computation. However, most molecular dynamics (MD) simulations have been performed on machines where this potential has not been realized due to slow communication limiting the effective use of large numbers of processors.

We have explored [3] the use of a new machine designed for quantum chromodynamics (QCD) by a team of elementary particle physicists [4] mainly at Columbia University, The Riken BNL Research Center and IBM. This machine consists of 12,288 IBM PowerPC processors with extremely fast nearest neighbor communication and 4 megabytes of memory located on each chip. For this reason,

it has been named QCDOC for quantum chromodynamics on a chip. Its peak speed is 10 teraflops ($10x10^{12}$ floating point operations per second) at a cost of roughly $500 per processor.

We project a dramatic level of fine scale parallelism for molecular dynamics simulations based on novel architectural and algorithmic principles, for this machine and its successor, Blue Gene / Light.

The architectural principle is the 6D mesh communication topology, supported by 24 wires per processor (bidirectional send/receive for each mesh dimension). This mesh can be thought of as a programmable distributed switch, and at the level of 10K processors or more, it vastly outperforms technologically feasible switches in terms of effective bandwidth and latency, as well as cost.

The algorithmic principle is fine-scaled parallelism enabled by the rapid communication, i.e. small problem sizes per processor. We plan to compute entirely within L2 cache, and completely avoid the bottleneck of communication delays to main memory. With fine-scaled parallelism, we assign about 10 particles per processor.

We project an ability to simulate 1-10 μsec of real time for all atom MD simulations of 10K to 100K atoms using this machine. We have developed algorithmic and hardware performance models to support this prediction.

References

[1] Allen, M.P. and Tildesley, D.J. *Computer Simulation of Liquid.* Oxford, New York, 1987.

[2] Frenkel, D. and Smit, B. *Understanding Molecular Simulation.* Academic, San Diego, 1996.

[3] Deng, Y., Glimm, J., Davenport, J., Cai, X., and Santos, E. Performance models on QCDOC for molecular dynamics with Coulomb potentials. Int. J. High Performance Computing Applications **18(2)**: 183-198 (2004).

[4] Chen, D., Christ, N.H., Cristian, C., Dong, Z., Gara, A., Garg, K., Joo, B., Kim, C., Levkova, L., Liao, X., Mawhinney, R.D., Ohta, S., and Wettig, T., hep-lat/0011004.

4.3.5 Pellet Ablation in Tokamak Refueling Process

R. Samulyak

The injection of frozen pellets of deuterium and tritium is considered the major mechanism for refueling of nuclear fusion reactors of the tokamak configuration. This problem is significantly important for US fusion facilities and ITER. The ablation of tokamak pellets and its influence on the tokamak plasma has been studied using both semi-analytical and computational approaches. Some analytical models, such as a neutral gas shielding (NGS) model by P. Parks and R. Turnbull [1] has achieved a significant success in explaining experimental data. However, it neglected several important physical and geometrical effects. Computational studies of the detailed ablation physics have been performed using 2D hydrodynamic codes [2,3], thus omitting important MHD effects. A large research effort has also been in the area of large-scale tokamak plasma studies in the presence of an ablating pellet. Such simulations have been performed using a finite element version of MH3D, a full MHD code for tokamak plasma simulations [4]. Details of the ablation were not considered in these works, and the pellet was given as a density perturbation of the initial conditions. MH3D simulations were not able to perform calculations at realistic pellet parameters. An improvement of such large-scale simulations was obtained through the use of an MHD code based on the CHOMBO

AMR package [5]. However, for the pellet ablation, the CHOMBO MHD code used the simplified NGS model, and described the ablation as a moving density source.

Therefore, we can conclude that previous pellet ablation studies have not provided accurate MHD models for all relevant physics processes and the MHD codes for tokamak plasma simulation need an improved model for the pellet ablation. We have recently developed and validated a model for the study of the pellet ablation physics using the FronTier MHD code and models for phase transitions and electron beam - plasma interaction.

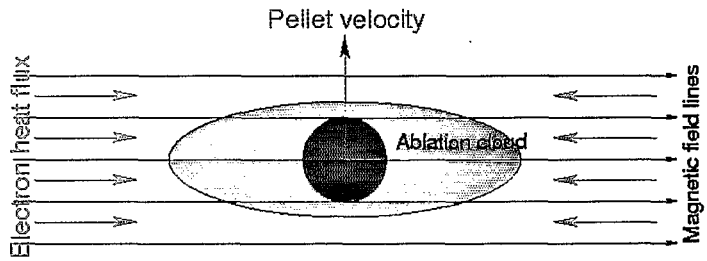


Figure 18. Schematic of processes associated with the ablation of a hydrogen pellet in tokamak.

The schematic of a pellet in a tokamak is shown in Figure 18. In our pellet ablation numerical model, explicitly tracked interfaces separate the solid pellet from the ablated gas, and the cold, dense ablation cloud from the highly conducting fusion plasma. Realistic equations of state are used in different geometrical regions corresponding to different states of matter. A surface ablation model based on thermodynamic properties of hydrogen and deuterium is used for the pellet surface. An electronic heat flux model for the calculation of the thermal energy deposition in the ablation cloud and on the pellet surface uses analytical approximations to the solution of the Fokker-Planck equation, and atomic physics processes in the ablation cloud such as dissociation, recombination, and ionization are taken into account. Pellet ablation simulations have been validated using experimental data and results of other theoretical and computational studies, such as analytical predictions of the NGS model and simulation results obtained with the 2D codes Pellet [2] and CAP [3]. We have studied the transonic ablation flow, its ionization and transformation by MHD forces into an elongated channel inside a tokamak. We have shown that a 1 Tesla toroidal magnetic field forms an ablation channel with the radius approximately equal to 7 pellet radii. Some simulation results are presented in Figures 19 and 20.

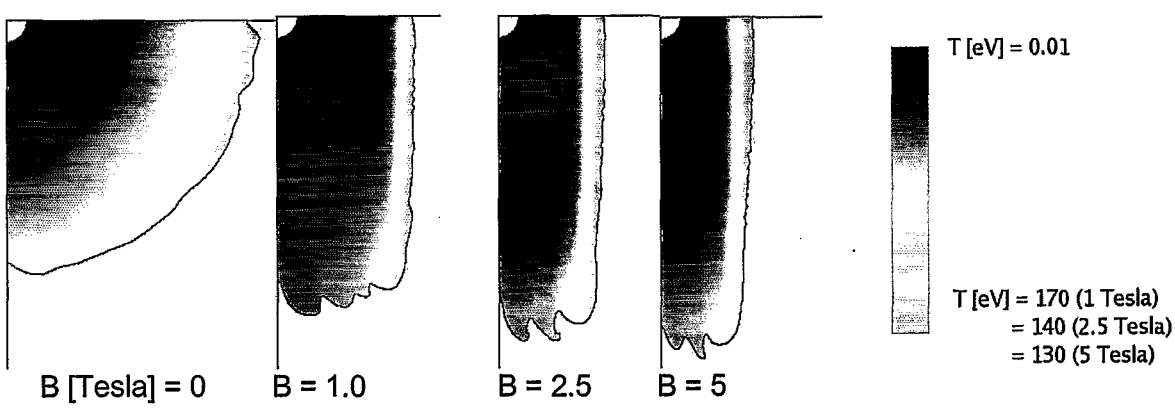
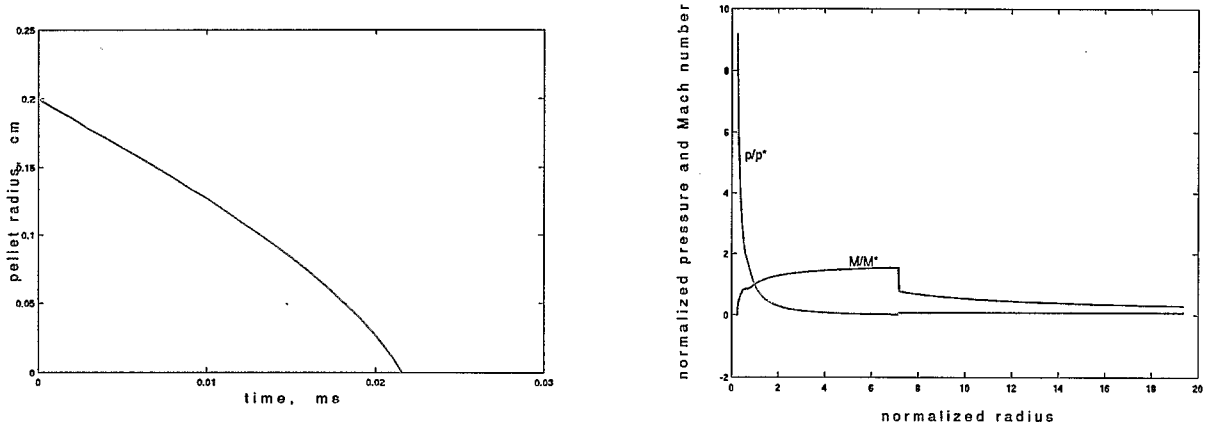


Figure 19. Formation of ablation flow channels near the pellet in magnetic fields ranging from 0 to 5 Tesla. Temperature distribution at 2 microseconds after the pellet injection is shown.



a) b)

Figure 20. a) Evolution of the pellet radius. b) Pressure and Mach number (normalized by the corresponding quantities at the sonic radius in the ablation cloud) along the longitudinal axes of the ablation cloud.

References

[1] Parks, P., and Turnbull, R. Effect of transonic flow in the ablation cloud on the lifetime of a solid hydrogen pellet in plasma. Phys. Fluids **21:** 1735-1741 (1978).

[2] MacAulay, A.K. Geometric, kinetic and atomic physics effects in a two-dimensional time dependent fluid simulation of ablating fuel pellets. Nuclear Fusion **34:** 43-62 (1994).

[3] Ishizaki, R., Parks, P., Nakajima, N., and Okamoto, M. Two-dimensional simulation of pellet ablation with atomic processes. Phys. Plasmas **11:** 4064-4080 (2004).

[4] Strauss, H.R., Park, W., Belova, E., Fu, G.Y., and Sugiama, L.E. 3-D MHD simulations of pellet injection and disruptions in tokamak plasmas Nuclear Fusion **39:** 2069-2076 (1999).

[5] Samtaney, R., Jardin, S., Colella, P., and Martin, D. 3D adaptive mesh refinement simulations of pellet injection in tokamaks. Princeton Plasma Physics Laboratory Report PPPL-3891, 2003.

4.3.6 Parallel Heisenberg Spin Model on Supercomputer Architectures

M. McGuigan and R. Bennett

The Heisenberg Spin Model we studied in [1] is used to describe magnetic materials. The model can be extended to a large number of atoms in order to compare with the bulk properties of magnetic materials whose measurement can involve millions or billions of atoms. In the future we will link our study of the Heisenberg model with density functional calculations of its parameters. An important application area is to high-density storage on nanomagnetic materials.

We measured the performance of a parallel Heisenberg spin model using the Monte Carlo method and the Metropolis algorithm on various supercomputer architectures. These architectures include IBM BlueGene/L, PSC Quadrics Cluster, SGI Altix and QCDOC. This assembly of supercomputer systems probes a variety of supercomputer approaches including shared memory, Linux clusters and specialty machines. BlueGene/L, although originally envisioned as a computer for biomolecular simulation, has an efficient implementation of MPI and can be applied to a variety of problems. PSC Quadrics Cluster is a Linux cluster with a quadric interconnect and is also multipurpose. SGI Altix is a

shared memory with Numaflex interconnect that can also be applied to distributed problems. QCDOC was built originally as a specialized machine for Lattice QCD but has an efficient message passing library called QMP and can also be applied to a wide range of applications, including computational biology [2] and nanoscience.

The Heisenberg spin model of magnetism is defined by the energy

$$E = -J\sum_{nn} \vec{S}_i \bullet \vec{S}_j + \vec{B}_{ext} \bullet \sum_i \vec{S}_i \qquad (1)$$

where $\vec{S}_i$ is a three component spin at lattice site i = (i_1 ,i_2 ,i_3), the sum over *nn* in Eq. (1) is over nearest neighbor lattice sites, and *J* is the nearest neighbor coupling. The number of lattice sites or spins along a given direction is given by *L*. We parallelized the Heisenberg model by using domain decomposition on large lattices up to 16,777,216 atomic grid points. The number of Monte Carlo steps invoked in the Metropolis algorithm is important in reducing the error in the computation so we wanted to study the number of Monte Carlo steps per second that can be achieved on supercomputer architectures. Previous studies of the Parallel Monte Carlo algorithm for the Ising model were performed in [3] where formulas for the number of Monte Carlo steps per second were obtained.

Two of the main properties of parallel computation we studied were strong scaling and weak scaling. Strong scaling means that we fix the problem size, vary the number of processors, and measure the speedup. This is very important for Monte Carlo simulations because one way of reducing the error in a Monte Carlo simulation is to increase the number of Monte Carlo steps. This can be done in a reasonable time frame by increasing the number of Monte Carlo steps per second. Weak scaling means we vary the problem size and the number of processors such that the execution time is the same. In this way we can obtain higher lattice resolution in the same amount of wall clock time. Results from a weak scaling study of QCDOC are shown in Figure 21. The results are in excellent agreement with weak scaling.

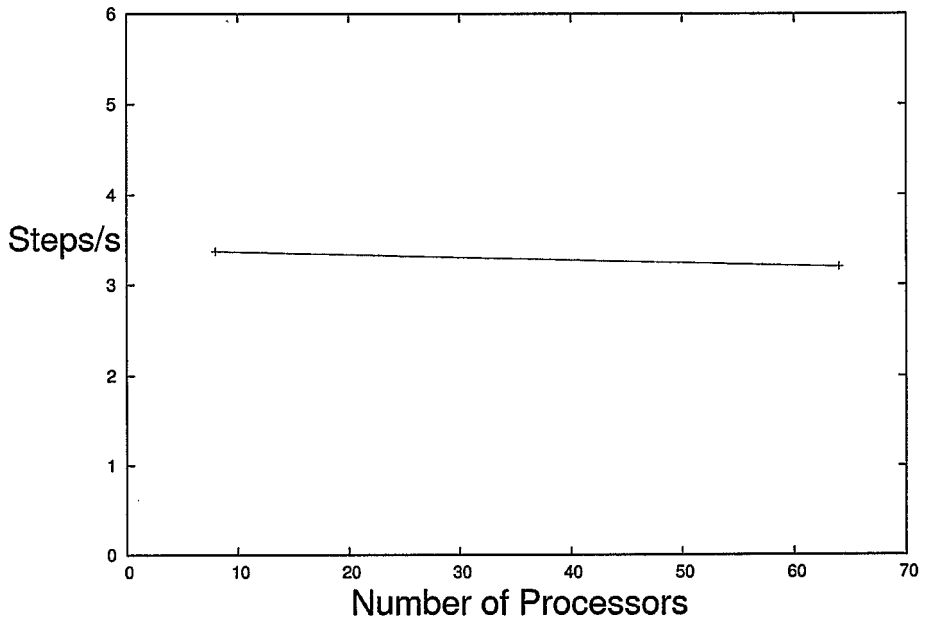


Figure 21. Weak scaling of QCDOC as a function of the number of processors. Performance on the vertical axis is measured in the number of Monte Carlo steps per second.

Using the data from the study we were able to fit a Laurent expansion of the form

$$(Steps/s)^{-1} = Time = aL^3/P + bL^2/P^{2/3} + c \qquad (2)$$

where P is the number of processors. One way to understand this formula is that the first term represents the time spent in computation and the other two represent time spent in communication between processors. This form is general enough to include the Ising model parallel performance derived in [3] as well as the Heisenberg model. Note the equation is consistent with weak scaling.

