

High Performance Computing and Visualization

Research and Development in Visual Analysis

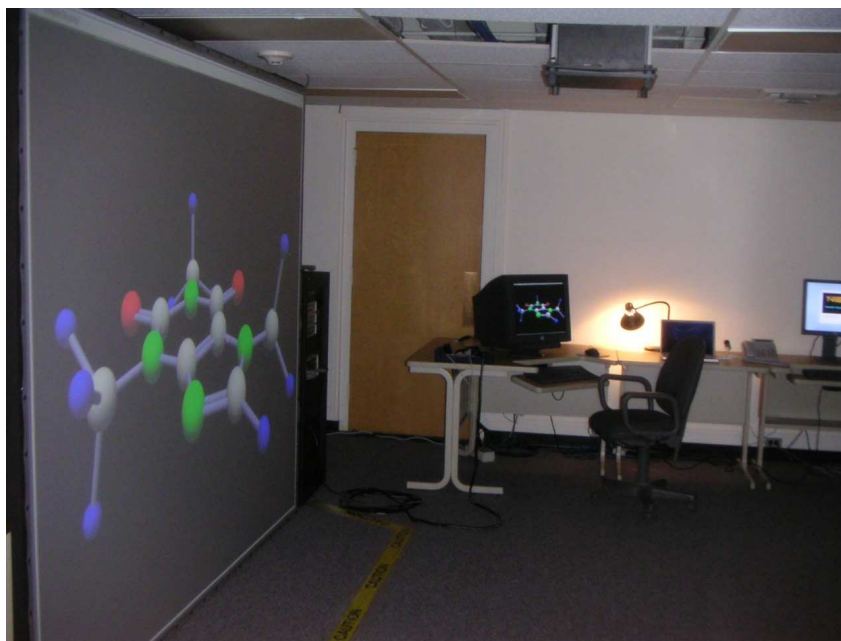
*Judith Devaney
Terrence Griffin
John Hagedorn
Howard Hung*

*John Kelso
Adele Peskin
Steve Satterfield*

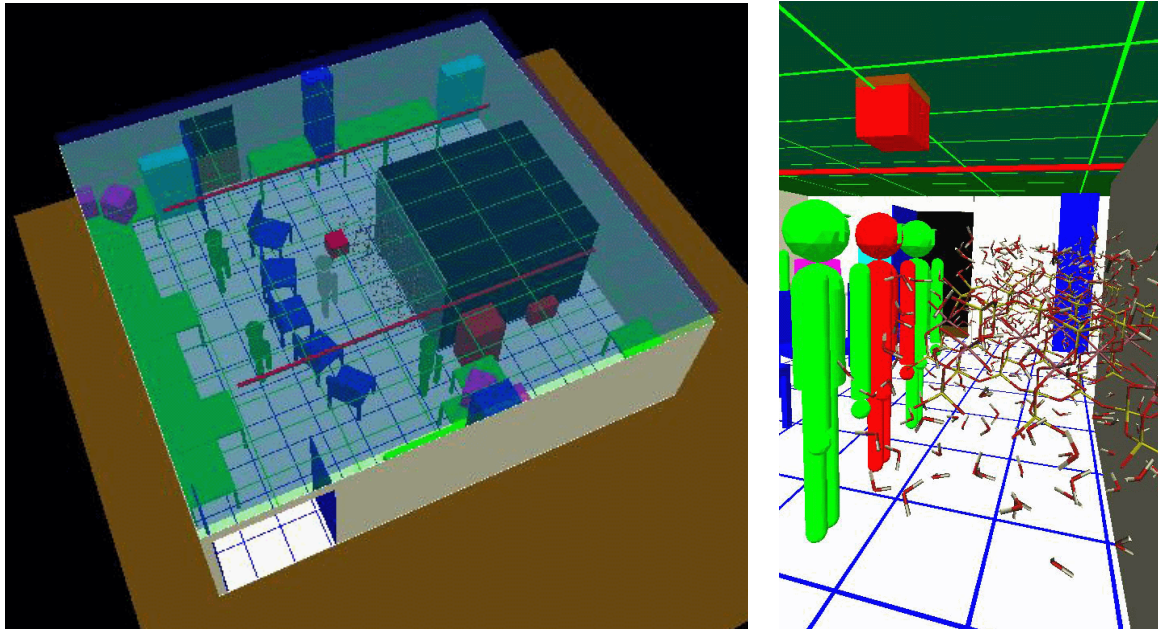
Computational and laboratory experiments are generating increasing amounts of scientific data. Often, the complexity of the data makes it difficult to devise *a priori* methods for its analysis, or the data is from new landscapes, such as the nano-world, where we have little experience. Moreover, there may be ancillary data, from databases for example, that would be helpful to have available. We are developing visual analysis capabilities in an immersive environment that allow NIST scientists to interact with data objects in a three-dimensional landscape rather than simply viewing pictures of them. With visual exploration, scientists can easily perceive complex relationships in their data, quickly ascertaining whether the results match expectations. This system functions as a unique scientific instrument.

Immersive Visualization Environment

The center of our visual analysis system is an immersive visualization environment. This environment is distributed and has software components and hardware components. A major goal of the project is to keep up with advances in hardware while insulating the scientists from these changes. The hardware for visualization is moving towards commodity cards and processors. We track these changes and update our environment at a pace that does not interfere with our ability to perform visual analysis on scientific data. A major accomplishment this year was the installation of an immersive environment at NIST Boulder, driven by an SGI with commodity graphics cards.



Boulder Immersive Environment



Top view (left) and side view (right) of mock up of Boulder immersive visualization facility.

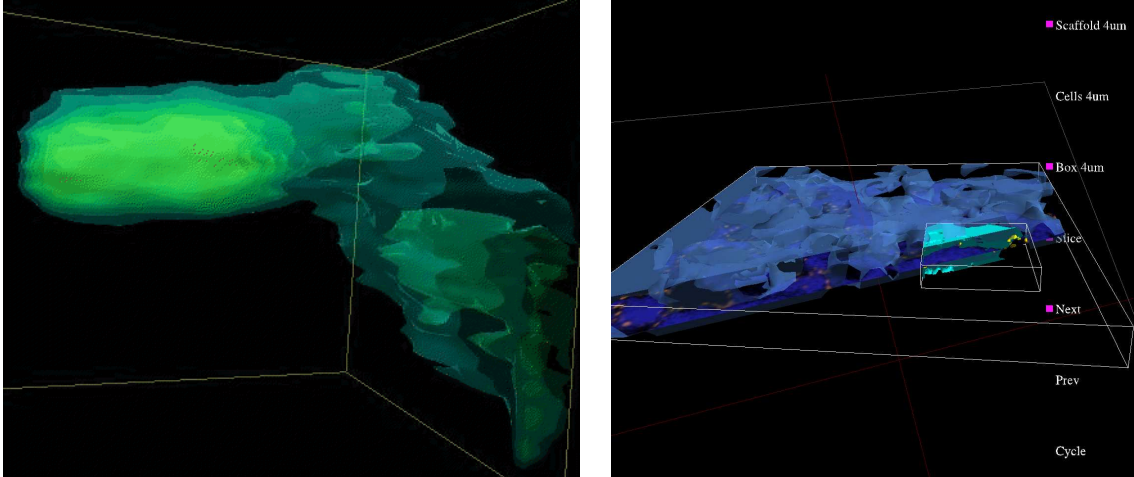
Another goal is to allow scientists to preview their data on their desktop with a Linux machine. These machines would be kept in sync with our primary environment. This year we have placed one machine on the desk of Nick Martys (BFRL) in support of the VCCTL project.

Visual Analysis Software (<http://math.nist.gov/mcsd/savg/software>)

Our software includes methods to get data into the immersive environment, methods to represent the data, and methods to interact with the data and the underlying computational environment. We analyze both the output of computational experiments as well as laboratory experiments. Following are some highlights of our accomplishments this year.

SAVG FileFormat. We have defined a file format and written an immersive visualization loader to input results of specific analyses we create directly, as well as to facilitate displaying of results we create with specialized packages. This file format is ASCII based. It describes geometric primitives including tri-strips, polygons, line segments, points, color, transparency, surface normals, and other rendering properties. The input file format is extensible and application independent, and will allow many applications to easily generate efficiently displayable data. We have written an OpenDX to SAVG converter and Kitware has written a program to output this format from VTK, their visualization toolkit. Kitware will include this in their next release. We have also written a toolbox that creates objects in this format and that contains tools for coloring, scaling, translating these objects. We have released the format and loader.

