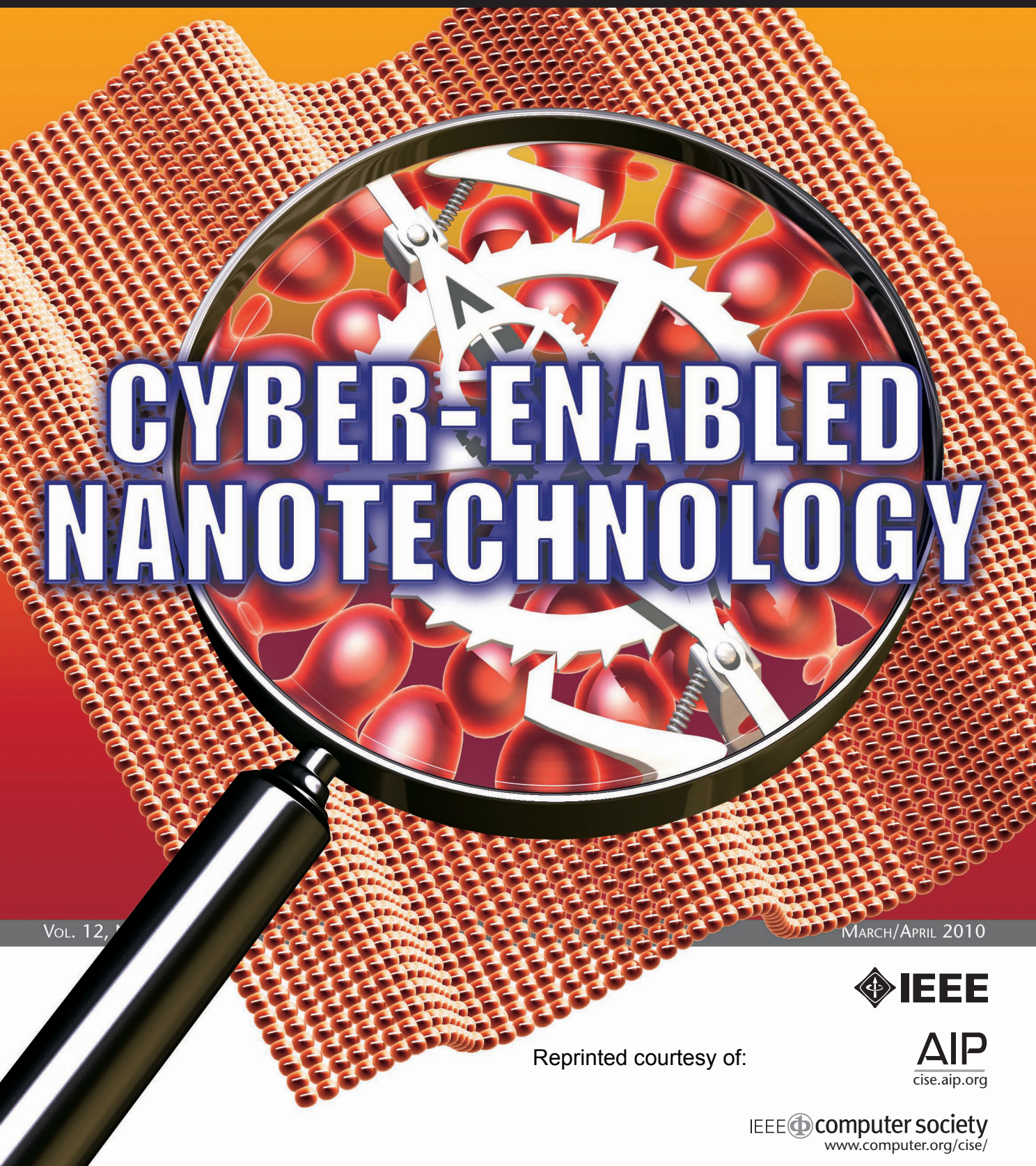
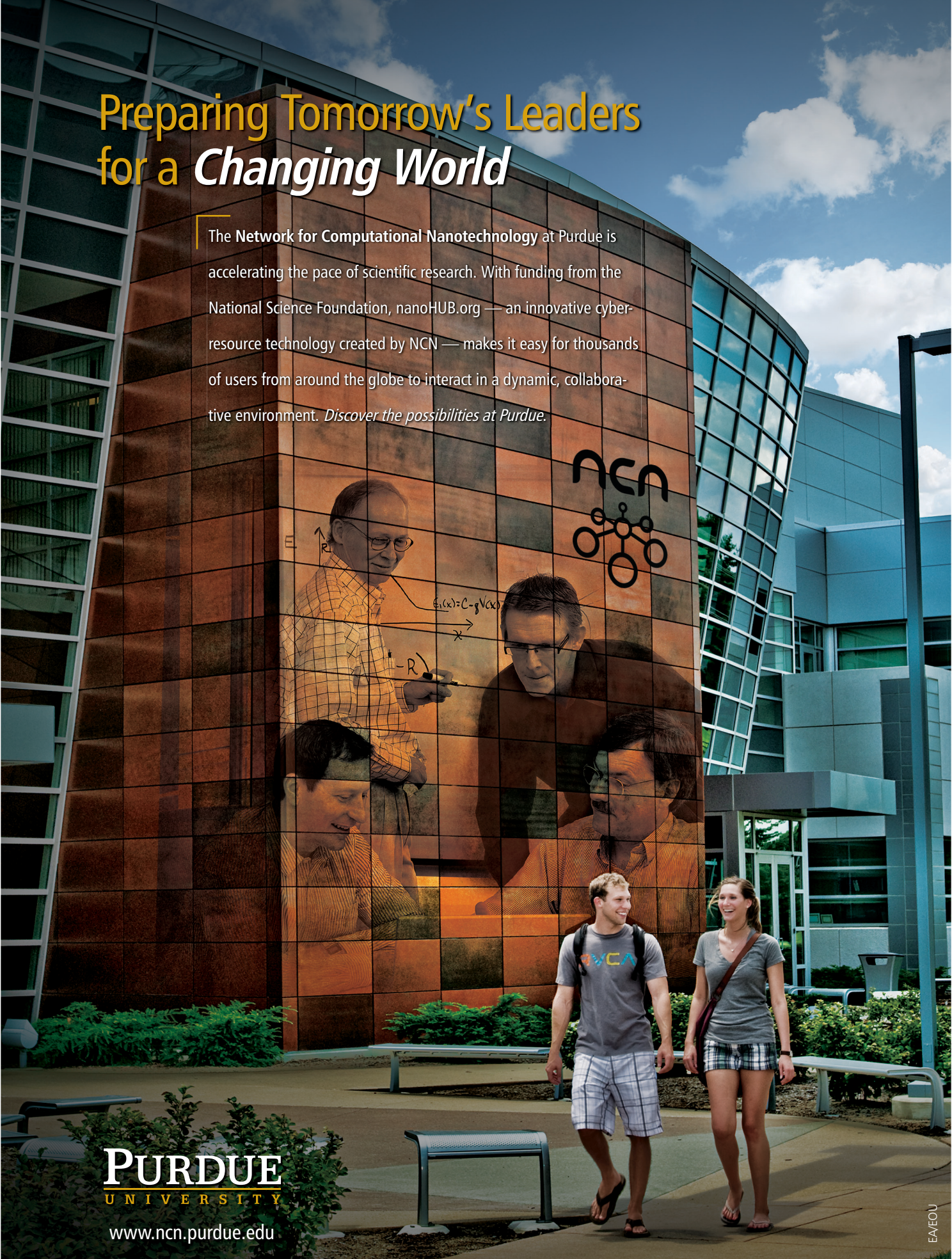
Computing in SCIENCE & ENGINEERING