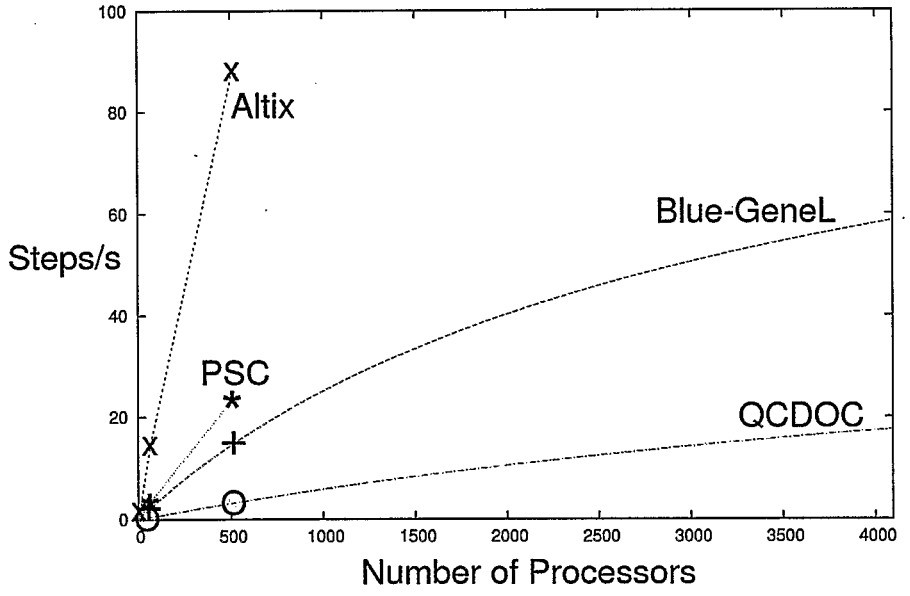


Figure 22. Strong scaling performance curves from various supercomputer architectures using measurements and fits to formula (2).

Fitting the data to the above formula we can obtain an estimate of the performance of the various architectures on large number of processors. The result is shown in Figure 22 for the 256^3 lattice and for a variety of computer architectures BlueGene/L, PSC Quadrics Cluster, SGI Altix and QCDOC applied to the Heisenberg spin model. In all systems we found excellent parallel performance. For example, our results indicate 60 Monte Carlo Steps per second on a 256^3 lattice is possible on a 4096 node BlueGene/L system. This a huge performance gain and will allow the study of large spin systems in the order of a day. On a workstation a similar run could take over a year and make comparison with experiment highly problematic. Our results demonstrate a dramatic increase in productivity in the study of magnetic systems using these leading supercomputer architectures.

References

[1] Bennett, R. and McGuigan, M. Parallel Heisenberg spin model performance on supercomputer architectures. ACM 2006 Symposium on Principles and Practice of Parallel Programming, Sept 2005, submitted.

[2] Deng, Y,. Glimm, J., Davenport, J., Cai, X., and Santos, E. Performance models on QCDOC for molecular dynamics with Coulomb potentials. Int. J. High Performance Computing Applications **18(2):** 183-198 (2004).

[3] Santos, E. and Muthukrishnan, G. Efficient simulation based on sweep selection for 2D and 3D Ising spin models on hierarchical clusters. *Proceedings of the 18th International Parallel and Distributed Processing Symposium (IPDPS'04)* (2004).

4.3.7 Uncertainty Quantification

J. Glimm and Y. Yu

Quantification of Uncertainty and Computer Assisted Decision Making

The need for computer assisted decision making is driven by two related factors. The first is the importance of complex scientific/technical decisions, such as those related to global warming, for which controlled experiments are not feasible. The second is the need for rapid or timely decisions, using incomplete information, such as in shortening the time to market of a product design cycle, mandating a reduction of the role of the human in the loop.

A key issue, and the central one considered here, is an accurate assessment of errors in numerical simulations [1]. Uncertainty quantification (UQ) can be viewed as the process of adding error bars to a simulation prediction. The error bars refer to all sources of uncertainty in the prediction, including data, physics and numerical modeling error. The requirement for UQ comes from the increasing use of simulation model based predictions to guide decision making. In this sense, the need for UQ is a natural consequence of simulation's attainment of a status parallel to that of experiment and theory. Our approach to uncertainty quantification uses a Bayesian framework. Specifically the Bayesian likelihood is (up to normalization) a probability, which specifies the probability of occurrence of an error of any given size. Our approach is to use solution error models as defining one contribution to this likelihood. We provide a scientific basis for the probabilities associated with numerical solution errors.

We have studied UQ for petroleum reservoir modeling [4] and for shock physics simulations [5,2,3], with a focus on statistical analysis of errors in numerical solutions. We seek error models which are robust and understandable. We propose statistical models of uncertainty and error in numerical simulations. We obtained methods to decompose the total simulation error into components attributed to various subproblems. We introduced wave filters to locate significant shock and contact waves. In 1D, a simple error model described the errors introduced at each wave interaction, and (for a spherical geometry) the power law growth or decay of the errors propagated between interactions. A scattering type formula was derived to allow summation of errors propagated through individual interactions to a final error after many interactions.

We have shown that linear models may be sufficient to describe errors even for deeply nonlinear problems. We have shown that complex problems can be decomposed, and models for the total errors can be expressed as a composition of error models for each of the components. In Figure 23 (left) we show the space-time trajectories of the principal waves, as detected by our wave filter program from the untracked solution in a spherical geometry (see also [5,2,3]). In Figure 23 (right), we plot the relative sizes of the variance for each of the six graphs, or trajectories of error initiation and propagation, leading to a component of the total error of the interaction 3 in Figure 23 (right).

The 2D shock interaction problem leads to chaotic interfacial mixing; the central UQ problem here is to define a methodology to describe solution errors for chaotic flow regimes. The challenge is to establish a reliable error analysis for chaotic simulations which do not converge in a pointwise sense, but rather add new complexity with each new level of mesh refinement (see, for example, Figure 24). The solution to this conundrum is to look for convergence in averaged quantities, i.e., the statistical moments, and the spatial, temporal, and ensemble averages which will define them.

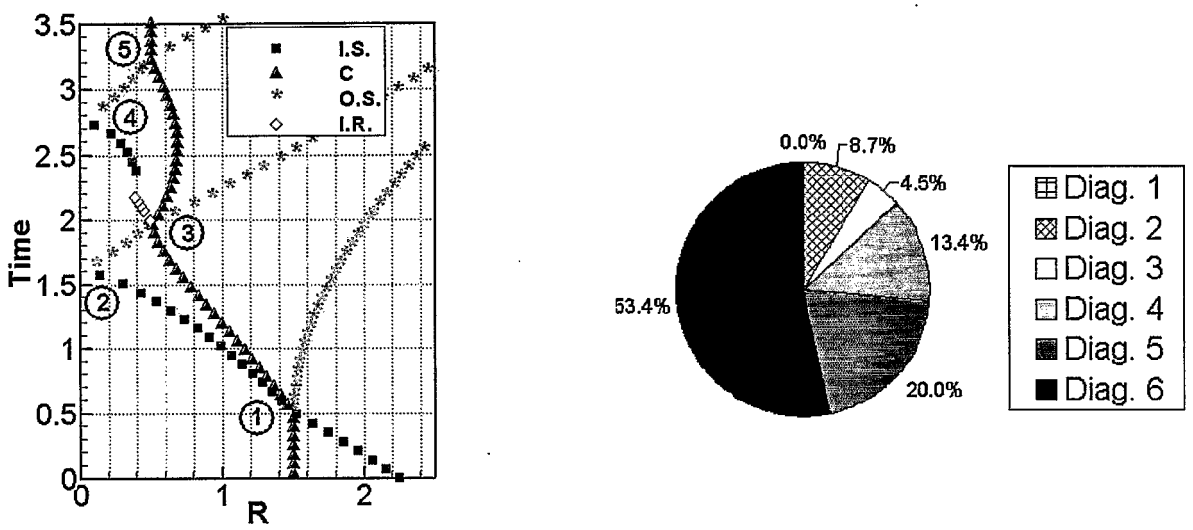


Figure 23. Left: The space-time trajectories of the principal waves, as detected by our wave filter program from the untracked solution in a spherical geometry. Here IS and OS denote inward and outward moving shocks, C is a contact, and IR is an inward rarefaction. Right: Variance of error or uncertainty at the shock contact interaction 3, associated with each of six paths (diagrams) for the error to arrive at interaction 3 starting at an earlier interaction or the initial data. Error determined from a 100 cell simulation of an ensemble of initial conditions.

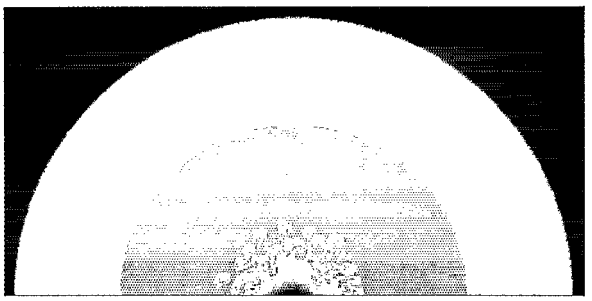


Figure 24. Density plot for a spherical implosion simulation with a perturbed interface. The outer orange-blue boundary is the edge of the computational domain. The red-orange circular boundary is an outgoing reflected shock, and the chaotic inner interface is the object of study. The grid size is 800 x 1600.

The raw data defines a fine scale problem, which can be computed, but not validated or understood. The fine scale data, through averaging, produces coarser scaled data, which is or should be subject to the normal tests of convergence and order of convergence studies, and which may in some cases satisfy its own averaged equations. Thus the problem is very much akin to turbulence modeling, which achieves repeatability only by use of averaged quantities or statistical descriptions of fluctuating quantities. Intrinsic variation, of importance in this study, is analyzed as statistical fluctuations, and these converge not pointwise, but in their statistical character, i.e., through their mean, variance, and possible higher moments. For some variables, convergence of the mean error will fail and then the higher moments of the error carry important fluctuation information (see [6]).

We have also studied the dependence of the solution error on the numerical algorithm [1]. For example, we find that use of Front Tracking reduces the error in a simple spherical implosion or explosion problem by an amount equal to mesh refinement to a factor of 4 to 8 per linear dimension.

References

[1] Dutta, S., Glimm, J., Grove, J.W., Sharp, D.H., and Zhang, Y. Error comparison in tracked and untracked spherical simulations. Computers and Mathematics with Applications **48:**1733-1747 (2004). University at Stony Brook preprint number AMS-03-10 and LANL report number LA-UR-03-2920.

[2] Glimm, J., Grove, J.W., Kang, Y., Lee, T.W., Li, X., Sharp, D.H., Yu, Y., Ye, K., and Zhao, M. Statistical Riemann problems and a composition law for errors in numerical solutions of shock physics problems. SISC **26:** 666-697 (2004). University at Stony Brook preprint number SB-AMS-03-11 and LANL report number LA-UR-03-2921.

[3] Glimm, J., Grove, J.W., Kang, Y., Lee, T., Li, X., Sharp, D.H., Yu, Y., and Zhao, M. Errors in numerical solutions of spherically symmetric shock physics problems. Contemporary Mathematics **371:** 173-179 (2005). University at Stony Brook preprint number SB-AMS-04-03 and LANL report number LA-UR-04-0713.

[4] Glimm, J., Hou, S., Lee, Y., Sharp, D.H., and Ye, K. Sources of uncertainty and error in the simulation of flow in porous media. Comp. and Applied Mathematics **23:** 109-120 (2004). SUNYSB-AMS-03-08. LANL preprint LA-UR-03-2328.

[5] Lee, T., Yu, Y., Zhao, M., Glimm, J., Li, X., and Ye, K. Error analysis of composite shock interaction problems. Conference Proceedings of PMC04, 2004. University at Stony Brook Preprint Number SB-AMS-04-08.

[6] Yu, Y., Zhao, M., Lee, T., Pestieau, MN., Bo, W., Glimm, J., and Grove, J.W. Uncertainty quantification for chaotic computational fluid dynamics. J. Comp. Phys., submitted, 2005. Stony Brook Preprint number SB-AMS-05-16 and LANL p0reprint number LA-UR-05-6212.

4.3.8 ViStA: Visual Statistical Analyzer

K. Mueller and W. Zhu

It is well known that once the data become too large in size and/or dimensionality, automated data mining approaches begin to fail. To address these shortcomings, our current efforts target a classify-refine mechanism that inserts the scientist into a tight loop in the data mining process. Interaction with the system is very natural and intuitive, since the user may reorganize, restructure, or fine-tune certain components of the model or portions of the data directly within the visual display. The changes in the data organization or model are promptly recognized by the statistics engine, which then performs the necessary parameter adjustment/tuning to reflect the changes made by the user. Subsequently, a new data organization or model is computed and presented visually to the user, who may then opt to begin a new iteration cycle. Both our SpectrumMiner (Sections 4.2.3) and our BrainMiner applications use this paradigm with great success [5,8]. Future work will also use this paradigm in our proteomics analysis (Section 4.4.2). In essence, ViStA is the general framework within which all of the different mining tools are unified. In the following we shall give a few examples of the visual tools that are currently available.

Viewing in the native domain. We have developed a 3D visualization interface that displays correlational data in the native domain. This tool is currently mostly used within the BrainMiner project, where the data and the functional relationships implied by the data are displayed within the brain anatomy. The correlation matrix is a 6-D object and a workable tool for its visualization is a challenge. With user selection of a single row or column, the problem reduces to three-dimensional

visualization. For the brain function correlation matrix, these data are presented, along with an MRI volume and a digitized version of the Talairach brain atlas. Both can be sliced in three orthogonal directions and can be overlaid on each other. A basic view with a few regions of interest (ROIs) is shown in Figure 25, which shows the Graphical User Interface (GUI) of our newly developed 3D brain visualization software, along with a basic view of a small number of ROIs embedded into a cut-out area of a normalized/standardized MRI brain. Similar to the 2D viewer, the colors of the ROIs denote the strength of the correlational relationship, on a rainbow scale. The root ROI is colored in yellow. The GUI allows the user to slide the cutting planes up and down and back and forth, to rotate the volume, and to select certain brain surfaces, such as white matter, gray matter, and skull, to be semi-transparently superimposed. The correlation thresholds can also be selected, and many more features are available. The number of ROIs to be displayed, however, can become quite large (about 120-140), which poses challenging problems in the visualization task: in a space too crowded with statistically significant ROIs it becomes very hard, if not impossible, for the user to tell the 3D positions of the individual ROIs. To overcome these difficulties, a number of techniques [1-2] were devised, some of which are illustrated and described in Figure 25.

Viewing in an abstracted feature domain. An alternative way to view and edit causal and hierarchical relationships is in a complementary abstracted feature-centric display. We illustrated this type of display in Section 4.2.3 by way of our SpectrumMiner domain application, but similar strategies will also be available to edit causal models for our BrainMiner domain applications. A particular challenge is imposed by time-varying data. Our interactive dendrogram viewer provides the following two mechanisms to cope with this kind of data: (i) the 4D time-slice selector and (ii) the 3D ThemeRiver. We shall describe these two displays in turn.

The 4D time-slice selector. The interactive dendrogram represents time-varying data as a cylindrical shape composed of a stack of circular time-slice dendrograms. The time-slice selector shown above the dendrogram in Figure 26 shows the unwrapped outer surface of this cylinder, with each horizontal slice capturing the leaf nodes of one time-slice dendrogram. Patterns in the data distribution over time are clearly visible. Stepping across the time slices will animate the data-related coloring of the dendrogram's arcs.