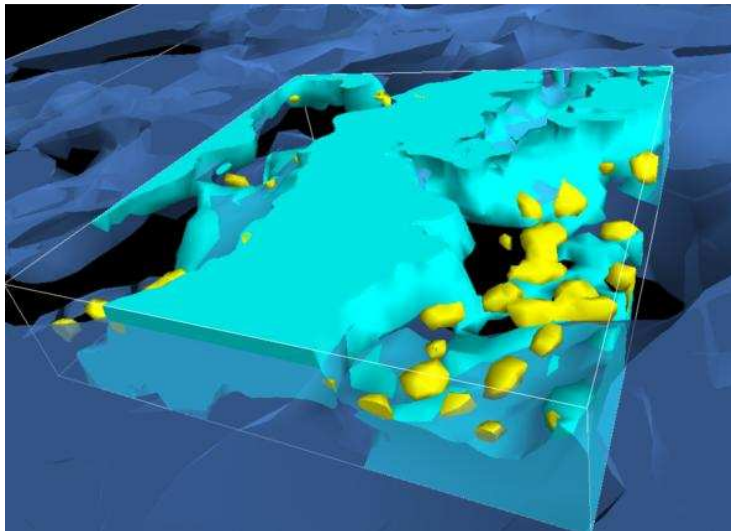
Transparency. We have created a nested transparent surface viewer for viewing of dense volumetric data in the RAVE. This viewer layers the colors for maximum view depth. The SAVG loader was written to display transparency with arbitrary nestings. The visualization of transparency is complex because what one sees depends on the sequence of transparent objects that intersect with the viewer's line of site. This sequence must be combined correctly to show the correct color to the viewer. There are efficiency issues as well. See figure below for an application.



Left: Bayesian reconstructions of a circuit with an electromigration void showing nested transparent isosurfaces. *Right:* A portion of the immersive display showing several features of the system (interactive menus, 2D cross-sections, interactive clipping plane, and transparency).

Clipping Planes. Many datasets are three dimensional dense datasets. Being able to see into them easily is a priority. We have created a set of interactive clipping planes that enable the user to prune away any data between them and the plane. This can also be a clipping box where all but the data in the box is made invisible. At the extreme of a very narrow box, the result is a three dimensional contour plane. The figure above (right) is an example with the tissue engineering data set.

Segmentation and Registration. We have incorporated the capability to perform segmentation and registration on data to derive additional representations. For example, we used level set segmentation to derive surfaces for the tissue engineering project. See the figure below.



Polymer and cell surfaces derived from segmentation of the tissue engineering project data using level sets. The blue is the polymer scaffold with the darker blue scanned at a higher resolution than the light blue. The yellow depicts the cells. The primary laboratory data is noisy, but the level set segmentation method produces a scientifically meaningful segmentation.

Navigation Through Differing Scales. In an immersive environment, the size of the objects in the virtual space relative to the scientist interacting with the objects becomes significant. For example, in the tissue engineering and cement fibers datasets, when data from both scales of the data are displayed simultaneously, some of the data was either too small to see properly or some was too large to see properly. The most important issue was the navigation through the differing scales of the data. We solved this by implementing a method for interactively changing the size of the data in the immersive environment in a continuous fashion. This enables the user to make small features large for detailed inspection or to make large volumes smaller to enable overviews of the data.

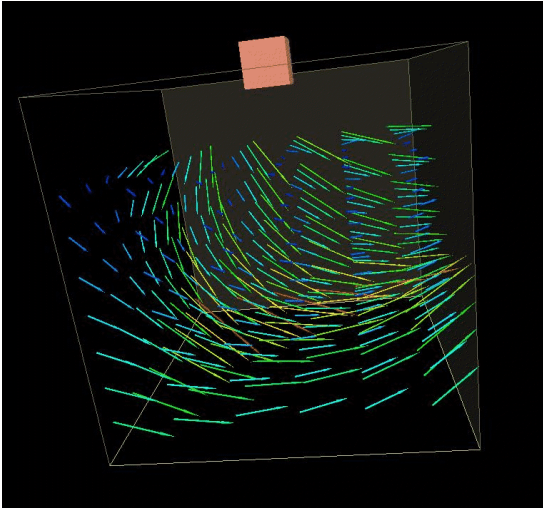
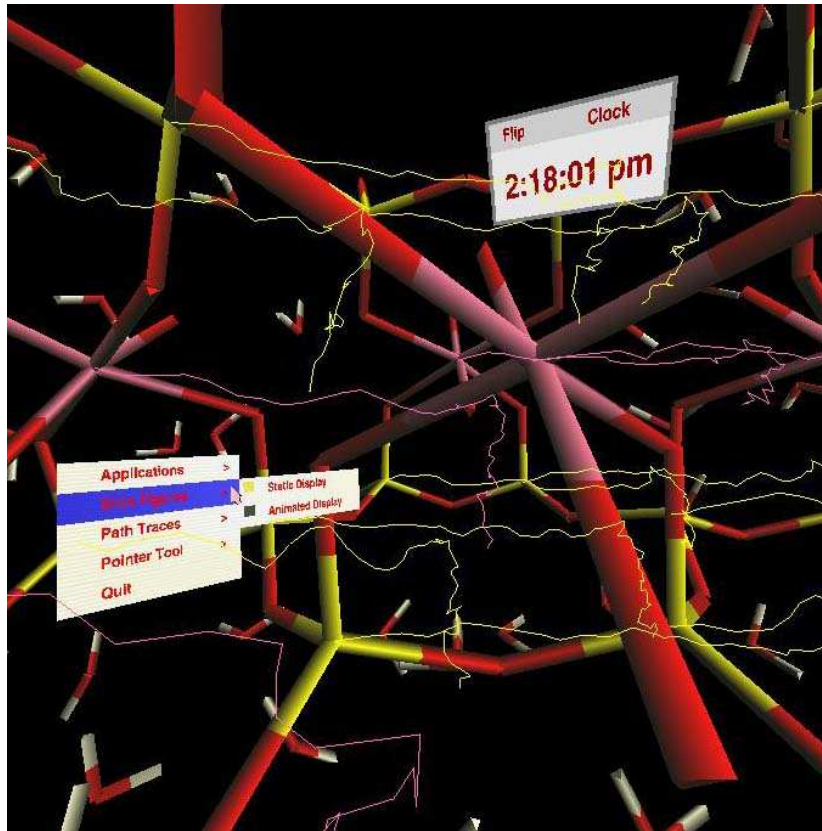


Image from tracker calibration project. Arrows show magnitude and direction of errors.

sampled from the space. This algorithm enables the interpolation of a single-valued function on points in three dimensional spaces. Three separate interpolations were needed to derive corrected values for the X, Y, and Z coordinates.

Technology Insertion. In order to spur commercial development of immersive visualization tools, we developed a solicitation entitled *Device Independent Interaction Framework for Immersive Scientific Visualization* for the FY 2004 NIST Small Business Innovative Research (SBIR) program. A company named Open Tech of Blacksburg, VA submitted a winning proposal. They will be developing tools that enable scientists to select computer hardware devices and interaction techniques that are most appropriate to their investigation, without necessitating modifications to their software to manage the particular devices, platform, and configuration selected. We have received a first delivery of menuing software for the immersive environment from Open Tech. Their interface works on the desktop as well as in the immersive environment. The windows are centered on the user's head and provide a natural method of interaction that includes response to gestures. The figure below provides an example.

Tracker Calibration. Virtual reality environments need a method of referencing where a user is positioned, so that the images projected are centered on the viewer's perspective. This is accomplished by using a tracking system. The electromagnetic tracking system is the most widely used today, but it is not without flaws. The idea behind tracking is very simple. Through the use of sensors, the position and orientation of a user is sent to the tracking mechanism. The problem is that when the coordinates are transmitted to the receiver, they are not precisely accurate, resulting in a distorted image projected to the user. This distortion is due to the sensitivity of the electromagnetic tracking system to metal objects. By calibrating the tracking system, the distortion is compensated by assigning a known value to offset the abnormality in transmission. The quadratic Shepard method for trivariate interpolation of scattered data was used on points



This figure demonstrates the OpenTech VEWL menuing system based on the OpenGL Polygon Based menu for the smart gel data.

Virtual Cement and Concrete Testing Laboratory (VCCTL)

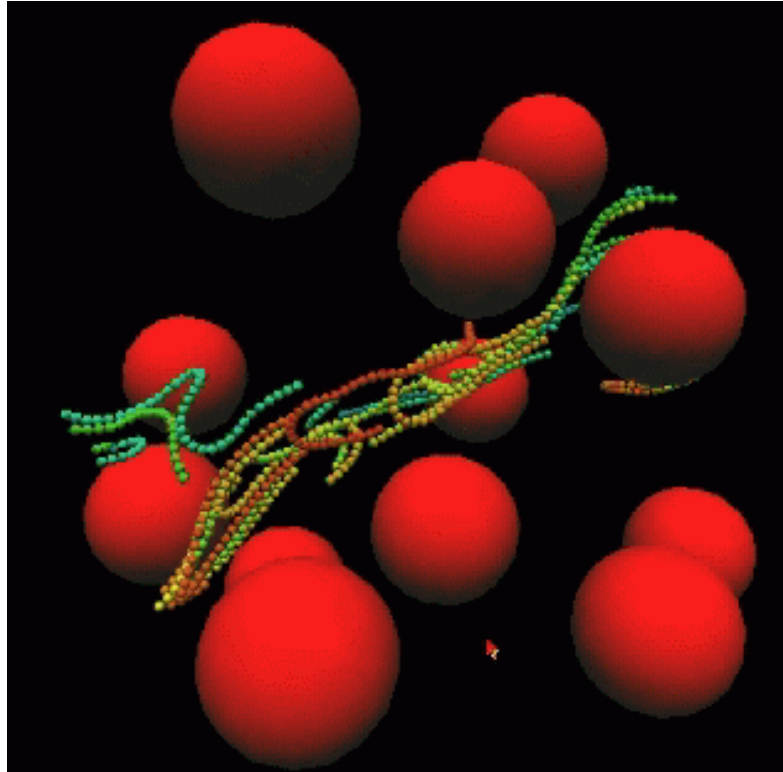
*Judith Devaney
William George
Terence Griffin
John Hagedorn
Howard Hung*