The 3D ThemeRiver. To show the fluctuations of different variables (i.e., nodes in the time-varying dendrogram) over time, we use the ThemeRiver paradigm [3] that visually illustrates multiple variables as parallel streams in which the width of the stream maps to node magnitude. A limitation of the existing ThemeRiver scheme is that only one attribute can be displayed per theme. We therefore have devised a 3D extension, which enables us to display two attributes of each variable in the data stream. The new 3D ThemeRiver [6] can display any ternary relationships within the data, and not just as a function of time. The surface itself is modeled as a smooth 3D

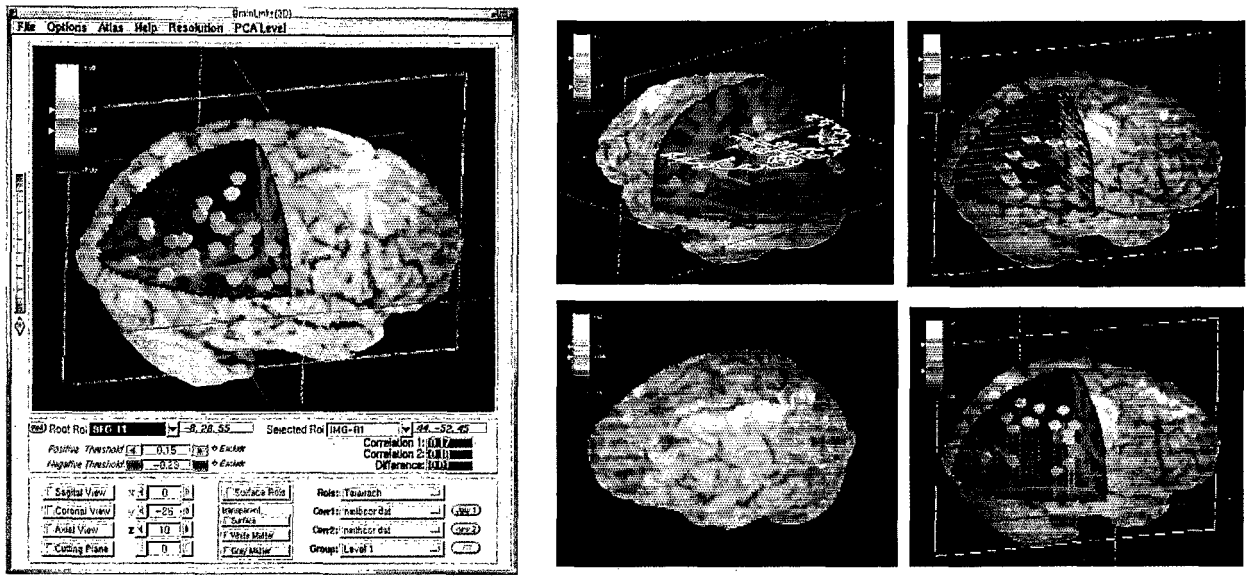


Figure 25. (*Left*) The present 3D visual interface of BrainMiner, along with a basic view of a small number of ROIs embedded into a cut-out area of a normalized/standardized MRI brain. The colors of the ROIs denote the strength of the correlational relationship, on a rainbow scale. The root ROI is colored in yellow. A legend is shown on the left to illustrate the color mapping. The two pointers in the legend can be used to select the range of statistical significance. Only ROIs within that range are displayed. The GUI also allows the user to slide the cutting planes up and down and back and forth, to rotate the volume, and to select certain brain surfaces, such as white matter, grey matter, and skull, to be semi-transparently superimposed. The correlation thresholds can be selected by moving the pointers in the color bar in the upper left corner of the display. When the sliders are moved, the application adjusts the color range such that the entire range of rainbow colors is always utilized. Two studies (baseline and drug) can be loaded simultaneously. The display can then be switched between three correlation matrices: those of the two loaded studies as well as the difference correlation matrix. The latter display shows the ROIs that have changed, under the influence of the drug, their statistical relationship with respect to the selected root ROI; (*center column top:*) superimposing a Talairach atlas slice (or an MRI slice) that can be slid up and down the volume; (*right column top:*) enhancing the ROIs by colored halos or coasters, where the colors code their height and depth on a rainbow color scheme, to aid the perception of 3D depth relations; (*center column bottom:*) near cortex ROI-correlations projected onto the cortex surface; (*right column bottom:*) depth cues provided drop lines.

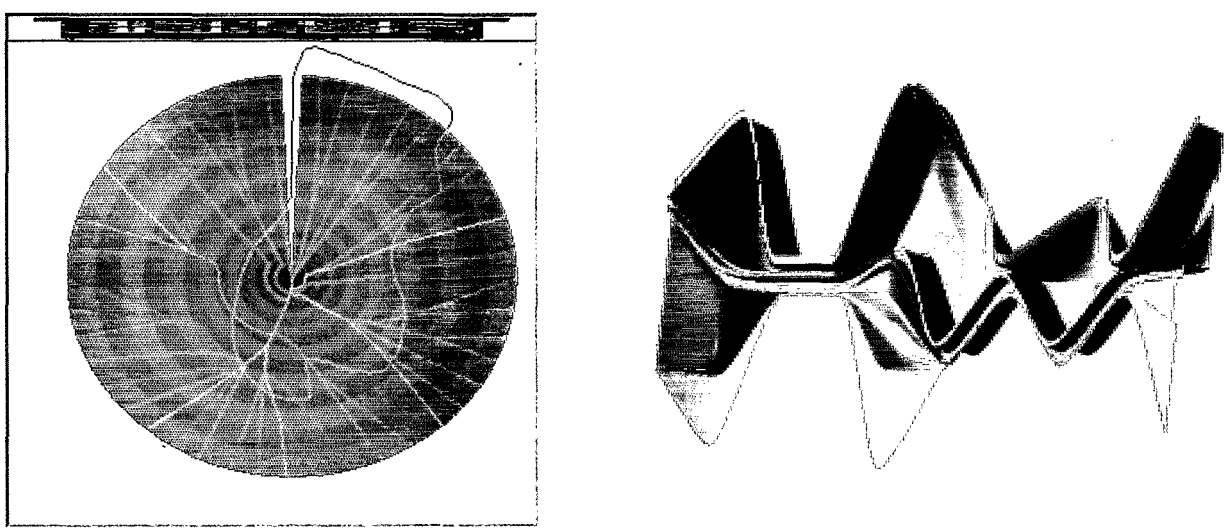


Figure 26. Left: The dendrogram with a user-drawn closed freeform line specifying the time axis-centric surface to be visualized with 3D ThemeRiver. Right: A 3D ThemeRiver visualization of 17 clusters of organic aerosols. Width encodes overall cluster distribution and the height encodes incidence of zinc. The horizontal axis represents the increasing concentrations of ozone in the atmosphere.

Focussed viewing while preserving visual context. In the dendrogram, a number of features are available to enable focus with context, including collapsible sub-trees and node coalescing, search by label and feature, node-content browsers, and non-linear rubber-sheet distortion using an interactive, graphics-hardware-accelerated technique (see Figure 27) [4,7].

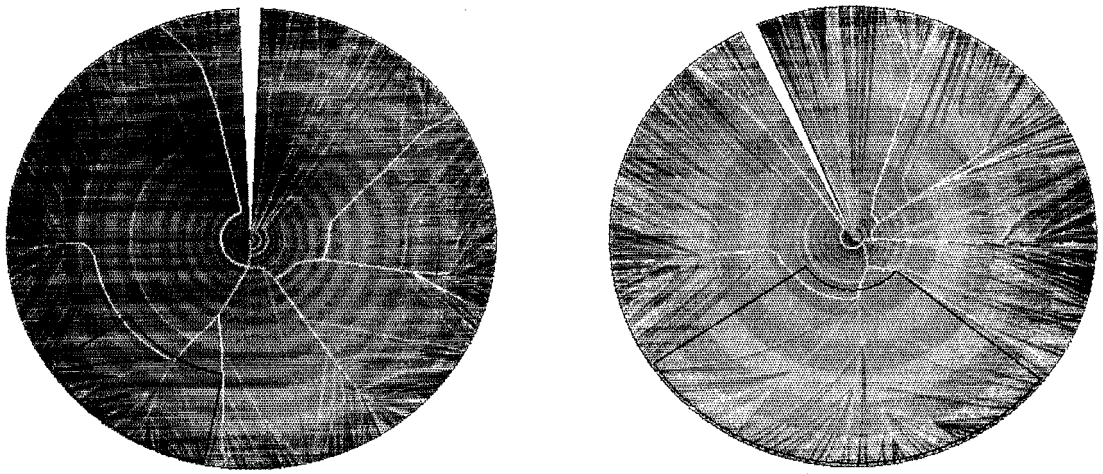


Figure 27. Rubbersheet zooming. Left: unwarped dendrogram; Right: a user-specified arc segment (shown as a black-framed area) is angularly expanded and radially shrunk. Notice the radial distortion of the dendrogram's polar rings.

References

[1] Mueller, K., Welsh, T., Zhu, W., Meade, J., and Volkow, N. Brain Miner: A Visualization Tool for ROI-Based Discovery of Functional Relationships in the Human Brain, *New Paradigms in Information Visualization and Manipulation (NPIVM) 2000*, Washington DC, November 2000.

[2] Welsh, T., Mueller, K., Zhu, W., Volkow, N., and Meade, J. Graphical strategies to convey functional relationships in the human brain: A case study. *Visualization '01*, pp. 481-485, San Diego, October 2001.

[3] Havre, S., Hetzler, E., Whitney, P., and Nowell, L. ThemeRiver: Visualizing thematic changes in large document collections. IEEE Trans. Visualization and Computer Graphics **8(1):** 9-20 (2002).

[4] Imrich, P., Mugno, R., Mueller, K., Imre, D., Zelenyuk, A., and Zhu, W. Interactive poster: Visual data mining with the interactive dendrogram. IEEE Information Visualization Symposium'02, October 2002.

[5] Volkow, N.D., Zhu, W., Felder, C., Mueller, K., Welsh, T., Wang, G-J, De Leon, M. Changes in Brain Functional Homogeneity in Subjects with Alzheimer's Disease. Psychiatry Research Neuroimaging, **114(1)**, 39-50 (2002).

[6] Imrich, P., Mueller, K., Imre, D., Zelenyuk, A., and Zhu, W. Interactive Poster: 3D ThemeRiver. IEEE Information Visualization Symposium '03, October 2003.

[7] Imrich, P., Mueller, K., Imre, D., Zelenyuk, A., and Zhu, W. Interactive Poster: A hardware-accelerated rubbersheet focus+context technique for radial dendrograms. IEEE Information Visualization Symposium'03, October 2003.

[8]. Zhu, W., Volkow, N.D., Ma, Y., Fowler, J.S., and Wang, G-J. Relationships of Ethanol Induced Changes in Brain Regional Metabolism and Motor, Behavioral and Cognitive Functions. Alcohol and Alcoholism. **39(1)**, 53-58 (2004).

4.3.9 Parallel Visualization of Large Data Sets

R. Bennett, M. McGuigan, G. Smith, J. Spiletic, and S. Tomov

In conjunction with the BNL Medical Department, we are developing a Digital Mouse Brain Atlas [1] that can be used by researchers as a standard reference to map anatomical regions and function associated with imaging studies. For example, in Figure 28 we show a probabalistic Atlas where different colors represent different regions of the mouse brain and darker shades indicate less probable results from an imaging study.

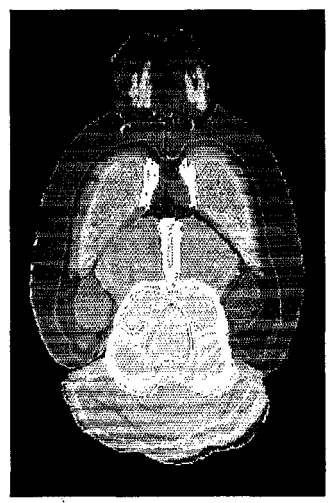


Figure 28: Image from a Probabilistic Mouse Brain Atlas where the colors indicate different anatomical regions and darker shades indicate a lower probability from measurement.

Parallel visualization allows one to handle scientific data sets that are too large or render too slowly on single processor systems. In [2] we implemented interactive parallel visualization by extending popular APIs such as Open Inventor and VTK to support commodity-based clusters. In this work we used Chromium, a popular software tool that provides scalable display technology, to implement our parallel rendering. Figure 29 shows a result from the parallel visualization on four processors of an X-ray tomograph taken of the thigh bone of a rat used in osteoporosis studies. The different colored isosurfaces are handled by separate processors.

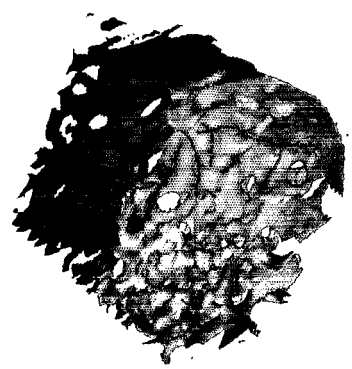


Figure 29. Parallel visualization of the thigh bone of a rat used in osteoporosis studies.

The high performance (40Gflops peak) and low cost ($400) of current graphics cards have given rise to great interest in using Graphics Processing Units (GPUs) for non-graphics scientific computations. In [3] probability-based simulations were implemented on the GPUs and their performance was benchmarked and compared against the CPU with an overall 3 times speedup for the GPU. Such probability-based simulations have a wide range of applications in physics, biology, chemistry and finance. They are computationally intensive and lend themselves naturally to lattice implementations on GPUs. Specifically we demonstrate the techniques using Monte Carlo simulation of the Ising spin model used in physics to study phase transitions in magnetic materials. The GPU Monte Carlo implementation used Cg (C for graphics) fragment programs, which are executed on the GPU by an OpenGL application. We use a standard "Dynamic texturing" programming model, where the computational domain is modeled by a texture, then a GPU fragment program uses the texture to render an image in an off-screen buffer, and finally the texture is updated from the resulting image. Special care is taken (on algorithmic and data representation level) to organize the computations in terms of 4D vector operations, which is important in the current GPUs in order to extract maximal performance. Figure 30 shows a direct visualization and computation on the graphics card of the 3D Ising model with red indicating spin up and blue spin down.

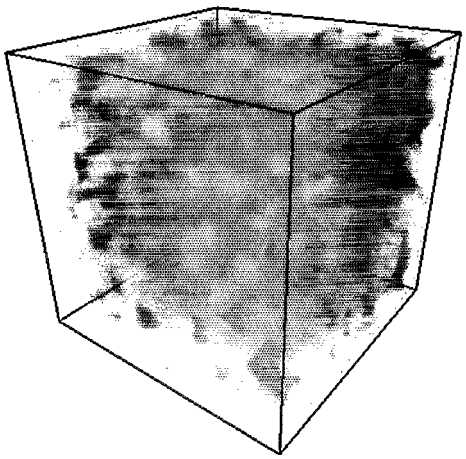


Figure 30. Graphics card computation of 3D Ising model.

References

[1] Benveniste, H., BlackLand, S.J., Ma, Y., Hof, P., Bennett, R., McGuigan, M., and Slatest, L. A three-dimensional digital Atlas database of the adult C57BL/6J mouse brain by magnetic resonance microscopy. Neuroscience **135:** 1203-1215 (2005).

[2] Tomov, S., Bennett, R., McGuigan, M., Peskin, A., Smith, G., and Spiletic, J. Application of interactive parallel visualization for commodity-based clusters using visualization APIs. Computers & Graphics **28(2):** 273-278 (2004).

[3] Tomov, S., McGuigan, M., Bennett, R., Smith, G., and Spiletic, J. Benchmarking and implementation of probability-based simulations on programmable graphics cards. Computers & Graphics **29(1):** 71-80 (2005).

4.3.10 Visualization Collaboratory

Building a Cross-Institutional Collaboratory for 3D Visualization in Technical Education and Training

M. McGuigan, G. Smith, J. Spiletic, BNL, and G. Fong, SUNY Alfred State College

This work builds upon the successes of a previous NSF ATE grant (DUE-0070910) [1]. Therein, two western New York colleges (Alfred State College (ASC) and Jamestown Community College (JCC) partnered with Brookhaven National Laboratory (BNL) in an informal consortium of students, teachers, and scientists focused on the transferal of a cutting-edge, three-dimensional (3D) visualization technology from BNL to the colleges' classrooms. The institutional participants for the current expanded project are the Computational Science Center and the Office of Educational Programs at Brookhaven National Laboratory, SUNY Alfred State College, Bergen Community College, SUNY Jamestown Community College, SUNY Nassau Community College, Rochester Institute of Technology, and SUNY Suffolk County Community College.

The objectives for the current expanded project [2] are: (1) provide students with technical and non-technical experiences not available at their home campuses; (2) update and reinvigorate participating faculty by providing professional development and research collaborations with BNL staff; (3) create 3D Visualization Theaters at the colleges to permit campus-wide involvement in project activities and results; (4) develop specific applications to promote use of the campus visualization facility within mainstream teaching; (5) provide students with an opportunity to collaboratively design and deliver a course focused on training interested campus clientele in the use of the Visualization Theater, thereby providing both a means to disseminate use of the facility on campus and to enhance students' communication skills and curriculum-specific competencies.