*Steve Satterfield
James Sims
Clarissa Ferraris (NIST BFRL)
Edward Garboczi (NIST BFRL)
Nicos Martys (NIST BFRL)*

The NIST Building and Fire Research Laboratory (BFRL) does experimental and computational research in cement and concrete. Recently, MCS D has been collaborating with BFRL in the parallelization of their codes and in creating visualizations of their data. In January 2001 the Virtual Cement and Concrete Testing Laboratory (VCCTL) Consortium was formed. MCS D assisted in this effort through presentations of our work with BFRL and demonstrations of visualizations in our immersive environment. The consortium consists of NIST (BFRL and ITL) and nine industrial members: Cemex Trademarks Worldwide, Ltd., Holcim (US) Inc., Master Builders Technologies, National Ready Mixed Concrete Association, Association Technique l'Industrie des Liant Hydrauliques (ATILH), International Center for Aggregate Research (ICAR), W.R. Grace, Sika Technology AG, and Portland Cement Association. The overall goals of the consortium are to develop a virtual testing system to reduce the amount of physical concrete testing, and to expedite the

research and development process. This will result in substantial time and cost savings to the concrete construction industry as a whole. MCS D continues to contribute to the VCCTL through collaborative projects involving parallelizing and running codes, creating visualizations, as well as presentations to the VCCTL current and prospective members. For example, in November 2003, J. Devaney made a presentation to the VCCTL entitled *Cement and Concrete: Parallel Computing, Scientific and Information Visualization*, and in November 2004 a presentation entitled *Automatic Labeling with Machine Learning*. The following projects are included in this effort.

This work is supported in part by the Virtual Cement and Concrete Testing Laboratory Consortium.



S. Satterfield worked with N. Martys (NIST BFRL) to create an immersive visualization showing the results of a dynamic simulation of concrete flow with fibers.

Computational Modeling of the Flow of Concrete. Understanding the flow properties of complex fluids like suspensions (e.g., colloids, ceramic slurries and concrete) is of technological importance and presents a significant theoretical challenge. The computational modeling of such systems is also a great challenge because it is difficult to track boundaries between different fluid/fluid and fluid/solid phases. We use a new computational method called dissipative particle dynamics (DPD), which has several advantages over traditional computational dynamics methods while naturally accommodating such boundary conditions. In DPD, the interparticle interactions are chosen to allow for much larger time steps so that physical behavior, on time scales many orders of magnitude greater than that possible with molecular dynamics, may be studied. Our algorithm (QDPD) is a modification of DPD which uses a velocity Verlet algorithm to update the positions of both the free particles and the solid inclusion. In addition, the rigid body motion is determined from the quaternion-based scheme of

Omelayan (hence the Q in QDPD). Parallelization of the algorithm is important in order to adequately model size distributions, and to have enough resolution to avoid finite size effects.

This year W. George, in collaboration with N. Martys of BFRL, completed a major modification to the QDPD application, consisting of re-write of the inner-most computational core to add a second level of parallelization on top of the higher-level MPI based message-passing parallelism. A simplification to the MPI level of parallelism is also being considered as a result of this modification. This effort is intended to expand the capabilities of this application in order to scale it from the current 10s of processors up to an application that will run on 1000s of processors. This will allow for the simulation of much larger systems as well as the inclusion of additional physics to the computation. This second level of parallelism has added multi-threading to the main computation in the inner-most loops of the algorithm. This change has been made to take better advantage of the hybrid architecture of current parallel machines, such as the IBM-SP, that consist of 10s to 100s of SMP nodes each of which consists of 2-16 CPUs.

For initial development and testing, J. Devaney and W. George have obtained user accounts at NERSC (National Energy Research Scientific Computing Center, Lawrence Berkeley National Laboratory), and have obtained a small allocation of 2000 CPU hours on the NERSC parallel machine Seaborg, a 6080 CPU IBM SP RS/6000 (380 16-CPU nodes). This account has been used to further develop this QDPD application and to compare the performance of the new multi-level parallel version with the pure MPI version.

In addition to the work on the IBM SP, we have performed a series of benchmark tests of the new multi-level parallel QDPD application on an 8-CPU SGI. The results so far, after some code tuning, show no performance improvement of the multi-level version of QDPD over the pure MPI version given the same number of CPUs. This is true regardless of the ratio of MPI processes to threads used. In all cases the pure MPI version of QDPD outperforms the multi-level parallel version. If similar results are obtained on the distributed memory IBM SP this direction of algorithm development will be changed.

Immersive Visualization of Concrete Aggregate

Flow. In support of the VCCTL, X-ray tomography has been used to create a database of aggregates spanning about four decades in size. Examples include cement particles, sand and, even, rocks. These realistic aggregate shapes can be incorporated into codes used to model the rheological properties of cement based materials. The purpose of this project is to develop techniques to display flows of multiple types of aggregate in an immersive visualization environment.

T. Griffin is developing a visualization of aggregate directly from the database to insure that the visualization and the codes reference the same aggregate particles. This will run directly on a Linux laptop or in the immersive environment. While the visualization references the same aggregate, it can be shown with differing levels of realism.

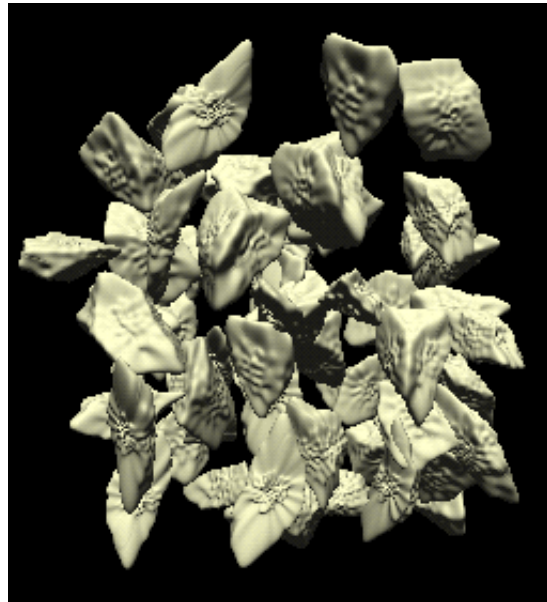


Image of modeled actual aggregate in a shear flow.

Parallelization of Fluid Flow in Complex Geometries. The flow of fluids in complex geometries plays an important role in many environmental and technological processes. Examples include oil recovery, the spread of hazardous wastes in soils, the processing of polymer blends, droplet breakup, phase separation and chemical analysis in confined geometries, and the service life of building materials. The latter two applications are of particular concern for our NIST collaborators in BFRL. The detailed simulation of such transport phenomena in varying geometries and subject to varying environmental conditions or saturation, is a great challenge because of the difficulty of modeling fluid flow in random pore geometries and the proper accounting of the interfacial boundary conditions. In order to model realistic systems, BFRL has developed a lattice Boltzmann (LB) algorithm that simulates multiple fluids, various forces, and wetting characteristics within arbitrary geometries. We have parallelized this algorithm using MPI to enable the study of large systems.

This year the code was extensively exercised in a study leading to the following paper: J. Hagedorn, N. Martys, and J. Douglas, "Breakup of a Fluid Thread in a Confined Geometry: Droplet-Plug Transition, Perturbation Sensitivity, and Kinetic Stabilization with Confinement," *Physical Review E* **69** (5) (2004). J. Hagedorn is currently collaborating with N. Martys on modeling the permeability of two 3D images of paper provided by the Avery Dennison Corporation.