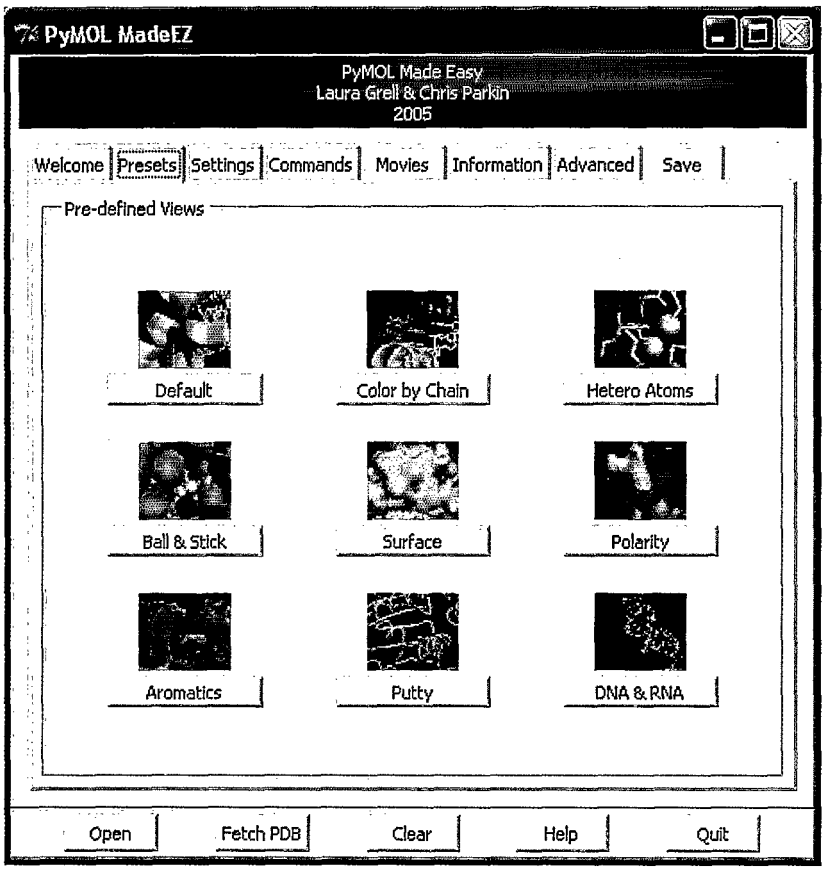
In the first year of this two-year project 18 students from the six campuses worked on nine projects, as developed by the participating faculty members of each college. Two of these projects are discussed here in detail.

PyMOL Made EZ

L. Grell, C. Parkin, Prof. P. Craig, Rochester Institute of Technology, and Leonard Slatest, BNL

PyMOL [3] is a molecular modeling program that can be used in a wide range of studies within the scientific community. Its ability to produce informative, detailed, stereoscopic images makes it a very powerful tool both in the laboratory and in the classroom. However, the PyMOL user interface is difficult to use, and the need for command line interactions is not user-friendly. Using tools from Python's Tkinter and PMW toolkits, a tabbed graphical user interface (GUI) plugin called PyMOL Made EZ has been created. The GUI contains a series of buttons and entry fields that allow users to easily select given attributes of a molecule and make changes to it with the click of a button.

The PyMOL Made EZ GUI features a notebook style design with separate tabs to distinguish one section from another. Some of its features are: (1) easy to use menus and buttons for the selection and manipulation of molecules, (2) nine predefined molecular views that provide clear and meaningful representations, (3) four movie settings that highlight key molecular aspects, (4) a toggle between PyMOL's normal-viewing and stereoscopic-viewing options, (5) one-click access to primary citation and sequence information, (6) a Chime/PyMOL command converter, and (7) the ability to hide PyMOL's internal interface.



Presets Tab in the PyMOL Made EZ

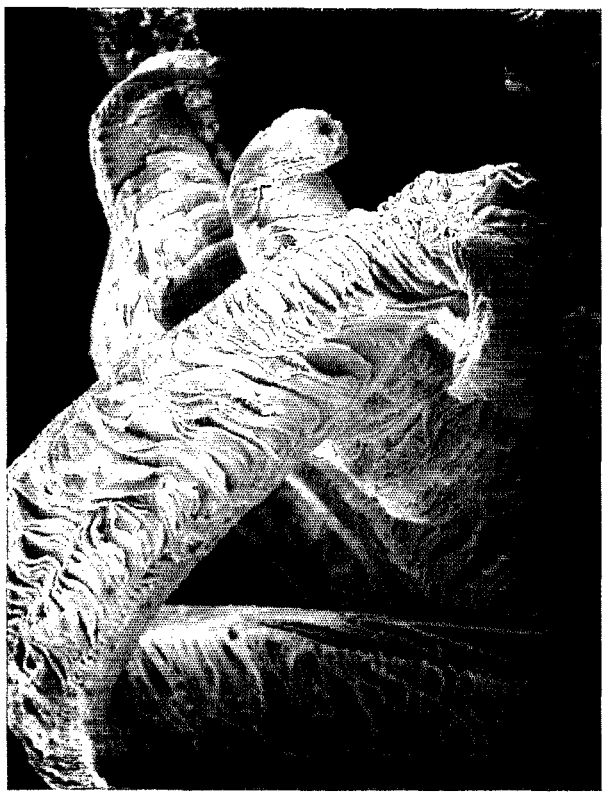
Stereoscopic Imaging of Scanning Electron Micrographs

V. Hall, K. Maggio, T. Smith, Prof. S. Beck, Nassau Community College, and J. Spiletic, BNL

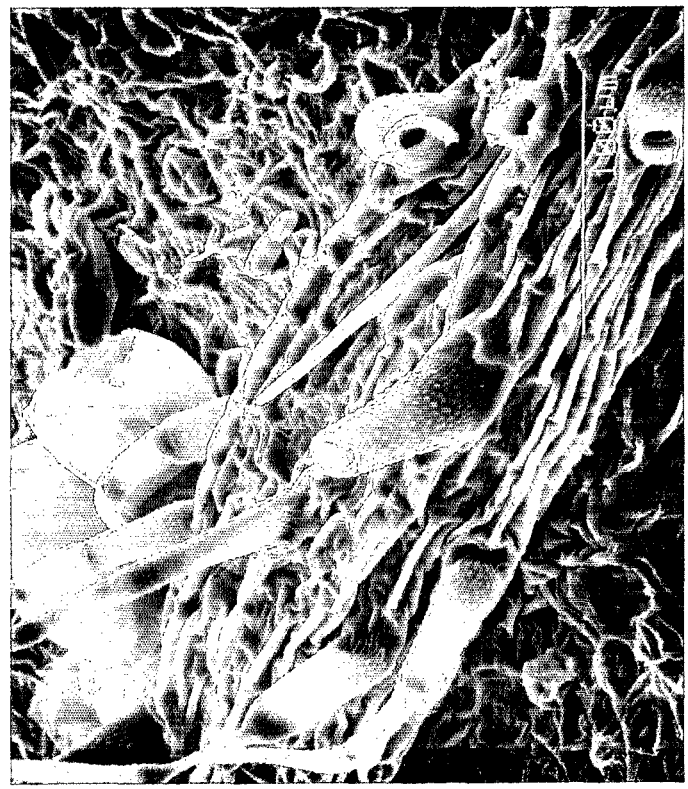
Scanning electron micrographs exhibit stereoscopic qualities due to the high depth of field inherent in the use of the scanning electron microscope (SEM). Using the program StereoPhoto Maker [4], selected micrographs, created with the SEM at Nassau Community College [5], are made viewable as stereoscopic images. A routine procedure has been developed that permits movement from trial-and-error to a method that yields consistent results. As in conventional photography, creation of a stereoscopic image using the SEM requires a left-eye and a right-eye image. SEM stage controls permit sample movement in the X and Y axes, 360 degree rotation, and a tilt of 90 degrees toward and 20 degrees away from the signal detector. Images were captured using successive 2.5 degree increments, horizontal and vertical shifting, and sample rotation. Since specimen rotation could not be measured directly, the SEM-generated scale bar was used to measure increments depending upon magnification. The resultant two-dimensional images were opened in StereoPhoto Maker and combined to form a stereoscopic image based on two micrographs of differing perspectives. In addition to the routine method of generating stereo pairs through sample tilting, other methods were also investigated. It was determined that lateral shifting does not allow for the production of a visually satisfying stereoscopic image. However, rotation produced stereoscopic images of equal quality to that of tilted samples. Correlations have been found between surface topography and the initial degree of tilt required.

The specimens shown in the images (Figure 31) possess characteristics that are desirable for stereoscopic viewing, such as, physical features of varying surface textures and heights, overlapping structures, and inherent depth. The fish gills specimen creates a quality stereoscopic image because

of the overlapping nature of its structures. The structures present in both the foreground and background provide depth to the photo. The leaf is an example of a specimen with varying surface features. Present in the foreground of the photo are hairs and extensive veining. In the background, there are several visible stomates that add variation to the surface topography.



Fish gills



Leaf

Figure 31.

References

[1] Fong, G. et al. A Two Year College Cooperative Applied Research Initiative for Faculty and Students in the Engineering and Science Technologies, NSF Proposal DUE 0070910 (1999).

[2] http://www.itd.bnl.gov/visualization/nsf_ate

[3] http://pymol.sourceforge.net

[4] http://stereo.jpn.org/eng

[5] http://www.sunynassau.edu/webpages/biology/becks.htm

4.3.11 Cluster Computing

N. D'Imperio and E. McFadden

Linux clusters have become important providers of computing cycles in the scientific world. (http://www.ccd.bnl.gov/bcf/cluster/). The CSC cluster, named Galaxy, consists of 374 Intel P3, P4 and Mac G5 processors running at speeds up to 3.2 GHz. Arranged in dual processor nodes, most

have 1 Gbyte of memory per node. A RAID storage facility with a 10 terabyte capacity is attached to this machine. Communication is gigabit ethernet through a Force10 E600 Terrascale switch with 240 ports. Parallel computations are performed using the message passing interface, MPI (www-unix.mcs.anl.gov/mpi) and through threads using Open MP.

Our goal is to continue to explore the frontier of the commodity component world, as processor, motherboard bus, and memory speeds increase, as fiber cable decreases in price, and as switch speeds increase in performance and decrease in price. We will determine optimal configurations and architectures, within the overall framework of our design. Results from this research will be published and posted on the network, to benefit others seeking a similar price/performance advantage.

The Galaxy computer provides computing cycles to CSC affiliates and projects. Convenient and assured access to local parallel computing is an essential requirement for the development of parallel simulation codes, and it is thus a key component of the CSC strategy to make state-of-the-art computing technology available to benefit BNL scientific programs.

4.3.12 Electrodynamics on QCDOC and BG/L

R. Bennett and N. L. D'Imperio

We have studied the performance of our electrodynamics code MAXSSIM (see 4.4.5) on several massively parallel machines. MAXSSIM is a parallel, scalable, finite-difference time-domain code which solves Maxwell's equations using the Yee algorithm [1].

It is possible to choose parameters such that communication among processors is a small fraction of the computation time. Thus it provides an important initial test bed for massively parallel machines.

We have ported the code to QCDOC, BlueGene/L, and to the Compaq alphaserver at the Pittsburgh Supercomputer Center (PSC). We chose a modest sized problem on a 72x72x2048 mesh and evaluated the performance on up to 1024 processors.

Figure 32 shows the actual number of floating point operations (sustained performance) as a function of the number of processors on the three machines. Scaling is quite good in all cases (this is a fixed problem size, so this is an example of strong scaling). However the percentage of peak performance which is achieved is relatively low – approximately 20% on the PSC machine and 5% on BG/L and QCDOC. We are currently investigating the use of other compilers to further understand this issue.

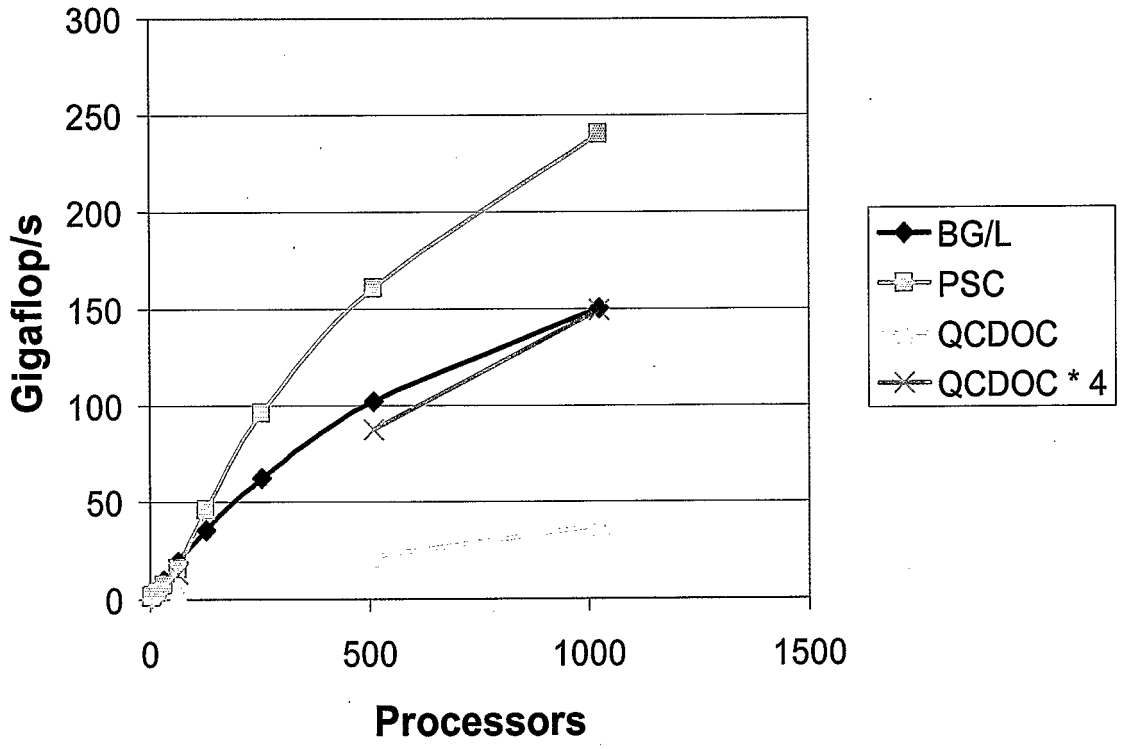


Figure 32. Achieved performance on QCDOC, BG/L, PSC Compaq. QCDOC*4 is 4 times QCDOC performance shown for clarity.

References

[1] Yee, K.S. Numerical solution of initial boundary value problems in isotropic media. IEEE Trans. Antennas and Propagation **14:** 302 (1996).

[2] Bennett, R. and D'Imperio, N.L. Performance Evaluation of a 3D Electrodynamics Simulation on Massively Parallel Computers. Submitted, 2005.

4.3.13 Three Dimensional FFT on QCDOC

Yuefan Deng and Bin Fang

Fast Fourier transforms (FFT) are critical components of many codes, including molecular dynamics. They are also known to stress the communication patterns in massively parallel machines. In this research we have studied a parallel implementation of the FFT algorithm [1,2] on QCDOC, a massively parallel machine originally designed for quantum chromodynamics (see section 4.3.4). These routines will be incorporated into our molecular dynamics code MDoC.

Figure 33 shows the speedup for three-dimensional FFT's of various problem sizes measured on QCDOC for up to 4k processors.

For the largest problem size the parallel speedup scales to larger number of processors than preliminary results for BlueGene/L reported by the IBM group [3] where saturation occurred beyond 512 processors.

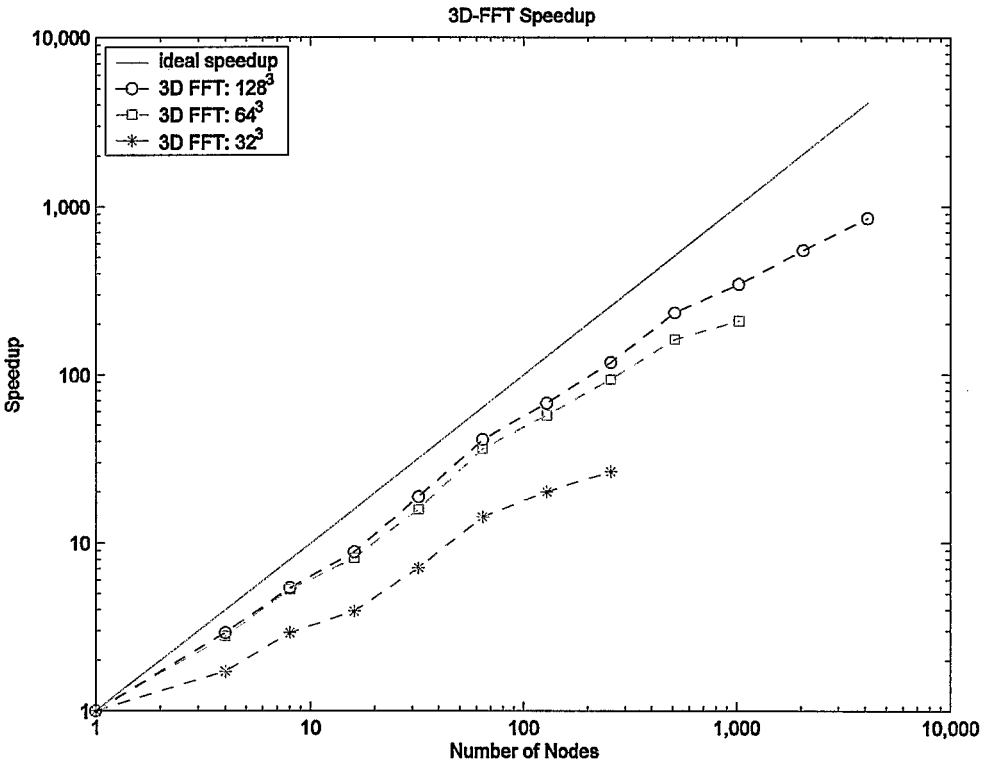


Figure 33. Speed-up for 32^3, 64^3 and 128^3 FFT versus number of nodes.

References

[1] Fang, B. and Deng, Y. Performance of 3D FFT on 6D QCDOC Torus Parallel Supercomputer. J. Comp. Phys. Submitted, 2005.

[2] Haynes, P. and Cote, M. Parallel fast Fourier transforms for electronic structure calculations. Computer Physics Communications **130:** 130 (2000).

[3] Eleftheriou, M., Fitch, B., Rayshubskiy, A., Ward, T.J.C., and Germain, R.S. Scalable framework for 3D FFTs on the Blue Gene/L supercomputer: Implementation and early performance measurements. IBM J. Res. & Dev. **49:** 457 (2005).

4.4 BASIC ENERGY SCIENCES

The Department of Energy's Office of Basic Energy Sciences supports research in materials and chemical sciences, geoscience, engineering, and energy biosciences. This broad-based program of fundamental research is also responsible for a number of national user facilities, including synchrotron x-ray and neutron sources and electron microscope facilities. CSC has engaged in numerous projects in computational fluid dynamics, many body physics, and optics, which are described here. Most projects are in collaboration with BES scientists and support the BES missions.

4.4.1 Superconductivity: Charge Distribution in the Superconductor Magnesium Diboride

J. Zheng, J. W. Davenport, Y. Zhu, and L. Wu

MgB_2 has recently been found to display superconductivity at the surprisingly high temperature of 39K [1]. This fact has stimulated a large research effort to determine the cause of the high transition temperature, as well as other properties of this otherwise unremarkable material. We have utilized

first principles density functional theory to calculate the charge density in MgB_2 and compare it with transmission electron microscope and synchrotron x-ray measurements carried out at Brookhaven. We used the full potential linear augmented plane wave (FLAPW) method to solve the density functional equations [2]. The calculated lattice constants are in good agreement with previous calculations and with experiments. The same technique had been used earlier to study the distribution in energy of the occupied and empty states [3].

MgB_2 forms in a hexagonal crystal structure, with the boron atoms arranged in a honeycomb fashion in planes that are structurally the same as graphite. The magnesium atoms are located in the hollow positions in parallel planes above and below.

Figure 34 shows a contour plot of the difference charge density in the boron plane [4]. The difference density is obtained by subtracting the density of isolated atoms from that of the compound. The red areas between the atoms show clearly the build up of the bond charge.

Experiments measure the structure factors of the charge distribution, which are related to the Fourier transform of the density. The calculations described here agree with the experiments to within 3%. Calculations using only the atomic density, which are typically used in diffraction experiments, disagree with the data by up to 28%, clearly showing the importance of quantum mechanical calculations of x-ray diffraction. We have also studied the details of the electron energy loss spectra finding excellent agreement with experiments [5].

We note that electron scattering is more sensitive than x-rays to the low Fourier coefficients of the charge density [6]. This fact has enabled the accurate comparisons of charge density described here.

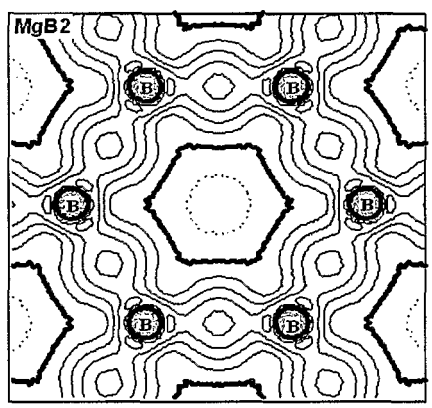


Figure 34. Charge density contours in the hexagonal compound MgB_2 calculated from first principles density functional theory.

References

[1] Nagamatsu, J., Nakagawa, N., Muranaka, T., Zenitani, Y., and Akimatsu, J. Superconductivity at 39K in magnesium diboride. Nature **410**: 63 (2001).

[2] Blaha, P., Schwarz, K., Madsen, G., Kvasnicka, D., and Luitz, J. WIEN2k, an augmented plane wave + local orbitals program for calculating crystal properties (Karlheinz Schwarz, Techn. Universitat Wien, Austria), 2001. ISBN 3-9501031-1-2.

[3] Zhu, Y., Moodenbaugh, A. R., Schneider, G., Davenport, J. W., Vogt, T., Li, Q., Gu, G., Fischer, D. A., and Tafto, J. Unraveling the symmetry of the hole states near the Fermi level in the MgB2 superconductor. Phys. Rev. Lett. **88:** 247002 (2002).

[4] Wu, L., Zhu, Y., Vogt, T. Su, H., Davenport, J. W., and Tafto, J. Valence electron-distribution in MgB_2 by accurate diffraction measurements and first-principles calculations. Phys. Rev. B **69**: 064501 (2004).

[5] Klie, R.F., Su, H., Zhu, Y., Davenport, J.W., Idrobo, J.-C., Browning, N., and Nellist, P.D. Measuring the hole-state anisotropy in MgB_2 by electron energy-loss spectroscopy. Phys. Rev. B **67:** 144508 (2003).

[6] Zheng, J-C., Zhu, Y., Wu, L., and Davenport, J. W. On the sensitivity of electron and X-ray scattering factors to valence charge distributions. J. Appl. Cryst. **38:** 648 (2005).

4.4.2 Scalable Localizable Density Functional Theory

K. S. Kang, J. W. Davenport, D. Volja, J. Zheng, D. Keyes, and J. Glimm

Nanoscale science and technology and biology are driving a search for new ways to calculate the electronic properties of clusters containing many thousands of atoms. Density Functional Theory (DFT) has been shown to yield accurate total energies, charge, and spin densities in molecules and crystalline solids but has been limited in the size system that can be treated.

The density functional equations consist of coupled Schrodinger and Poisson equations, which must be solved self consistently. Usually one chooses a basis set (for example Gaussian type orbitals or plane waves) thereby converting the Schrodinger equation into an algebraic eigenvalue problem. The number of such basis functions will scale linearly with N, the number of atoms in the system. Solving the eigenvalue problem, which is formally dense, will then scale as N^3. Hence for large systems, there is a premium on efficiency, or reducing the number of basis functions required per atom. Gaussian type orbitals (GTOs) have the advantage of being highly localized in space, decaying like $\exp(-ar^2)$. However, they are not maximally efficient, in the sense that many Gaussians are required to accurately represent the wave functions. Plane waves have the disadvantage that they are not localized at all, but rather extend over the whole system. In addition, their use generally requires a "supercell" in which the system is repeated periodically in space, leading to inaccurate treatment of surface and edge effects in some cases.

We solve these problems by using a mixed numerical and localized analytical basis set. We partition the space into nonoverlapping spheres surrounding each atom and an extra-atomic region outside the spheres. Inside the spheres the basis consists of numerical solutions of the Schrodinger (or Dirac) equation for the spherical part of the potential. Outside the spheres, the basis is given by Slater type orbitals, which have the form

$$\phi_{nlm}(\vec{r}) = r^{n-1}\exp(-\zeta r)Y_{lm}(\vec{r})$$

where the Y_{lm} are spherical harmonics. These functions are not overly localized as are GTOs and achieve the same accuracy more efficiently, typically 5x fewer basis elements per atom.

At each sphere boundary these functions are matched onto the numerical solutions inside the sphere. Such a scheme has already been implemented for periodic systems and is known as LASTO, the Linear Augmented Slater Type Orbital method [1]. It is a local orbital version of the Linear Augmented Plane Wave (LAPW) method, considered the most accurate method for solving the DFT equations.

The Poisson equation is solved on a numerical grid using sparse matrix techniques and the hypre library [2]. Hypre is a set of highly parallel preconditioners and solvers suitable for large sparse

systems. The eigenvalue problem is solved using ScaLAPAK [3], a set of codes for the solution of large eigenproblems.

Figure35 shows the density of states in a 13-atom cobalt-nickel cluster, $CoNi_{12}$. The bond lengths were chosen to be the same as in bulk nickel. The discrete eigenvalues were broadened with a Gaussian.

For the largest systems we use an O(N) technique known as "divide and conquer" [5]. In this method, a large cluster is partitioned into subsystems, and the charge density is calculated for each. The computational effort scales like SN_s^3, where S is the number of subsystems and N_s is the number of atoms in the subsystem. This scheme is possible because the charge density (more generally, the density matrix) is localized in space even when the eigenfunctions are not [6].

Using divide and conquer, along with the localized basis set provided by STOs, we expect to be able to calculate the charge and spin density of clusters containing up to 5000 atoms.

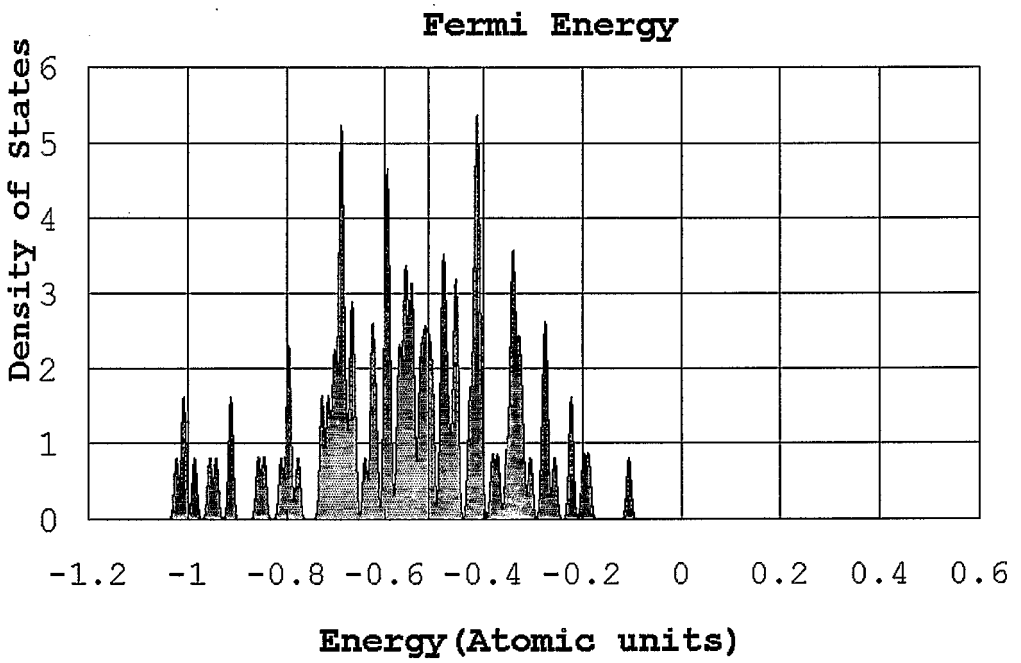


Figure 35. Density of states versus energy for $CoNi_{12}$.

References

[1] Fernando, G.W., Davenport, J.W., Watson, R.E., and Weinert, M. Full-potential linear augmented-Slater type orbital method. Phys. Rev. B **40:** 2757 (1989).

[2] hypre, high performance preconditioners, www.llnl.gov/CASC/hypre/

[3] ScaLAPAK, http://netlib2.cs.utk.edu/scalapack/

[4] PETSc, Portable, Extensible Toolkit for Scientific Computing. www.mcs.anl.gov/petsc/

5] Yang, W. Direct calculation of electron density in density-functional theory. Phys. Rev. Lett. **6:** 1438 (1991).

[6] Kohn, W. Density functional and density matrix method scaling linearly with the number of atoms. Phys. Rev. Lett. **76:** 3168 (1996).

4.4.3 Direct Numerical Simulation of Multiphase Flows with Phase Transitions

T. Lu, R. Samulyak, Z. Xu, and J. Glimm

An accurate description of cavitation and wave propagation in cavitating and bubbly fluids is a key problem in modeling and simulation of hydrodynamic processes in a variety of applications ranging from marine engineering to high-energy physics. We are interested in the study of cavitation and bubbly fluids occurring in high speed liquid jets such as diesel jets in fuel injectors and liquid mercury targets that interact with high intensity proton pulses. Such targets are key components of advanced accelerators such as the Spallation Neutron Source (http://www.sns.gov) and Muon Collider/Neutrino Factory (www.cap.bnl.gov/mumu). Another important application is the simulation of cavitation in a high-speed liquid lithium of helium jet in a device proposed for the plasma disruption mitigation in tokamaks.

Modeling of the cavitation is a complex multiscale and multiphysics problem involving the description of thermodynamic properties of liquids under strong external fields and nonlinear wave phenomena in multiphase system. In the direct numerical simulation of bubbly flows, a liquid-vapor or liquid-non-dissolvable-gas mixture is represented as a system of one-phase domains (vapor bubbles, for instance) separated by free interfaces. The FronTier code is capable of tracking simultaneously a large number of interfaces and resolving their topological changes (the breakup and merger of droplets) in two- and three-dimensional spaces. Though computationally intensive, such an approach is potentially very accurate in treating important effects in bubbly flows including bubble oscillations, heat transfer, drag, viscosity, and surface tension. The method makes it possible to resolve spatial scales smaller than the typical distance between bubbles and model some non-equilibrium thermodynamic features such as finite critical tension in cavitating liquids. The direct method has been validated through the comparison of numerical simulations with theoretical predictions and classical experiments on linear (sound) and nonlinear (shock) waves in bubbly fluids [2,3] (see Figure 36) and applied to the study of the Muon Collider and Spallation Neutron Source targets (see Sections 4.5.1 and 4.5.4).

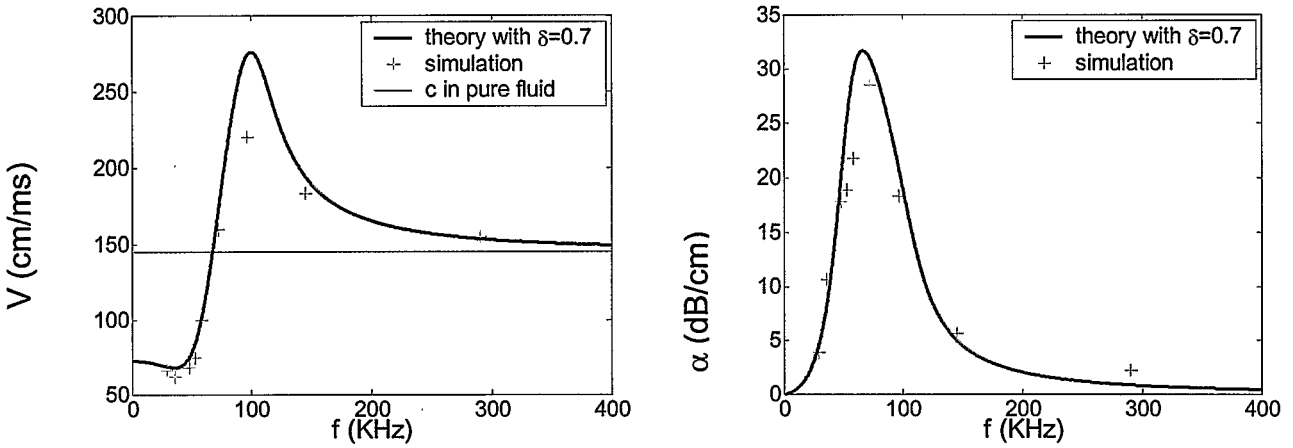


Figure 36. Phase velocity and attenuation rate of linear waves in bubbly water.