Visualization of Smart Gels

Steve Satterfield

Carlos Gonzalez (NIST CSTL)

<http://math.nist.gov/mcsd/savg/vis/gel/>

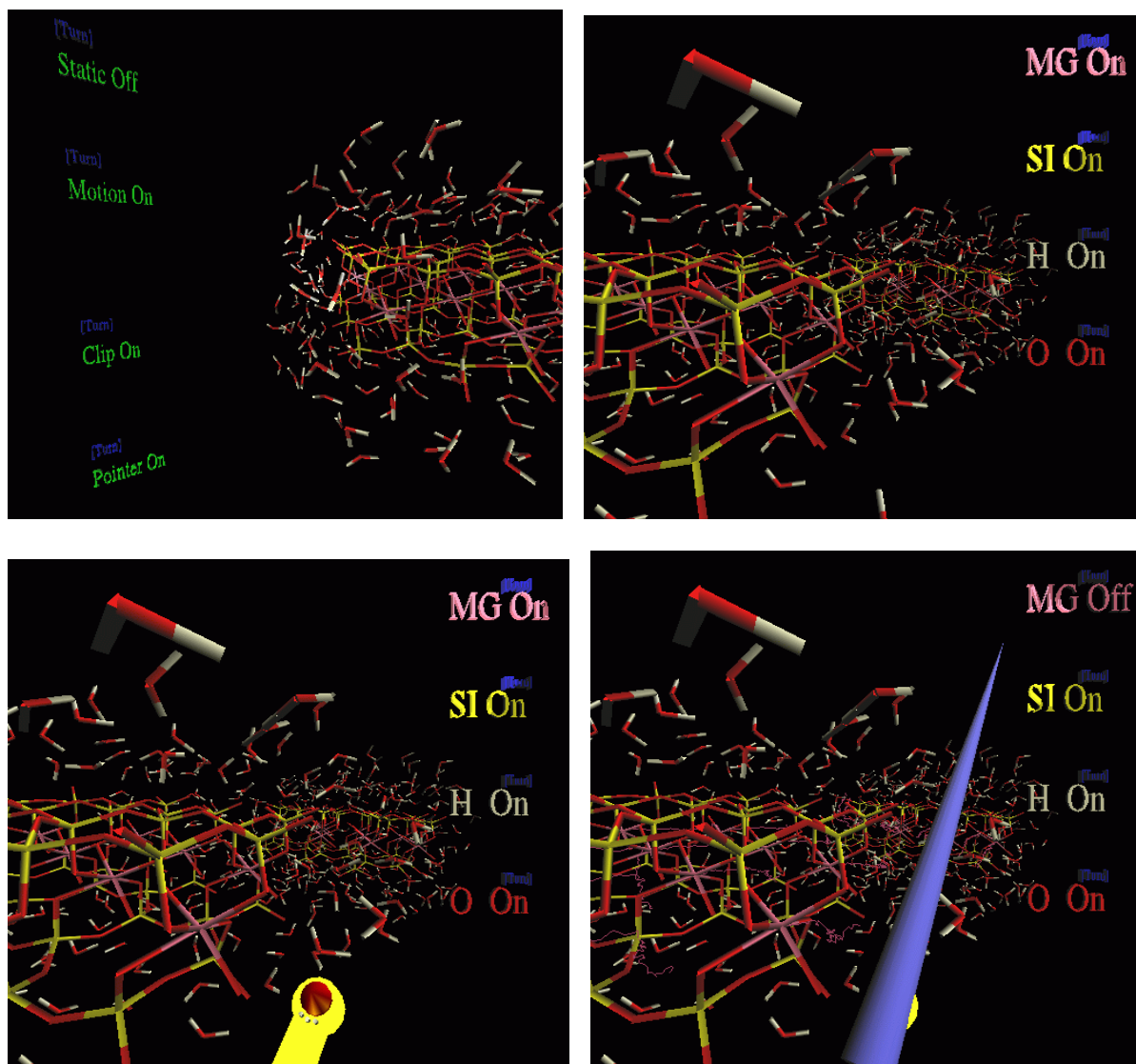
A smart gel is a material that gels in response to a specific physical property. For example, it may gel at a specific temperature or pressure. The mechanisms that create a gel in response to given stimuli are not well understood. Developing this understanding is the key to being able to create materials that gel under precise control. The potential for applications of smart gels is enormous. For example, they could be useful in applications such as an artificial pancreas that releases insulin inside the body in response to high sugar level. Smart gels might someday be used to make exotic foods, cosmetics, medicines, and sensors.

The NIST team is studying a subclass of these materials called shake gels. Through some complex and as yet unknown process, these watery mixtures of clays and polymers firm up into gels when shaken, and then relax again to the liquid phase after some time has passed. A shake gel might be used, for example, in shock absorbers for cars. The material would generally be a liquid but would form a gel when the car drove over a pothole; the gel thickness would adjust automatically to the weight of the car and the size of the pothole. A more esoteric application might be the formation of gelled areas within a liquid where holograms could be created using a laser.

Working with our collaborators in CSTL, we have created a dynamic immersive visualization of the results of a molecular level computational simulation of shake gel formation. This visualization provided NIST scientists with unprecedented insight into the processes controlling the formation of shake gels. This work received wide notoriety during the last year, including the following, each of which described Smart Gel work of Carlos Gonzalez and immersive visualization of S. Satterfield. Film reports were done in the MCS D RAVE immersive visualization lab.

- Segment in opening video of the SC 2003 Conference, November, 2003, Phoenix, AZ.
- News report on local Fox Network affiliate (Channel 5) on October 30, 2003
- *Design News* article in their October 20, 2003 issue.

- A segment of the science series NEXT@CNN that aired on CNN on Saturday January 3. J. Devaney and S. Satterfield were both interviewed. An associated article was posted on the CNN site³.
- *HPCWIRE* article (106741) in their January 9, 2004 issue.



Four images of the shake gel visualization illustrating menu controls and pointing devices which have been developed by MCSD to enable interactive exploration of the data.

³ <http://www.cnn.com/2004/TECH/science/01/02/coolsc.visualization/index.html>

3D Chemical Imaging at the Nanoscale

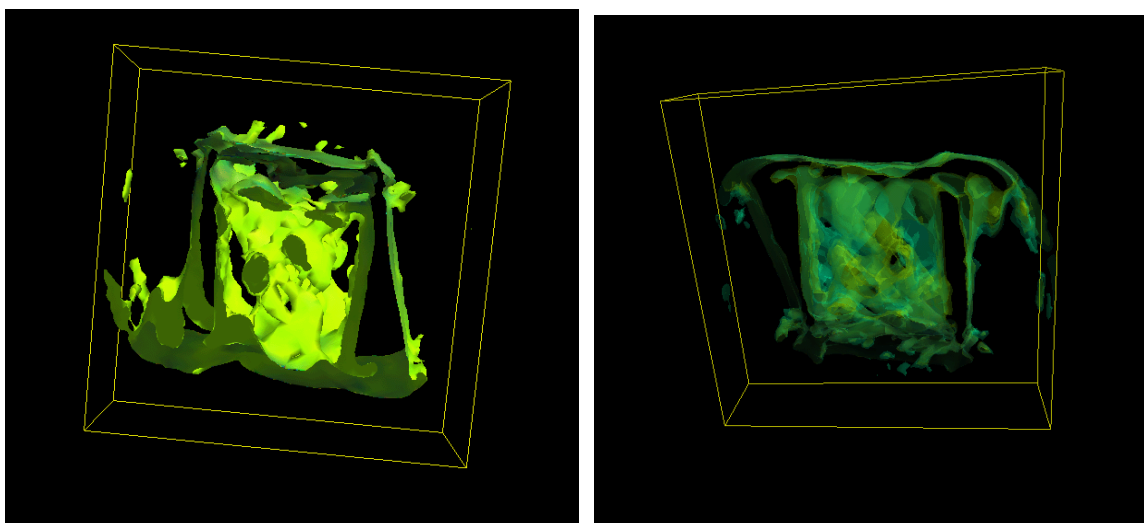
*William George
Steve Satterfield
John Hagedorn
John Kelso
Adele Peskin*

*Judith Devaney
Eric Steel (CSTL)
John Henry Scott (CSTL)
John Bonevich (MSEL)
Zachary Levine (PL)*

A quantitative understanding of the distribution of chemical species in three dimensions including the internal structure, interfaces and surfaces of micro- and nanoscale systems is critical to the development of successful commercial products in nanotechnology. Current nanoscale-chemical 3D measurement tools are in their infancy and must overcome critical measurement barriers to be practical. This project is developing intermediate voltage electron microscope measurement approaches to attain three-dimensional chemical images at nanoscale-resolution. These approaches will be broadly applicable to nanoscale technologies from microelectronics to pharmaceuticals and subcellular biomedical applications.

MCS D collaborators are working on data management and visual analysis techniques and tools to enable the analysis of imagery to be generated by this project. Among the particular capabilities under development are the following.

- Techniques for the interactive visualization of 3D isosurfaces in an immersive environment
- Segmentation techniques for 3D datasets
- Standardization of file formats



Visualizations of reconstructed tomographic data of a photonic crystal sample. This is a photonic band gap material, approximately 3 micrometers across. An artificially periodic crystal was created by exposing photoresist to four laser beams. The sample was sectioned with a focused ion beam, then a tilt series of 76 views was obtained using a transmission electron microscope. The views were aligned to five fiducial markers, then reconstructed using a Bayesian algorithm. The reconstruction was Fourier filtered. *Left:* the structure is represented as an opaque surface approximating the surface of the structure. *Right:* the structure is represented with two transparent surfaces in order to reveal additional internal detail in the data.