Numerical simulation of cavitation presents an additional level of complexity compared to the simulation of wave phenomena in fluids containing small non-dissolvable gas bubbles. The problem is associated with the dynamic creation and collapse of bubbles in the computational domain. The corresponding numerical models and software routines have been developed and implemented in the FronTier code. The phase transition rate is proportional to the deviation of the vapor pressure from the saturated pressure

$$M_{ev} = \frac{\alpha}{\sqrt{2\pi RT}}(p_{sat}(T) - p_{vap}),$$

where α is the condensation coefficient. The Riemann problem for phase transition has been studied using the method of viscous profiles. The transient waves in realistic (thermal conductive) phase transitions with Riemann data have been demonstrated, in particular, by the analytical solution for the linear waves, as shown in Figure 37.

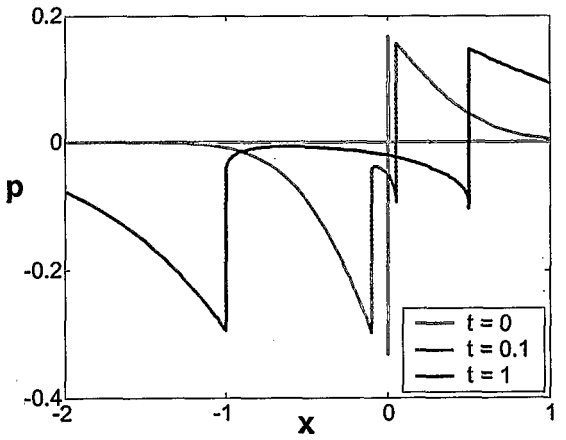


Figure 37. Analytical solution for the transient linear pressure wave induced by the temperature difference in the Riemann data.

To account for the phase transition induced mass transfer across the liquid - vapor interface, a numerical scheme has been developed in the frame of front tracking. A non-local Riemann solver governing the evolution of interfaces has been implemented and a numerical technique has been introduced to account for the thin thermal layer. The algorithm has been validated and applied to physical problems such as the jet breakup in diesel engines.

References

[1] Lu, T., Samulyak, R., and Glimm, J. Direct numerical simulation of bubbly flows and application to cavitation mitigation. J. Fluid Eng., submitted, 2005.

[2] Xu, Z., Kim, M., Oh, W., Glimm, J., Samulyak, R., Li., X., Lu, T., and Tzanos, C. Atomization of a high speed jet. Phys. Fluids, submitted, 2005.

4.4.4 Electrodynamics Simulations: Photonic Devices and RF Cavity Design

N. D'Imperio and J. Glimm

Photonic structures have numerous applications: light emitting diodes, low loss optical devices with increased extracting capability (high Q, low volume and thin cavities, surface grating couplers), lasers, waveguides, low-loss waveguide bends, junctions, couplers and many more. Photonic crystals are a novel class of optical materials fabricated with at least two different dielectric permittivities in a periodic arrangement. They have the ability to suppress, enhance, or otherwise control the emission of light in a selected frequency range by altering the density of states. A complete photonic band gap (PBG), i.e,. a range of frequencies for which light cannot propagate through the crystal in any direction, is a spectral region where the density of states in an infinite crystal vanishes.

RF Cavities are important in accelerator design, providing the means by which the beam is accelerated using active cavities and taking measurements using passive cavities.

We have developed MAXSSIM, a parallel, scalable, finite-difference time-domain (FDTD) [1] code capable of simulating various electrodynamics problems of interest, including photonic crystals. The code can handle complex geometry, realistic initial and boundary conditions, finite size effects including absorbing boundary conditions using the Perfectly Matched Layer (PML) [2], dispersive and nonlinear media, and surface waves. Our primary focus is on the simulation of realistic experimental structures, such as waveguides, laser cavities, and frequency doublers. Currently, we are applying our results on defect modes and localization effects to laser structures that are laterally confined by a photonic crystal and capped with Bragg reflectors longitudinally (PBG-VCSEL's). Design studies comparing PBG-VCSEL's with dielectric post VCSEL's are currently underway in collaboration with Joseph Haus and Andrew Sarangan, both of the University of Dayton.

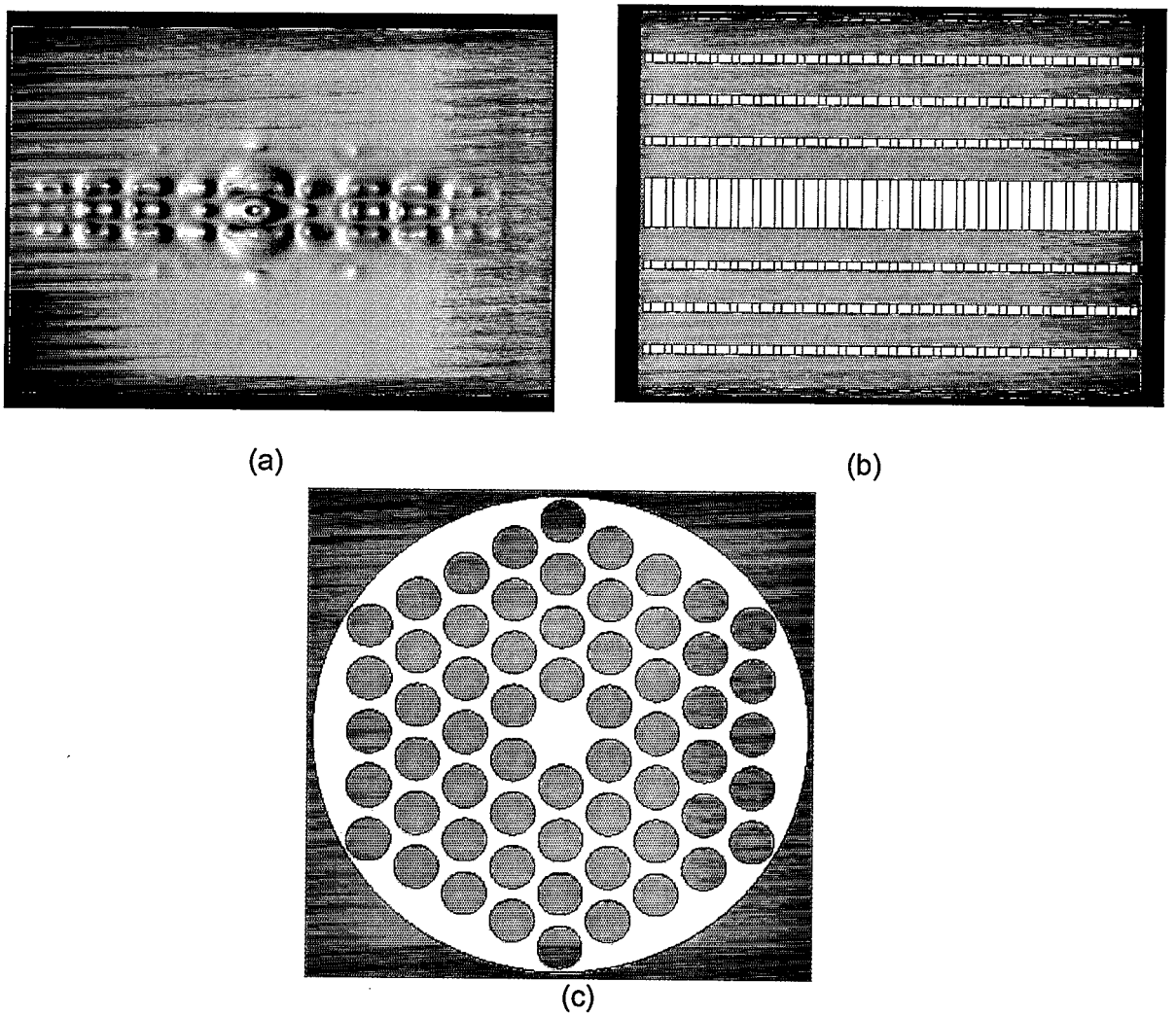


Figure 38. (a) Longitudinal slice showing the electric field energy excited by an electric dipole located in the defect layer of the VCSEL. Energy confinement due to the Photonic Crystal is very high and single mode excitation is evident. (b) Longitudinal cross section of the PBG-VCSEL. Dark blue regions indicate air, light blue and red regions are alternating Gallium Arsenide/Aluminum Arsenide dielectric layers. (c) Transverse cross section of the PBG-VCSEL illustrating the photonic crystal implemented as etched holes.

MAXSSIM has also been applied to RF Cavity modeling and design. One area of interest is studying the effects of cavities on beams as they traverse the structure. See Figure 39. Currently, we are collaborating with the Bates Accelerator Facility of MIT to study the Relativistic Stern-Gerlach Interaction as a tool for attaining spin separation [5].

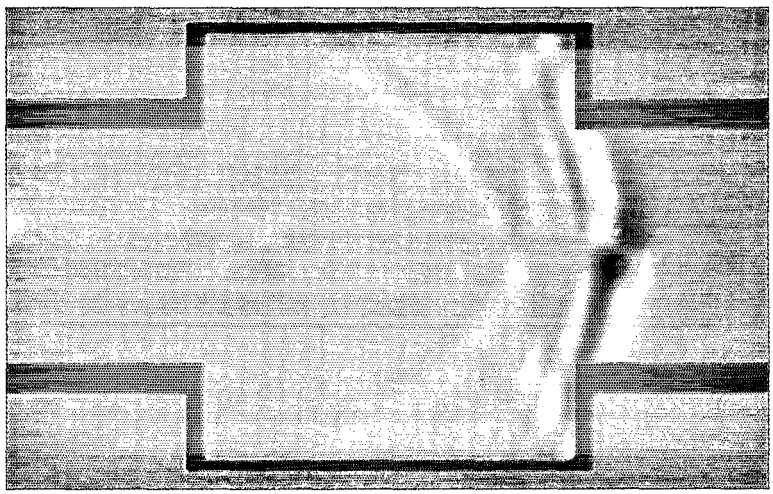


Figure 39. The y component of the electric field generated by a Gaussian beam passing through a rectangular cavity with circular pipes attached. A cross section of the cavity and pipes has been superimposed on the image to more clearly show the boundaries.

References

[1] Yee, K.S. Numerical solution of initial boundary value problems in isotropic media. IEEE Trans. Antennas and Propagation **14:** 302 (1996).

[2] Stojic, N., Glimm, J., Deng, Y., and Haus, J.W. Transverse magnetic defect modes in two-dimensional triangular-lattice photonic crystals. Physical Review E **64:** 056614 (2001).

[3] Smith, D.R., Dalichaouch, R., et al. J. Opt. Soc. Am. B **10:** 314 (1993).

[4] Sakoda, K. Numerical study on localized defect modes in two-dimensional triangular photonic crystals. J. Applied Physics **84:** 1210 (1998).

[5] P. Cameron et al., Design and Test of a Prototype Cavity for a Stern-Gerlach Polarimeter. BNL C-A/AP Technical Note #195, March 2005.

4.4.5 DFT Study of Isocyanide Adsorption on Gold (111) Surface.

Y. Gilman and and P.B. Allen

Isocyanides are molecules terminated with the -NC group, which can serve as the "alligator clip" to connect a molecule to metallic electrodes. Isocyanides may find application in molecular electronics. Several groups successfully fabricated self-assembled monolayers of diisocyanide molecules sandwiched between two gold electrodes and studied electrical transport through them [1]. At Stony Brook University, work is under way to fabricate a molecular transistor based on a single diisocyanide molecule.

We study the details of adsorption of two simple isocyanides, HNC and CH_3NC on the gold (111) surface by density functional theory (DFT) methods. The standard slab calculation scheme is employed, in which the real surface is modeled by the surface of a slab of several atomic layers of gold (see Figure 40). Monolayers at various coverages are considered: one molecule per one Au surface unit cell, one per three, and one per four. We use the WIEN2K [2] DFT code based on the full-potential linearized augmented plane wave (FPLAPW) method. The generalized-gradient approximation (GGA) of Perdew et al. (PBE) is chosen for the exchange-correlation potential.

We find that molecules can adsorb only at the site on top of a Au atom, which agrees with experimental evidence [3]. The adsorption is weak with the adsorption energy of only 0.2 eV for both HNC and CH_3NC molecules. When the local density approximation (LDA) is chosen for the exchange-correlation potential, the results are qualitatively different: adsorption is possible at all sites, with hollow site preferred, and the adsorption energy is about 1 eV. In the absence of adsorption energy measurements, it is difficult to tell whether GGA description of these systems is accurate. We repeated the calculations for the CO molecule (isoelectronic to CNH) for which experimental values of adsorption energies on Au are known. We find that while the LDA overestimates adsorption energies, GGA underestimates them by about 0.2 eV.

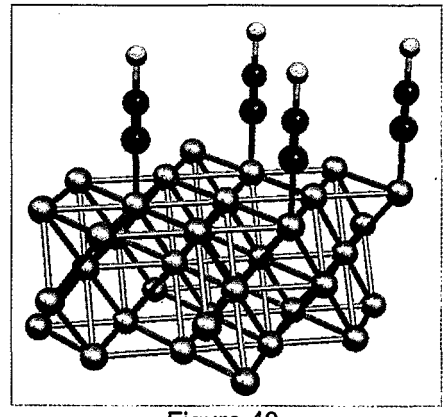


Figure 40.

[1] Chen, J., Calvet, L.C., Reed, M.A., Carr, D.W., Grubisha, D.S., and Bennett, D.W. Chem. Phys. Lett. **313:** 741 (1999); Lee J.-O., Lientschnig, G., Wiertz, F., Struijk, M., Janssen, R.A.J., Egberink, R., Reinhoudt, D.N., Hadley, P., and Dekker, C. Nano Letters **3:** 113 (2003).

[2] Blaha, P., Schwarz, K., Madsen, G.K.H., Kvasnicka, D., and Luitz, J. WIEN2k, An Augmented Plane Wave + Local Orbitals Program for Calculating Crystal Properties (Karlheinz Schwarz, Techn. Universit.at Wien, Austria), 2001. ISBN 3-9501031-1-2.

[3] Robertson, M.J. and Angelici, R.J. Langmuir **10:** 1488 (1994).

4.5 HIGH ENERGY AND NUCLEAR PHYSICS

New major physics facilities are central to the vitality of the nation's HENP programs. Due to their expense and complexity, simulation codes are integral to the design process. Design errors have been a major contributor to past instances of project cancellations and missed target performance. For this reason, design teams will increasingly use high performance computing, developed in national, multi-laboratory collaborations to address simulation issues in the design, operation, and control of accelerators.

4.5.1 Neutrino Factory/Muon Collider Target

R. Samulyak, J. Glimm, and J. Du

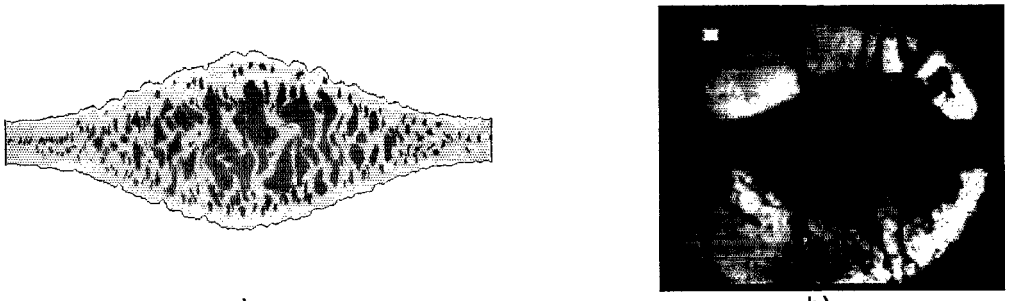
In order to understand the fundamental structure of matter and energy, an advance in the energy frontier of particle accelerators is required. Advances in high energy particle physics are paced by advances in accelerator facilities. The aim of the multi institutional research group Neutrino Factory/Muon Collider Collaboration is to explore the feasibility of a high energy, high luminosity muon-muon collider [1] and a neutrino factory. For more information, visit the Neutrino Factory/Muon Collider Collaboration home page (http://www.cap.bnl.gov/mumu). However, several challenging technological problems remain to be solved in the collider design in order to achieve potential advantages of greatly increased particle energies over traditional electron-positron machines (linear colliders). One of the most important problems is to create an effective target able to generate high-flux muon beams. The need to operate high atomic number material targets in particle accelerators that will be able to withstand intense thermal shock has led to the exploration of free liquid jets as potential target candidates for the proposed Muon Collider. The target will contain a series of mercury jet pulses of about 1 cm in diameter and 30 cm in length. Each pulse will be shot at a velocity of 30 m/s into a 20 Tesla magnetic field at a small angle to the axis of the magnetic field. When the jet reaches the center of the magnet it will be hit with a 2 ns proton pulse. Every proton pulse will deposit about 100 J/g of energy in the mercury.