In addition, W. George has designed an architecture and API for a 3D Chemical Imaging service. This architecture is client/server based using Jini (a Java based distributed computing infrastructure) to link a client application on the user's workstation to a compute server on a parallel machine. A parallel Monte Carlo Simulator, probably in Fortran/MPI, will be the main computation. Other applications will be needed to prepare the input and post-process the output from this simulator. The computation will run on an MCSD multiprocessor SGI system using semi-dedicated CPUs.

This project is sponsored by the NIST Competence Program.

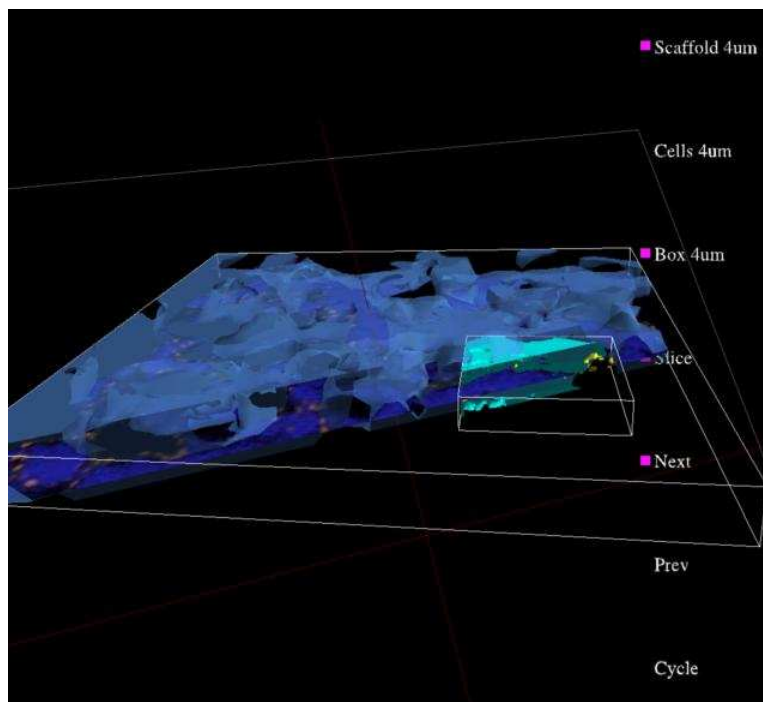
Multi-Modal Imaging and Visualization for Integrating Functional and Structural Information

*John Hagedorn
Steven Satterfield
John Kelso
Adele Peskin
Judith Devaney*

*Joy Dunkers (NIST MSEL)
Marcus Cicerone (NIST MSEL)
Lyle Levine (NIST MSEL)
Gabrielle Long (NIST MSEL)*

We can no longer advance material science simply by studying model systems that are idealized in dimension and function. We must comprehend realistic, complex, three-dimensional systems in terms of their structure, function, and dynamics over a broad scale from nanometers to millimeters. In this collaboration with the NIST Materials Science and Engineering Lab, which is sponsored by the NIST ATP program, we are combining data from different measurement devices that reflect both functional and structural information. The multiple data sets taken from a single sample are then combined and visualized in our interactive, immersive, virtual reality environment in order to gain new insights into the physics and materials science of complex systems.

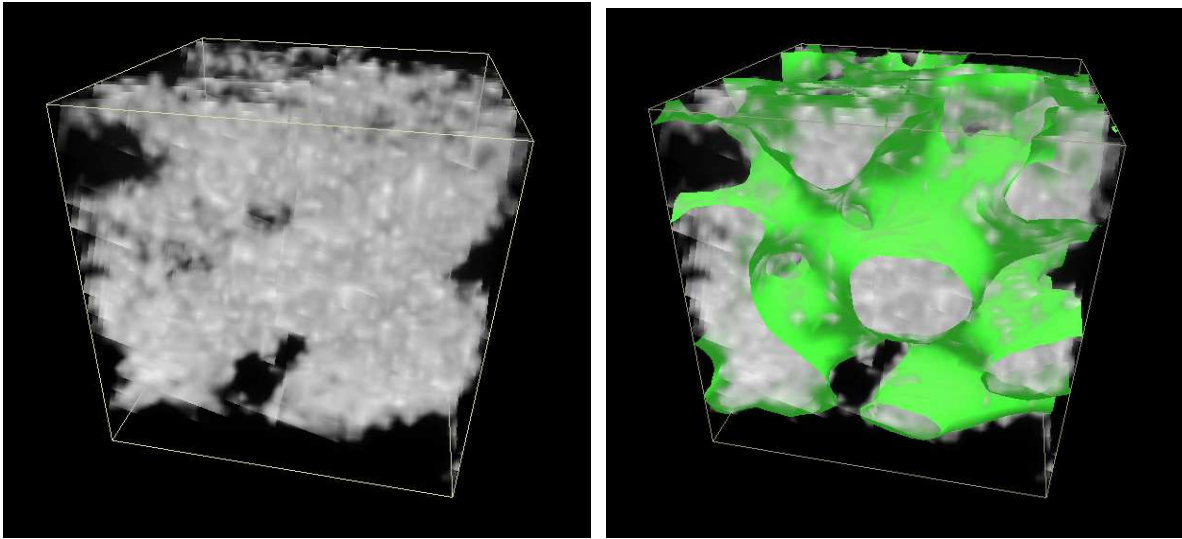
Our collaborators are gathering measured data using a variety of techniques, including optical coherence tomography (OCM) and confocal fluorescence (CFM) imaging. When data from these different devices is combined in a manner that is visually apparent, unprecedented insight towards the comprehension of complex relationships among large amounts of correlated data can be obtained. These methods have applications to the characterization of biomaterials, the failure analysis of polymer composites, and the reliability of semiconductor devices. In this project we are concentrating on an application in tissue engineering. The system being imaged is that of cells growing on polymer scaffoldings.



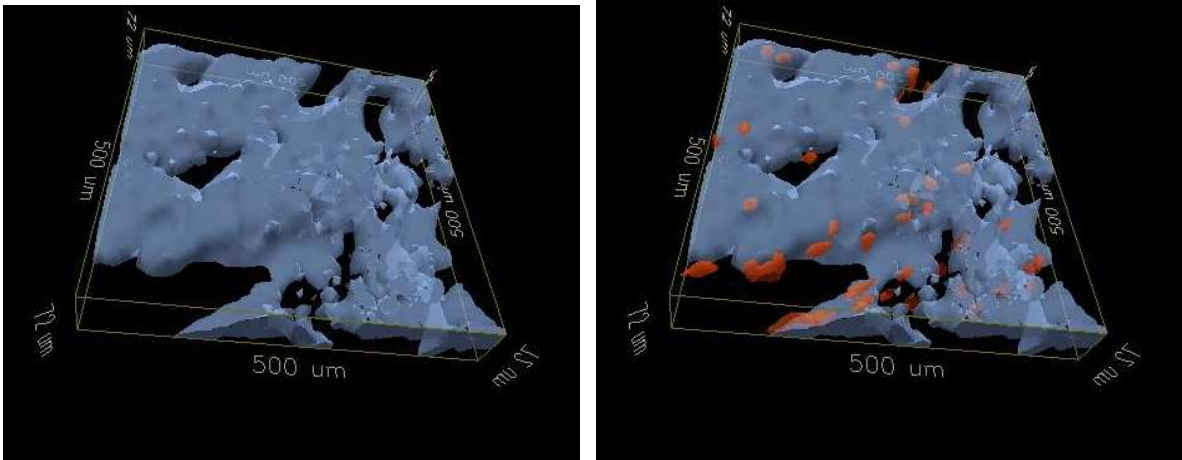
A portion of the immersive display showing several features of the system: interactive menus, 2D cross-sections, interactive clipping plane, and transparency.

Visual analysis of data of this type presents a variety of challenges. The data is three-dimensional. Separate noisy data sets from different instruments must be cleaned, combined, and registered. Data of different resolutions must be combined for use in a single visualization. Complex porous surfaces within the volume must be identified. Interactive volume-rendering techniques, surface-rendering techniques, and combined volume/surface-rendering techniques must be available. We have developed a flexible framework that supports such complex visual analysis of combined CFM and OCM data in our immersive environment. Our display makes use of transparency, lighting effects, simultaneous viewing of 2D and 3D data, annotation, clipping, and menus. The figures below illustrate a variety of these features.

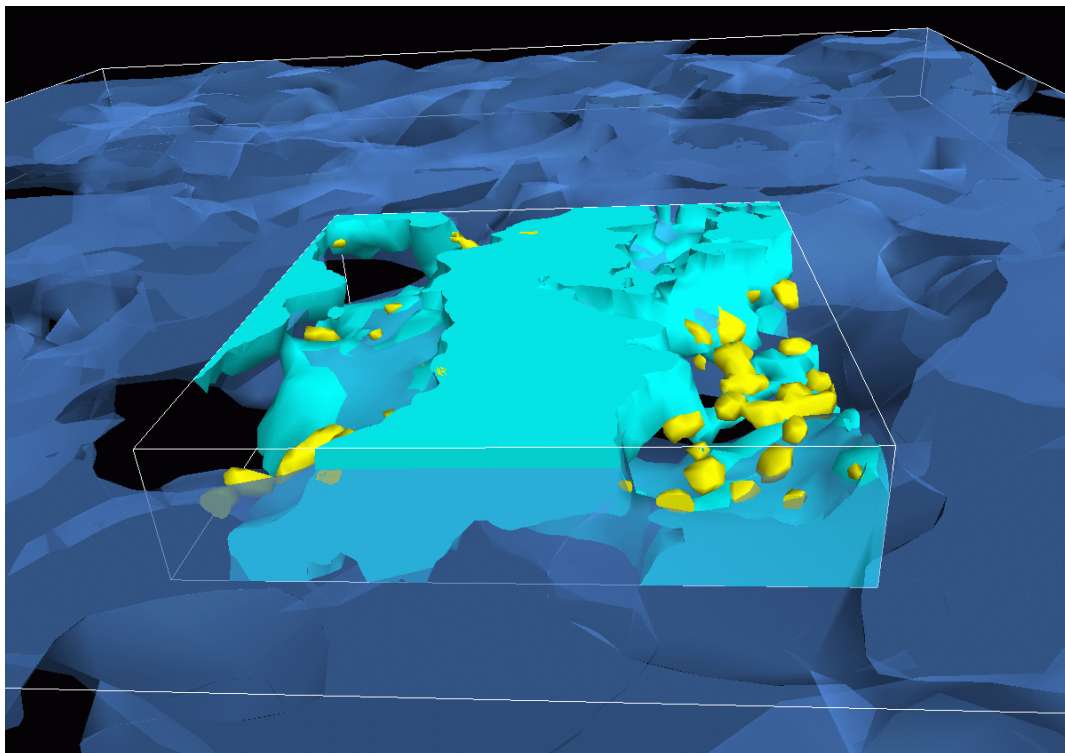
This project is sponsored by the NIST Advanced Technology Program.



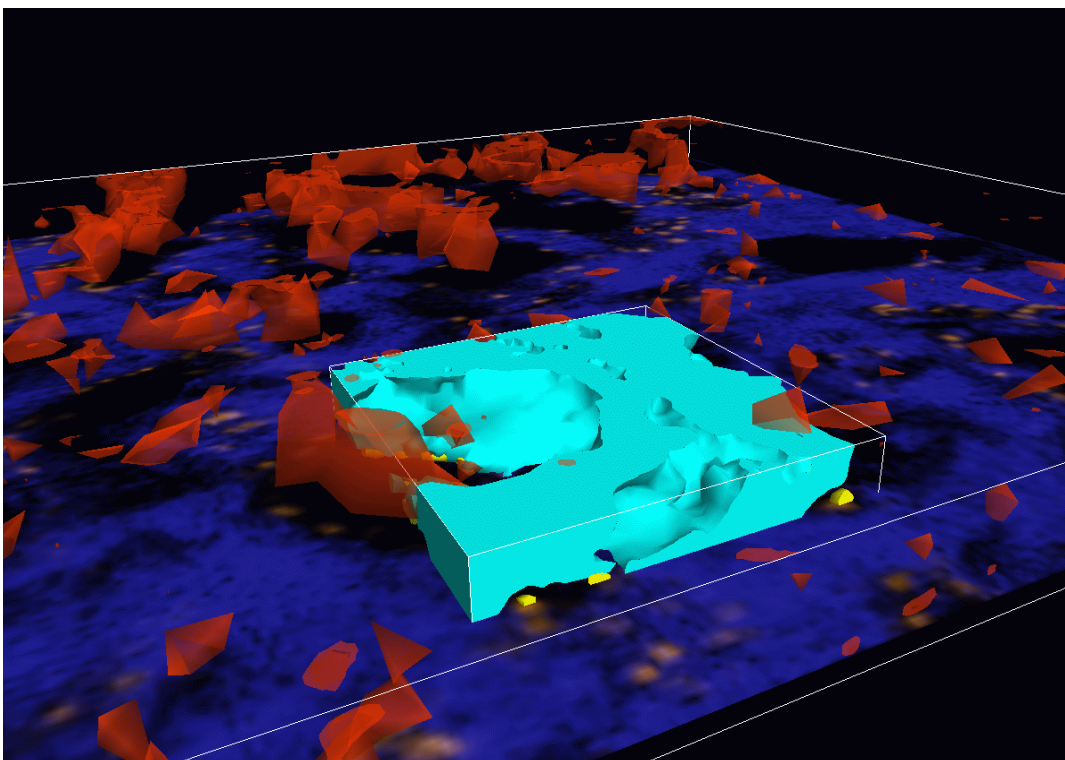
Left: Volume visualization of scaffolding. *Right:* A segmentation of the scaffolding from the pore space. Both are directly viewable in the RAVE.



Left: The surface of the scaffolding as a transparent surface. *Right:* Cells (in red) are shown growing on the scaffolding. This is a fusion of data from two separate instruments. It is directly viewable in the RAVE as scaffolding only, cells only, or scaffolding and cells.



High resolution (opaque blue) scan of polymer scaffolding with high resolution scan of cells (yellow) combined with lower resolution scan (transparent blue) of polymer scaffolding



High resolution (opaque blue) polymer scaffolding with cells (yellow) scans combined with cross section of lower resolution scans of polymer scaffolding (dark blue) and cells (red)

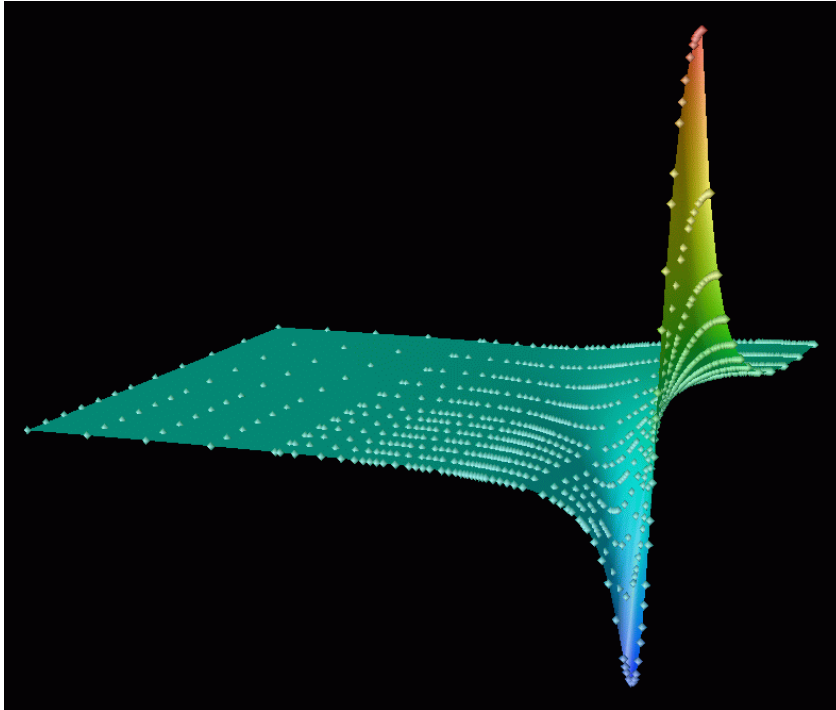
Physics Models for Transport in Compound Semiconductors

Howard Hung

Terrance Griffin

Herbert S. Bennett (NIST EEEL)

Physics models for carrier transport in semiconductors are essential inputs of computer programs that simulate the behavior of microelectronic and optoelectronic devices. Such simulations increase understanding, reduce times-to-market, and assist in making selections from among competing or alternative technologies. As devices shrink in size to nanometers, performing experimental measurements becomes more costly and time-consuming. This means that computer simulations will become more essential for advances in future nanotechnologies. Unlike many physics models that are based on using variations in parameters to fit experimental data, the NIST physics models developed in this project are based on quantum mechanical calculations with no fitting parameters to account for dopant ion effects and many-body physics effects.



This figure summarizes the computed theoretical calculations of the real part of the complex electric susceptibility for n-type GaAs with a dopant density of 10^{18} cm^{-3} and at a temperature of 300K.

The calculations include many body quantum effects and bandgap narrowing due to dopant ion carrier interactions. The many body quantum effects treat both electron-electron and electron-hole interactions. The results are unique because all other reported treatments for the electric susceptibility 1) do not treat these effects self-consistently; 2) are Taylor series expansions in either (Q/A) or (A/Q) , where Q is the magnitude of the normalized wave vector and A is the normalized frequency used in such measurement methods as Raman spectroscopy; and 3) do not give the structure shown in the figure below. These results will change

the way researchers and process engineers interpret non-destructive measurements to extract the carrier concentrations of GaAs wafers. The wafer carrier concentration is a key figure of merit associated with a go-no-go decision for determining whether a wafer meets specifications and should undergo further processing.