Numerical simulations of hydro and MHD processes in the target can reduce the amount of costly experiments and help to optimize target parameters. Such simulations present a challenging problem of computational science. They require mathematical modeling of complex flows undergoing phase transitions (cavitation) and numerical methods for solving MHD equations in complex geometries. Our numerical studies have been performed using FronTier-MHD, a magnetohydrodynamic extension of the front tracking hydro code FronTier (for more details, see Section 4.3.3). Numerical simulations of the Neutrino Factory/Muon Collider target have already achieved important results. In our previous work, we addressed problems of the interaction of mercury jet with proton pulses, the evolution of Richtmyer-Meshkov instabilities on the jet surface and the jet breakup, stabilizing effect of the magnetic field, and cavitation on mercury caused by strong rarefaction waves. Simulations agreed with experiments carried out at the BNL Alternating Gradient Synchrotron and CERN.

During last year, we studied cavitation on the mercury jet using both a homogenized equation of state for multiphase fluids and the direct numerical simulation (see Section 4.4.3), as well as the influence of magnetic fields on jet evolution and cavitation. In numerical simulations, we used new elliptic solvers for the MHD system developed recently for the FronTier code (see Section MHD). Numerical results show that pressure waves in the mercury jet caused by the proton energy deposition lead to the formation of a two-phase cavitation domain in the center and the jet (see Figures 41(a) and 41(b). However, the jet cavitation and expansion are strongly reduced in a magnetic field and almost completely suppressed when the magnetic field strength reaches 20 Tesla [see Figures 42(a) and 42(b)] [1]. These results are in agreement with numerical simulations performed earlier using simple approximations. The jet expansion velocity at B = 0 calculated from the graph 1(a) is in a very good agreement with experimentally measured values, which give us a confidence in predictive capabilities of the code. All past experiments have been performed without magnetic fields but the preparation for a new series of experiments involving mercury jets interacting with powerful proton beams and in a 15 Tesla solenoid are underway at CERN, Geneva.

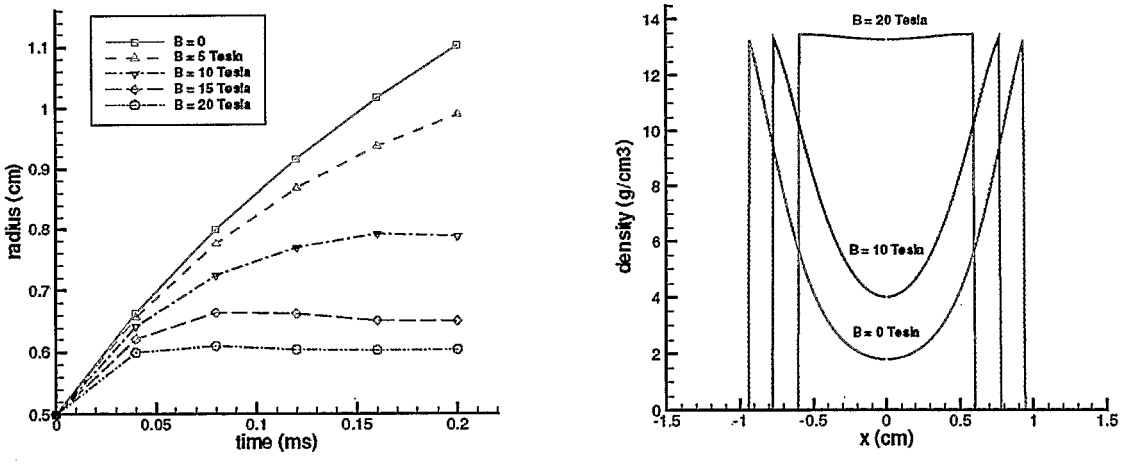
Using perturbation theory for the incompressible system of MHD equations and numerical solutions of equations for first order corrections, we also studied the shape of the mercury jet as it enters a 15 Tesla solenoid at a small angle to the magnetic field axis. This problem is very important for the preparation of new targetry experiments at CERN, as the jet shape may affect the efficiency of the proton pulse interaction with mercury and the pion production rate. We showed that the circular jet

cross-section attains an elliptical shape due to the interaction with eddy currents induced by the transverse component of the magnetic field. The cylindrical shape of the jet is restored by the surface tension forces in the vicinity of the solenoid exit (Figure 43).



a) b)

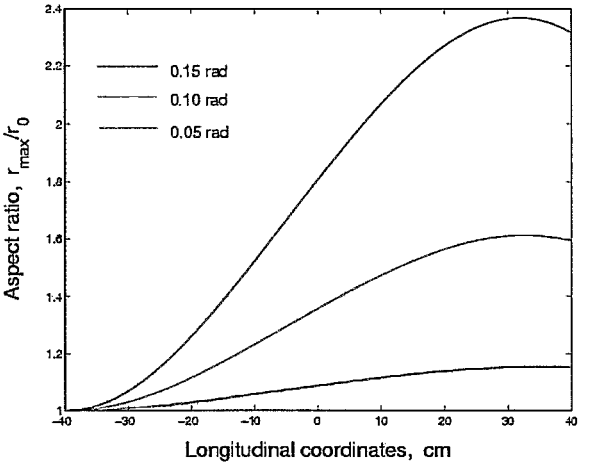
Figure 41. Comparison of cavitation simulations in the mercury target obtained using the direct numerical simulation method, and experimental results. a) Numerically computed density distribution in the mercury jet. b) Experimental image of the mercury jet.



a) b)

Figure 42. Mercury jet evolution in magnetic fields ranging from 0 to 20 Tesla. a) Evolution of the mercury jet radius. b) Average density profile in the jet cross-section at 130 microseconds.

Figure 43. Aspect ratio of the cross-section of the mercury jet entering a 15 Tesla solenoid at three different angles with respect the solenoid axis. Zero of the longitudinal coordinate corresponds to the center of the 1 meter long solenoid. The jet nozzle is placed inside the solenoid at l = -40 cm in coordinate system.



Reference

[1] Samulyak, R., Du, J., Glimm, J., and Xu, Z. A numerical algorithm for MHD of free surface flows at low magnetic Reynolds numbers. J. Comp. Phys., submitted, 2005.

4.5.2 Modeling of Wake Fields and Impedances in Accelerators

R. Samulyak

The electromagnetic interaction of an intensive charged particle beam with its vacuum chamber surroundings in an accelerator plays an important role for the beam dynamics and collective beam instabilities. Wake fields, generated by a moving particle in the accelerator pipe and objects such as RF cavities, bellows, stripline monitors etc., affect the motion of particles in the tail part of the beam causing the parasitic loss, beam energy spread, and instabilities. The effect of wake fields is usually of the same order of magnitude as the space charge effect. While the space charge forces approach zero in the ultrarelativistic limit, wake fields remain finite for an ultrarelativistic beam due to resistivity of the accelerator walls and non-smoothness of the chamber (existence of RF cavities, bellows etc.). The effect of wake fields is an increasingly important issue since operating regimes are continually moving towards higher currents and smaller bunches. To avoid collective beam instabilities that limit the accelerator performance, an accurate numerical modeling of wake fields and their interaction with the beam is necessary.

In the traditional approach for including wake field forces in an accelerator code, the total impedance budget for the accelerator ring is calculated or experimentally measured and the corresponding forces are applied to tracked particles once per beam turn. Such a cumulative force approach is not sufficient for the simulation of beam instabilities caused by wake fields. It is less accurate than the 3D computation of the space charge, which has already been developed in advanced accelerator modeling codes including the MaryLie/Impact code. We have developed a model which accounts for the fine structure of particle beams and distributes wake fields in the accelerator chamber. The corresponding theoretical model is based on the expansion of the particle beam in terms of the multipole moments and the notion of the wake function, which allows elimination of the complex temporal behavior of the electromagnetic field between the incident charge creating the wake field and the test charge. The wake function describes the response of the accelerator chamber element to a delta-functional pulse carrying m-th moment. Wake functions are independent of the beam properties and defined totally by properties of the accelerator chamber [1].

The wake field algorithm is coded as a parallel Fortran 90 module which performs the charge deposition of macroparticle beams on a 3D grid, expansion of the corresponding charge distribution into miltipole moments, computation of wake functions and wake field forces, and interpolation of the wake field forces from the grid to macroparticles. The module has been implemented in MaryLie/Impact code and used for the study of wake field effects on the beam dynamics in simple focusing-defocusing channels and circular accelerators. In the current numerical implementation, most of the accelerator chamber elements (resistive pipe, RF cavity etc.) have associated analytical wake field models valid under certain approximations. Analytical wake field models are beneficial for the study of long-range wake fields and their multiturn effect on the beam dynamics in circular accelerators. To study wake field effects in accelerator elements which cannot be accurately approximated by analytical models, wake functions in a tabular format can also be used. The corresponding data can be obtained through accurate numerical solutions of the full Maxwell system of equations using commercial (MAFIA) or public domain electromagnetic codes.

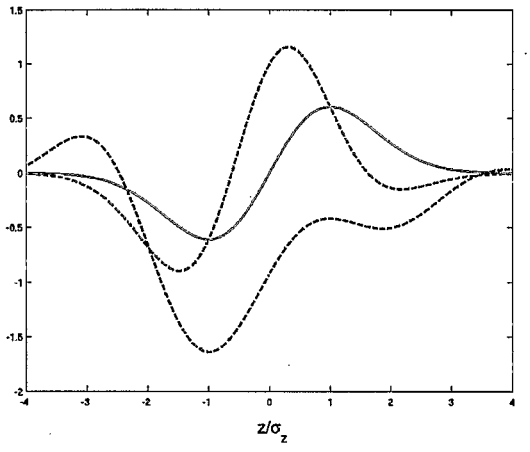


Figure 44. Energy spread due to the longitudinal space charge forces (red line, in units of $q/\gamma^2\sigma_z^2$), zero-mode resistive wake fields (blue line, in units of $\frac{q}{b\sigma_z^{3/2}}\sqrt{\frac{c}{2\pi\sigma}}$), and the resonator model wake fields (green line, in units of $\frac{qR_s c^2}{Q\omega_R\sigma_z^2}$) for a short bunch with the Gaussian longitudinal distribution and a uniform disk transverse distribution.

We have also been working on the development of numerical tools for the study of electron cooling and intrabeam scattering which represent the most important problems of the future upgrade of high-energy accelerators such as RHIC at BNL and Tevatron at FNAL. The self-consistent Langevin approach for solving the Fokker-Plank equation, which computes the diffusion and damping coefficients from first principles based on the Rosenbluth potentials, presents significant computational challenges. Since the distribution function is defined in the six-dimensional phase space, the computation of convolution integrals in the velocity space is required for every spatial point. This was accomplished using an extension of the Particle-in-Cell (PIC) technique in [2].

To reduce the computational time for electron cooling problems which require simulations of long physical times, we have been working on numerical methods for a set of reduction equations. The axial symmetry approximation reduces the dimensionality of the system by two. Using further simplifying assumptions, the system can also be reduced to a set on one-dimensional, coupled, nonlinear, integro-differential equations. The axisymmetric approach may especially be beneficial for the study of non-magnetized electron cooling. Such a cooling method is being evaluated now at BNL as an alternative for the future RHIC upgrade.

References

[1] Chao, A.W. and Tinger, M. *Handbook on Accelerator Physics and Engineering*. World Scientific, 1998.

[2] Ryne, R., Qiang, J., Dragt, A., and Habib, S. Self-consistent Langevin simulation of Coulomb collisions in charged-particle beams, Tech. Paper SC2000.

4.5.3 Unified Accelerator Library: SIMBAD

N. D'Imperio and A. Luccio

Working in collaboration with computational physicists at BNL's Collider Accelerator Department, we have continued development on a Particle-in-Cell (PIC) [1] code to model collective beam effects in two and three dimensions. The code was formerly known as ORBIT and is being integrated into the Unified Accelerator Libraries (UAL) [2] as the SIMBAD component for the modeling of space charge.

UAL provides a framework in which the tracking of particles takes place by pushing a "bunch" of macropartcles through a lattice using the TEAPOT code, which has also been integrated in UAL. Space charge is calculated separately using SIMBAD. Once all particles have reached a certain location, their charge density is calculated by binning to a grid. The potential Φ is found by solving the Poisson equation with the perfectly conducting wall boundary conditions:

$$\Delta\Phi = -\rho/\varepsilon,\ \rho_{wall}(x,y,z) = 0.$$

Space charge force components (with coefficients to account for both the electrostatic and magnetic action) are calculated as derivatives of the potential and applied to each macro particle in the transverse direction. Longitudinal space charge is calculated by binning the particles longitudinally and following the formalism presented in [1]. In a ring with long longitudinal bunches, the transverse motion can be uncoupled from the longitudinal, and the Poisson equation can be solved in parallel in many longitudinal beam segments.

The parallelization of SIMBAD is implemented differently in 2-d as opposed to 3-d. In 2-d, the parallelization is implemented by dividing the macro particles among the processes and collectively calculating the forces on each mesh. In 3-d, each process takes a number of longitudinal slices and calculates the space charge forces only within its own slices. This parallelization requires load balancing that considers both the number of particles and the number of slices. This is accomplished using a genetic algorithm [3]. Figure 45 shows parallel decomposition of the bunch and the resulting division of the beam.

The code was used to model high intensity beams in the Alternating Gradient Synchrotron at BNL to study its suitability as a proton driver [3]. Figure 46 shows the rms transverse emittance vs. turn number. The space charge couples the two transverse coordinates while the two emittances reach a steady state after some turns. A Montague resonance is observed.

UAL has been ported to run on the Blue Gene Light supercomputer where beam-beam interaction will be simulated. In addition, the code will be used to study the long-term impedance effects of the SIS100 accelerator currently being designed and built at GSI in Darmstadt, Germany.

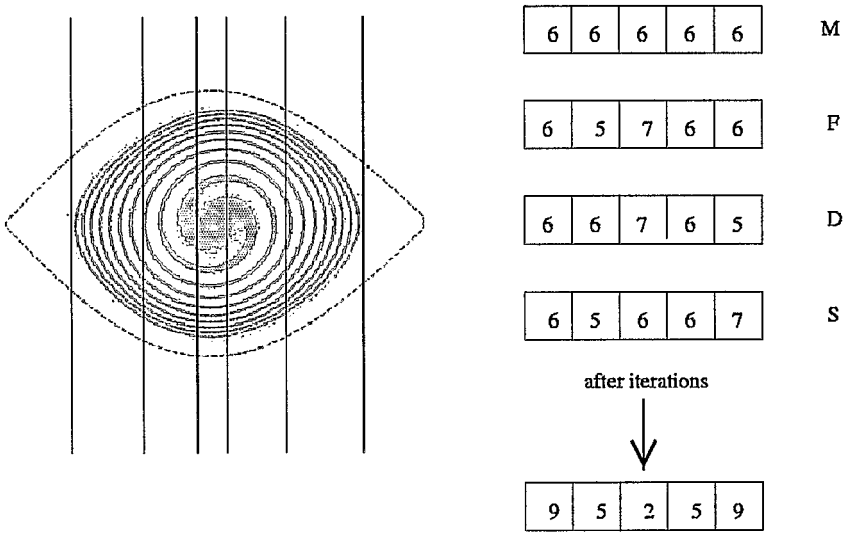


Figure 45. Bunch decomposition and load balancing. (a) A bunch of macro particles in an RF bucket with accompanying lines to indicate the boundaries for each of 5 processes. (b) Genetic algorithm produces an optimal balance between number of particles per process and number of slices per process.

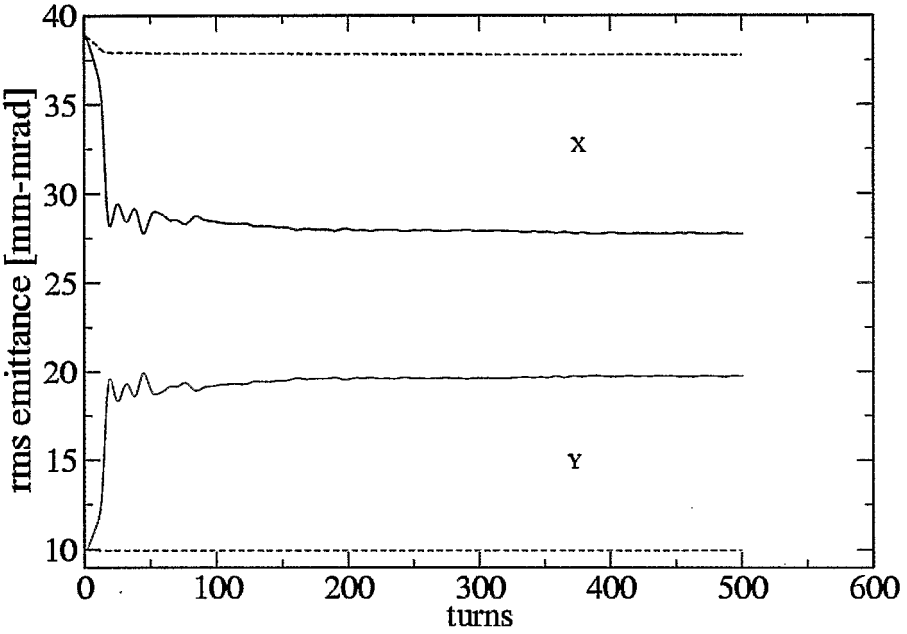


Figure 46. Emittance vs. turns for the AGS at low and high intensity. The dashed lines are low intensity, solid lines indicate high intensity. Transverse coupling is very noticeable at high intensity.

References

[1] Chao, A.W. *Physics of Collective Beam Instabilities in High Energy Accelerators.* Wiley, 1993.

[2] Malitsky, N. and Talman, R. Unified Accelerator Libraries. In *Computational Accelerator Physics,* AIP Conference Proceedings 391, J.J. Bisognano and A.A. Mondelli, Eds., pp. 337-342 (1997).

[3] Luccio, A. and D'Imperio, N. Simulation of the AGS as a proton driver. ICFA Beam Dynamics Mini Workshop on Space Charge Simulation, Oxford, UK, Trinity College, April 2-4, 2003, www-bd.fnal/icfa/workshops (2003).

4.5.4 Spallation Neutron Source

R. Samulyak and T. Lu

The Spallation Neutron Source (SNS) is an accelerator-based neutron source being built in Oak Ridge, Tennessee, by the U.S. Department of Energy (http://www.sns.gov). The SNS will provide the most intense pulsed neutron beams in the world for scientific research and industrial development.

The proposed liquid mercury target design for the Spallation Neutron Source (see Figure 47) includes a main flow region inside a stainless steel structure where mercury enters from the sides, flows around a baffle into the proton beam path, and exits out the center. A cooling jacket that wraps from bottom to top around the target is used to cool the target window through which the proton beam enters. The stainless steel target structure is approximately 0.5 x 0.4 x 0.15 m^3.

One of the most important issues associated with using liquid metals as targets for pulsed proton beams is withstanding the loads caused by the rapid pressure increase resulting from the intense heating of the liquid metal from a single pulse of protons. This heating occurs essentially instantaneously compared to acoustic time scales; therefore, the mercury undergoes a large pressure increase. In addition to a set of difficult engineering problems associated, for instance, with the design of windows able to withstand large thermal gradients and shocks, recent experiments with an SNS target prototype uncovered yet another problem critical to the target lifetime. They showed pitting of stainless steel surfaces that were in contact with mercury subject to large pressure pulses induced by the collapse of cavitation bubbles [1]. Due to the cavitation-induced erosion, it will be necessary to replace the target after two weeks of operation at frequency 60 Hz of a 1 MW proton pulse. To extend the target lifetime, future research efforts will be concentrated in two areas, each of which should lead to reduction of the erosion damage:

- Evaluation of cavitation resistant materials and coatings.
- Investigation of mitigation techniques such as introduction of non-dissolvable bubbles into the system.

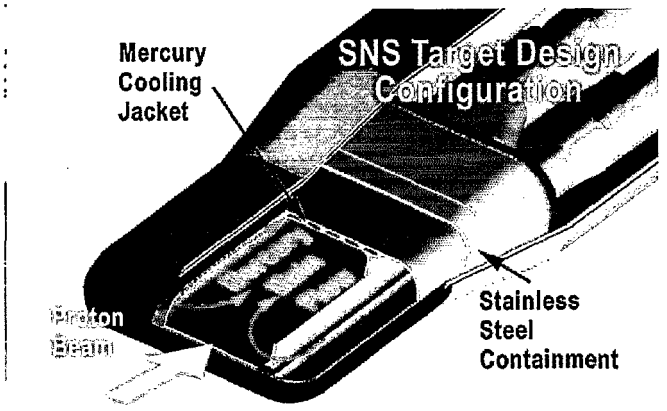


Figure 47. SNS mercury target design.

We have applied the direct numerical simulation technique for bubbly fluids (see Section 4.4.3) to the study of pressure mitigation through the injection of non-dissolvable gas bubbles near the target front window. We have found that while the bubbly layer indeed causes a significant reduction of pressure during 200 microseconds, large transient pressure oscillations exist for a short period of time (< 100 microseconds) after the proton beam energy deposition (see Figure 48). We have studied the formation and evolution of cavitation bubbles in mercury caused by the pressure distributions depicted in Figure 48. The collapse pressure of cavitation bubbles was calculated by solving the Keller equation. The mitigation efficiency, estimated by performing statistical averages of pressure

peaks, was found to be dependent on the parameters of the bubbly layer such as the volume fraction and average bubble size. For example, a bubbly layer with the average bubble size R = 0.5 mm and a 0.53% volume fraction reduces the integral effect of cavitation induced pressure peaks by 50 times.

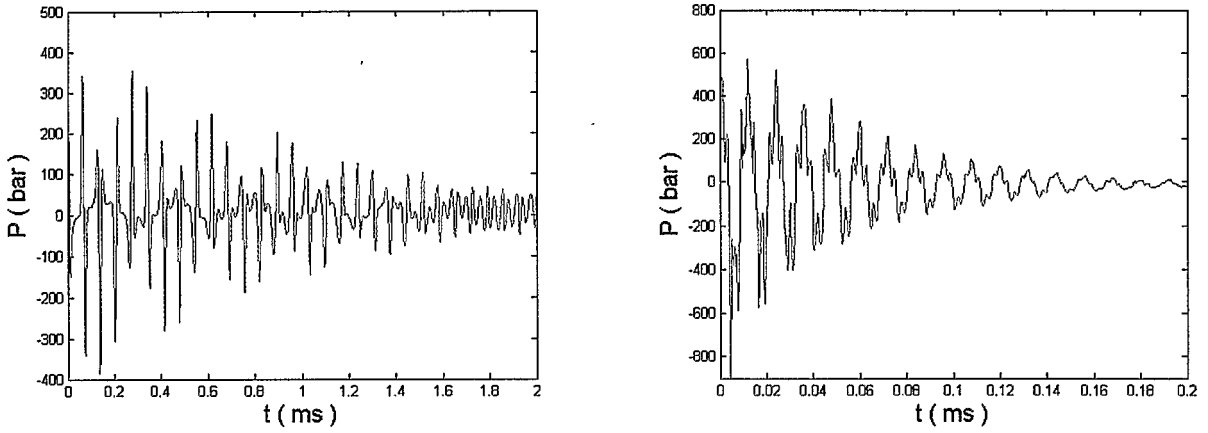


Figure 48. Proton pulse induced pressure peaks on the entrance window in the pure mercury (left) and mercury containing gas bubbles (right).

References

[1] Status Report on Mercury Target Related Issues, SNS-101060100-TR0006-R00, July 2002.

[2] Lu, T., Samulyak, R., and Glimm, J. Direct numerical simulation of bubbly flows and application to cavitation mitigation. J. Fluid Eng., 2005 (submitted).

4.5.5 Grid Computing: MonALISA

Development and Use of MonALISA High Level Monitoring Services for the STAR Unified Meta-Scheduler

E. Efstathiadis, L. Hajdu, J. Lauret, and I. Legrand (CalTech)

As a Particle Physics Data Grid (PPDG) cross team-project we study, develop, implement and evaluate a set of tools that allow Meta-Schedulers to take advantage of a consistent set of shared information (such as information needed for complex decision making mechanisms) across both local and Grid Resource Management Systems. We demonstrate the usefulness of such tools within the MonALISA monitoring framework and the STAR Unified Meta-Scheduler.

We define the requirements and schema by which one can consistently provide queue attributes for the most common batch systems and evaluate the best scalable and lightweight approach to access the monitored parameters from a client perspective and, in particular, the feasibility of accessing real-time and aggregate information. Client programs are envisioned to function in a non-centralized, fault tolerant fashion. We believe that such developments could highly benefit Grid laboratory efforts such as the Grid3+ and the Open Science Grid (OSG).

The MonALISA (Monitoring Agents in A Large Integrated Services Architecture) system provides a distributed monitoring service. It is based on a scalable Dynamic Distributed Services Architecture (DDSA) that is implemented using JINI/JAVA and WSDL/SOAP technologies. The scalability of the system derives from the use of autonomous, multi-threaded station servers to host a variety of loosely coupled, self-describing, dynamic services, the ability of each service to register itself and then to be

discovered and used by other services or clients that require such information, and the ability of all services and clients subscribing to a set of events (state changes) in the system to be notified automatically. The framework integrates several existing monitoring tools and procedures to collect parameters describing computational nodes, applications and network performance. It has built-in SNMP support and network-performance monitoring algorithms that enable it to monitor end-to-end network performance, as well as the performance and state of site facilities in a Grid.

The core of the MonALISA monitoring service is based on a multithreaded system (the monitoring service) used to perform the many data collection tasks in parallel, independently. It is designed to easily integrate existing monitoring tools and procedures and to provide this information in a dynamic, self-describing way to any other services or clients. MonALISA services are organized in groups and their group attribute is used for registration and discovery. Each service registers with a set of JINI Lookup Discovery Service (LUS), as a member of a group, and having a set of attributes. The LUSs are also JINI services and may be registered with other LUSs, resulting in a distributed and reliable network for registration of services. Services also provide the code base for the proxies that other services or clients will need to instantiate for using it.

A generic framework for building pseudo-clients for the MonALISA services was developed. This has been used for creating dedicated web service repositories with selected information from specific groups of monitoring services. The pseudo-clients use the same LUSs approach to find all the active MonALISA services from a specified set of groups and subscribe to these services with a list of predicates and filters. These predicates or filters specify the information the pseudo-client wants to collect from all the services. Pseudo-clients store received values from the running services in a local MySQL database. A Tomcat based servlet engine is used to provide a flexible way to present global data and to construct on the fly graphical charts for current or customized historical values, on demand. Multiple Web Repositories can easily be created to globally describe the services running in a distributed environment.

Queue monitoring data collected using custom MonALISA modules at each site are cached locally using pseudo-clients with fail-over capabilities. This way the monitoring data from all services that have joined a group are available to policies implemented into the Meta-Scheduler that choose the appropriate queue for the submitted job.

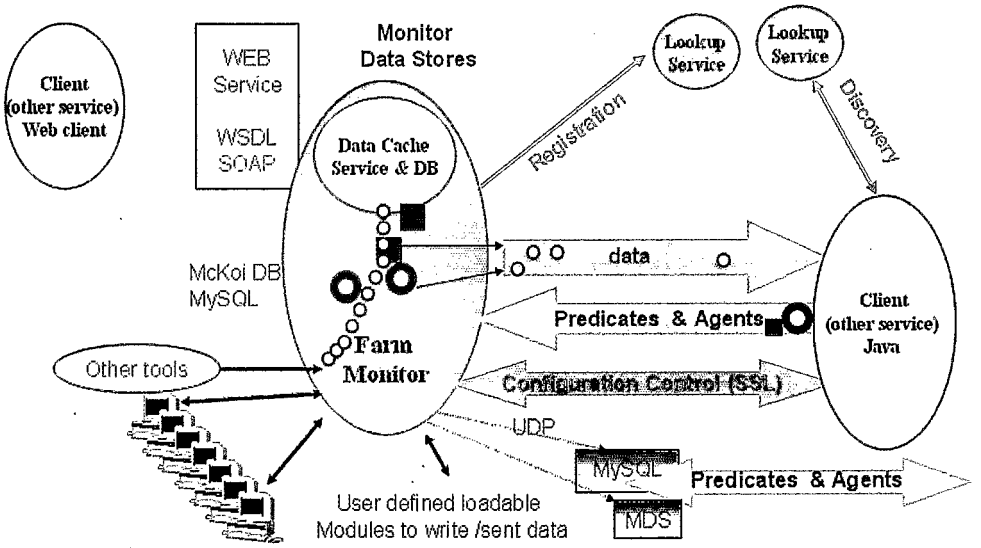


Figure 49. The MonALISA monitoring service.

In addition, a MonALISA service has been deployed at BNL for the monitoring needs of the UltraLight project [4]. UltraLight is a collaboration of experimental physicists and network engineers whose

purpose is to provide the network advances required to enable petabyte-scale analysis of globally distributed data. Current Grid-based infrastructures provide massive computing and storage resources, but are currently limited by their treatment of the network as an external, passive, and largely unmanaged resource. The goals of UltraLight are to:

- Develop and deploy prototype global services which broaden existing Grid computing systems by promoting the network as an actively managed component.
- Integrate and test UltraLight in Grid-based physics production and analysis systems currently under development in ATLAS and CMS.
- Engineer and operate a trans- and intercontinental optical network testbed, including high-speed data caches and computing clusters, with U.S. nodes in California, Illinois, Florida, Michigan and Massachusetts, and overseas nodes in Europe, Asia and South America.

MonALISA can be used to monitor and control network devices, such as routers and photonic switches. Since it gathers information system-wide, MonALISA is able to generate global views of the prevailing network connectivity, to identify network or end-system problems and act on them strategically, or locally as required. Services that take decisions based on these (global) system views can be created and deployed: for example, mobile agents that are able to provide optimized dynamic routing for distributed applications have recently been added to MonALISA.

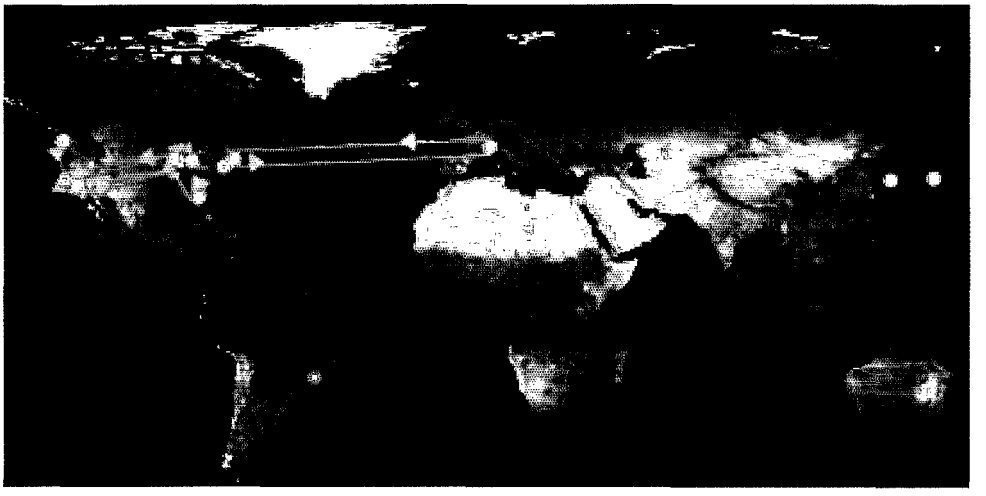


Figure 50. The UltraLight testbed includes sites in the Americas, Europe and Asia.

References

[1] Efstathiadis, E. et al. Proceedings of the Computing in High Energy Physics (CHEP) 2004 conference, Interlaken, Switzerland, Sept. 26 - Oct. 1st, 2004 (to be published).

[2] http://www.ivdgl.org/documents/document_server/uploaded_documents/doc--998--iVDGL_Star_monitoring.ppt

[3] http://indico.cern.ch/getFile.py/access?contribId=393&sessionId=7&resId=0&materialId=slides&confId=0

[4] http://ultralight.caltech.edu/web-site/ultralight/html/index.html