We are collaborating with EEEL to develop efficient computational methods for this problem, and to develop evocative displays of the results in our immersive environment. In particular, H. Hung has worked with H. Bennett of EEEL to develop new Fortran subroutines to evaluate Lindhard electric susceptibility integrals with singular integrands. They calculated the electric susceptibility for GaAs as a function of frequency for given reciprocal wave vectors at 300K

in n-type GaAs with a dopant density of 10^{18} cm^{-3} . These results will be used by CSTL to develop an algorithm for non-destructively determining the carrier density of GaAs wafers from Raman measurements. Displays of the results were developed using Dataplot, IDL, and OpenDX. T. Griffin created an OpenDX viewer for the 3D output which was then converted to a format readable by DIVERSE and the results viewed in the RAVE immersive environment.

Computation and Visualization of Nano-structures and Nano-optics

James Sims

John Hagedorn

Howard Hung

John Kelso

Steve Satterfield

Adele Peskin

Garnett Bryant (NIST PL)

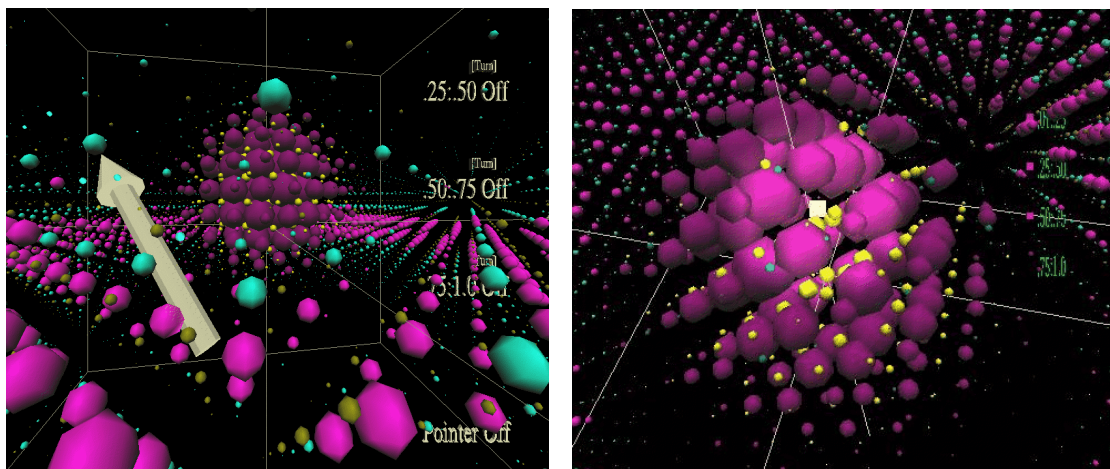
<http://math.nist.gov/mcsd/sav/parallel/nano/>

<http://math.nist.gov/mcsd/sav/vis/nano/>

Accurate atomic-scale quantum theory of nanostructures and nanosystems fabricated from nanostructures enables precision metrology of these nanosystems and provides the predictive precision modeling tools needed for engineering these systems for applications including advanced semiconductor lasers and detectors, single photon sources and detectors, biosensors, and nanoarchitectures for quantum coherent technologies such as quantum computing. Theory and modeling of nanoscale and near-field optics is essential for the realization and exploitation of nanoscale resolution in near-field optical microscopy and for the development of nanotechnologies that utilize optics on the size-scale of the system. Applications include quantum dot arrays and quantum computers. Atomic-scale theory and modeling of quantum nanostructures, including quantum dots, quantum wires, quantum-dot arrays, biomolecules, and molecular electronics, is necessary to understand the electronic and optical properties of quantum nanostructures and nanosystems fabricated from component nanostructures.

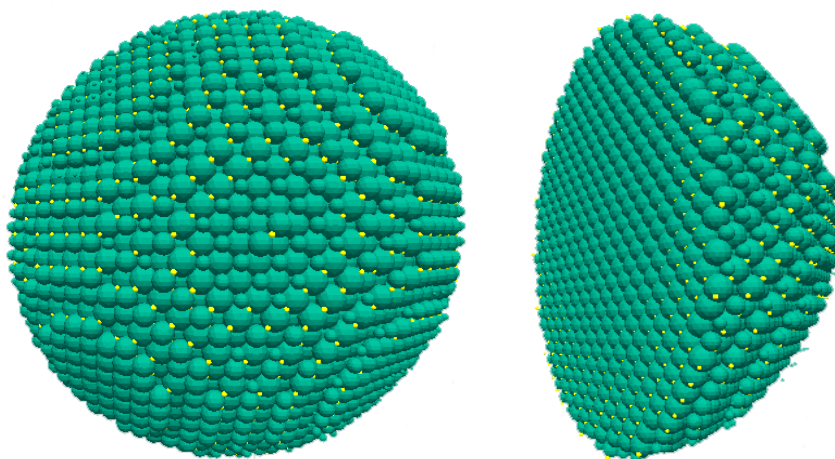
We are working with the NIST Physics Lab to develop computationally efficient large-scale simulations of such nanostructures. We are also working to develop immersive visualization techniques and tools to enable analysis of highly complex computational results. Among the accomplishments of the past year are the following.

- J. Sims completed and benchmarked a parallelization of the simulation code across one dimension. The code shows a speedup of 27 on 41 processors of our Linux cluster. On a 96 processor SGI it shows a speedup about twice as fast as a 32 processor run, so the code is scaling well up to the limit of the 1D treatment (around 60+ processors).
- J. Sims also completed a second parallelization effort that includes spin-orbit interaction, which introduces complex arithmetic and hence is even more in need of parallelization. The complex version of PARPACK is being used in this effort (some bugs in PARPACK were discovered in the process). J. Sims and H. Hung will be co-authors with G. Bryant on a talk on this to be given at the March American Physical Society meeting. This work required parallelization for size as well as speed, with matrices spread among processors. The latest runs are diagonalizations of 3.7 million by 3.7 million matrices.
- H. Hung, S. Satterfield, J. Hagedorn, J. Kelso, and A. Peskin and have developed visualizations of the atomic structure of lattices of electrons. The display of s-electron orbitals has been made. Data sets with of up to 377,777 atoms have been successfully processed and displayed in the RAVE.



Left: Visualization of s-orbitals of a pyramidal structure. *Right:* visualization of a quantum dot.

- J. Kelso investigated how to most efficiently display large numbers of these objects, including techniques such as billboarding, small-feature culling, level of detail and image based rendering.
- J. Sims has begun work on the next major step the computational part of the project, i.e., calculations on arrays of nanoparticles. The basic idea is to consider each nanoparticle as part of its own cluster, using the same input data as now, but as the computation proceeds information from neighboring atoms in each cluster has to be distributed to the appropriate processor in neighboring clusters, thereby “stitching” the calculations on the clusters in the array together. This involves using some of the communication facilities of the previous code, but also some new ones. For the new problems presented by stitching together structures, Sims is using a set of irregular communication routines that are available from Steve Plimpton (part of the Zoltan package).



These visualizations of HgS S-orbitals showed a state that had never been seen before: “I’m astonished that it is so constant throughout. I have never seen a state like this before.” -- Garnett Bryant, PL

Computation of Atomic Properties with the Hy-CI Method

James Sims

Stanley Hagstrom (Indiana U.)

<http://math.nist.gov/mcsd/savg/parallel/atomic/>

Impressive advances have been made in the study of atomic structure, at both the experimental and theoretical levels. For atomic hydrogen and other equivalent two-body systems, exact analytical solutions to the nonrelativistic Schrödinger equation are known. It is now possible to calculate essentially exact nonrelativistic energies for helium (He) and other three-body (two-electron) systems as well. Even for properties other than the nonrelativistic energy, the precision of the calculation has been referred to as “essentially exact for all practical purposes”, i.e., the precision goes well beyond what can be achieved experimentally. These high-precision results for atomic two-electron systems have been produced using wave functions that include interelectronic coordinates, a trademark of the classic Hylleraas (Hy) calculations done in the 1920s. The challenge for computational scientists is to extend the phenomenal He accomplishments (the ability to compute, from first principles alone, any property of any two electron atom or its ion to arbitrary accuracy) to molecules and to atomic systems with three, four, and even more electrons. Where three electron atomic systems (i.e., lithium (Li) and other members of its isoelectronic series) have been treated essentially as accurately as He-like systems, demand on computer resources has increased by 6,000 fold. Because of these computational difficulties, already in the four-electron case (i.e., beryllium (Be) and other members of its isoelectronic series) there are no calculations of the ground or excited states with an error of less than 10 microhartrees (0.00001 a.u.). This is where a technique developed by Sims and Hagstrom in a series of papers from 1971 to 1976 becomes important. They developed the Hy-CI method, which includes interelectronic coordinates in the wave function to mimic the high precision of Hy methods, but also includes configurational terms that are the trademark of the conventional Configuration-Interaction (CI) methods employed in calculating energies for many-electron atomic (and molecular) systems. Because of this, the Hy-CI method has been called a hybrid method. This is the power of the method, because the use of configurations wherever possible leads to less difficult integrals than in a purely Hy method, and if one restricts the wave function to at most a single interelectronic coordinate to the first power, then the most difficult integrals are already dealt with at the four electron level and the calculation retains the precision of Hy techniques, but is greatly simplified.

Two papers on this work appeared this year.

- J. S. Sims and S. A. Hagstrom, Erratum: Comment on “Analytic Value of the Atomic Three-electron Correlation Integral with Slater Wave Functions,” *Physics Review A* **68** (2003), p. 059903.
- J. S. Sims and S. A. Hagstrom, “Math and Computational Science Issues in High-Precision Hy-CI Calculations I. Three-electron Integrals,” *Journal of Physics B: At. Mol. Opt. Phys.* **37** (7) (2004), pp. 1519-1540.

In the second paper, Sims and Hagstrom discuss changes they have made to their Hylleraas-Configuration Interaction (Hy-CI) methodology to most effectively use modern day computers to increase the size (number of terms) and accuracy of the calculations. The availability of cheap CPUs, which can be connected in parallel to significantly (orders of magnitude) enhance both the CPU power and the memory that can be brought to bear on the computational task, has made techniques that appeared hopeless only five years ago doable (assuming the linear dependence problem can be obviated with extended precision). The goal is to extend techniques which are known to give the most accurate upper bounds to energy states to four and more electrons. The first step in this process was to efficiently evaluate the only difficult integral arising when using the Hy-CI technique in the case of the number of electrons greater than or equal to three, the three-electron triangle integral. Sims and Hagstrom focus on recursive techniques at both the double precision and quadruple precision level of

accuracy while trying to minimize the use of higher precision arithmetic. Also, they investigate the use of series acceleration to overcome problems of slow convergence of certain integrals defined by infinite series. They find that a direct + tail Levin u-transformation convergence acceleration overcomes problems that arise when using other convergence acceleration techniques, and is the best method for overcoming the slow convergence of the triangle integral.

They are now at work on Papers II and III in this series, "Math and Computational Science Issues in High Precision Hy-CI Variational Calculations II. Four-electron integrals," and "Math and Computational Science Issues in High Precision Hy-CI Variational Calculations III. Nuclear attraction and kinetic energy integrals". The latter paper is coming out of a need effort to modify their codes to handle the more general case of non-spherically symmetric Slater-type orbitals (STOs). Once these papers are finished, they will be ready to tackle the difficult matrix assembly problem, and then to do a benchmark Be calculation. Progress to date has included finding a new Be radial limit (s-orbital CI, no r_{ij}) that is better than any published result. Also integrals are being calculated in blocks, each block independent of all others, so different blocks can be calculated on different processors in the final big runs.

In addition to atoms, work is underway to extend high precision electronic energy calculations to diatomic molecules. A calculation is underway on the hydrogen molecule, the two electron molecular analog of helium. Progress to date has been to calculate the ground state energy with an accuracy of nanohartrees (10^{-9}) which is better than all but one previous calculation, an explicitly correlated gaussian (ECG) calculation. The goal is to achieve accuracy on the order of 10^{-12} , worthy of publishing in the *Journal of Chemical Physics*. Sims has completed a parallelization of the code. Building the H and S matrix and the matrix diagonalization (solving) are both parallelized. We focus on diagonalizing the matrix since that dominates the calculation.

All results are for a 3485-term wave function with a total energy of -1.1744 7571 4018 7855 1344 4640 2249 9 a.u. We have achieved a speedup of 24.5 on 32 processors of our Linux cluster, a good scaling (76.4 % scalability), and a more favorable speedup of 30 on 32 processors (93.7 % scalability) on a 96 processor SGI. Furthermore that scaling holds well up through the 90 processor run, where we get a factor of 80 on 90 processors, an (89.1 % scalability). This parallelization is relevant not only to H_2 but atomic states (Be, etc.) as well.

Finally to verify that they can do a better H_2 calculation than ECG, they have done an H_2^+ calculation and get, with 63 terms,

No. terms	E (R=2.0)					
63	-0.6026	3421	4494	9464	1673	<i>Our result</i>
160	"	"	"	911		<i>ECG (Cencek and Kutzelnigg)</i>

So they have achieved energy with 18 digits accuracy, five more digits than ECG. So it looks like for H_2 as well as the other 2-electron diatomics that Hy-CI will beat ECGs in the long run.

In order to achieve this level of accuracy, they have been refining their QDE (Quad-double with extended exponent) code to calculate on both the Linux cluster and our SGI to 32 digits accuracy. This code is based on Hida and David Bailey's C++ package but has the advantage of being Fortran 90, so no interface between C++ and Fortran is needed. This is the extended precision package needed for the final Be benchmark.

Screen Saver Science (SSS)

William L. George

Samuel Small (Johns Hopkins U.)

Jacob Scott (U. of California, Berkeley)

Angel L. Villalain Garcia (U. of Puerto Rico)

<http://math.nist.gov/mcsd/savg/parallel/screen/>

The SSS project aims to develop a computing resource composed of a heterogeneous set of PCs, scientific workstations, and other available computers, that can be easily used by scientists to execute large highly distributed, compute intensive applications. Each individual computer in this system would make itself available for participating in a computation only when it would otherwise be idle, such as when its screen saver would be running. This project is based on Jini, open software architectures built in Java and intended for the development of robust network services.

There are several goals to this project. First, we hope to utilize the idle processing power of the many PCs and workstations we have available here at NIST to execute production scientific codes. The compute power of personal PCs and workstations continues to increase and they have become increasingly capable of executing large compute intensive applications due to faster processors and larger main memories. Second, the research on Grid computing has been accelerating and this SSS computing environment will allow us to develop and experiment with new highly parallel and distributed algorithms more suitable for grid type environments. Finally, the use of Java for scientific applications is of interest in general and so the development of applications for SSS will give us the opportunity to explore this topic on actual production quality applications.

Up until recently, this type of project would have required a large investment in software development just to become minimally functional and so was not practical, especially for a small team of programmers. However, with the introduction of Jini, and more specifically the Jini based network service called Javaspaces, the most difficult parts of this project have now become trivial. Javaspaces is a portable, machine independent, shared memory system that expands upon the tuple-space concepts developed in the 1980s by David Gelernter of Yale University.

The SSS project began in the summer of 2002. Substantial progress has been made in designing, implementing, and testing the basic SSS infrastructure. A generic compute server has been implemented and a new Jini service, a remote file server, has been developed to provide SSS applications with basic file I/O capabilities.

In this last year, W. George has restructured the server side of SSS to facilitate the addition of security features and to simplify the installation of SSS on user's PCs and workstations. In the summer of 2004, W. George, in collaboration with SURF student Angel Garcia, began adding much needed security features to SSS including login authentication, integrity verification of downloaded code, confidentiality (encryption) in all network communication within SSS, and trust verification of the provider of the SSS Jini service. More work is needed to complete the security framework for SSS before this system can be fully deployed.

W. George and J. Devaney collaborated with Ray Mountain, of the Physical and Chemical Properties Division (838) of the Chemical Science and Technology Lab to develop a Monte Carlo style simulation for the study of the clustering of water molecules. R. Bohn and W. George, have ported a Fortran Monte Carlo simulation to Java as a first step in creating this SSS application. After some performance tuning, initial testing has shown that the Java version performs as well as the optimized Fortran version, with respect to speed, on our SGI and Sun workstations. Because this application depends heavily on input and output of large files, the completion of this SSS application awaits the completion of the SSS Secure Remote File Service.

W. George presented an invited talk on SSS at Bowie State University on Feb. 26, 2004.

Interoperable MPI (IMPI)

William George

John Hagedorn

Judith Devaney

<http://impi.nist.gov/>

The Interoperable Message Passing Interface (IMPI) project supports the vendors of MPI, as they implement the IMPI 0.0 protocols, by maintaining the NIST IMPI conformance tester, managing the IMPI mailing list (interop@nist.gov), maintaining the IMPI specification document and its errata, and in general promoting the implementation of IMPI by the current MPI vendors.

This year J. Devaney and W. George consulted with Andrew Lumsdaine and Jeff Squires of the University of Indiana Open Systems Lab concerning the support of IMPI in the new “Open MPI” library, which will be replacing LAM/MPI and several other open source MPI implementations in 2005. IMPI is seen as an important capability for this new MPI implementation. LAM/MPI supports IMPI and we expect Open MPI to support IMPI also.

The company MPI Software Technology, Inc. completed its NIST Phase II SBIR project and submitted their final report on “Collective, Performance-Oriented Algorithms for Interoperable MPI. This research helped advance and improve their implementation of IMPI within their commercial product MPI/Pro.

IMPI has begun to have impact in the Grid computing research community. This year W. George consulted with Yutaka Ishikawa, from the University of Tokyo, regarding the use of Interoperable MPI (IMPI) in a Grid computing environment they are developing. Additionally NSF is considering funding the development of IMPI support within one or more MPI research groups.

The NIST IMPI tester has been under active use by several sites over the last 12 months, including MPI Software Technology and others. This tester is used by developers of MPI libraries as they implement the IMPI protocols.

Finally, an article on IMPI was developed which appeared in Dr. Dobb’s Journal: W.L. George, J.G. Hagedorn, and J.E. Devaney, “Parallel Programming with Interoperable MPI,” *Dr. Dobb’s Journal* **357** (Feb. 2004), pp 49-53.