CYBER-ENABLED NANOTECHNOLOGY

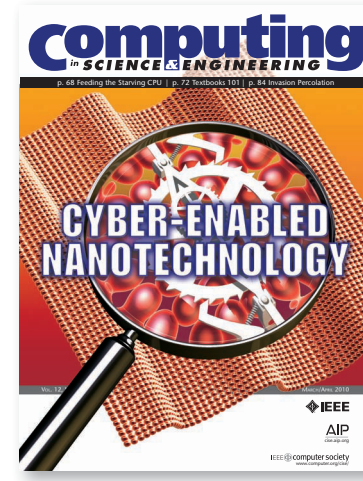
VOL. 12, [illegible] MARCH/APRIL 2010


Preparing Tomorrow's Leaders
for a Changing World
The Network for Computational Nanotechnology at Purdue is accelerating the pace of scientific research. With funding from the National Science Foundation, nanoHUB.org — an innovative cyber-resource technology created by NCN — makes it easy for thousands of users from around the globe to interact in a dynamic, collaborative environment. Discover the possibilities at Purdue.
ncn
PURDUE
UNIVERSITY
www.ncn.purdue.edu
EA/EOU

Computing in SCIENCE & ENGINEERING

March/April 2010
Vol. 12, No. 2

CYBER-ENABLED NANOTECHNOLOGY



Cover Illustration: Giacomo Marchesi
www.giacomomarchesi.com

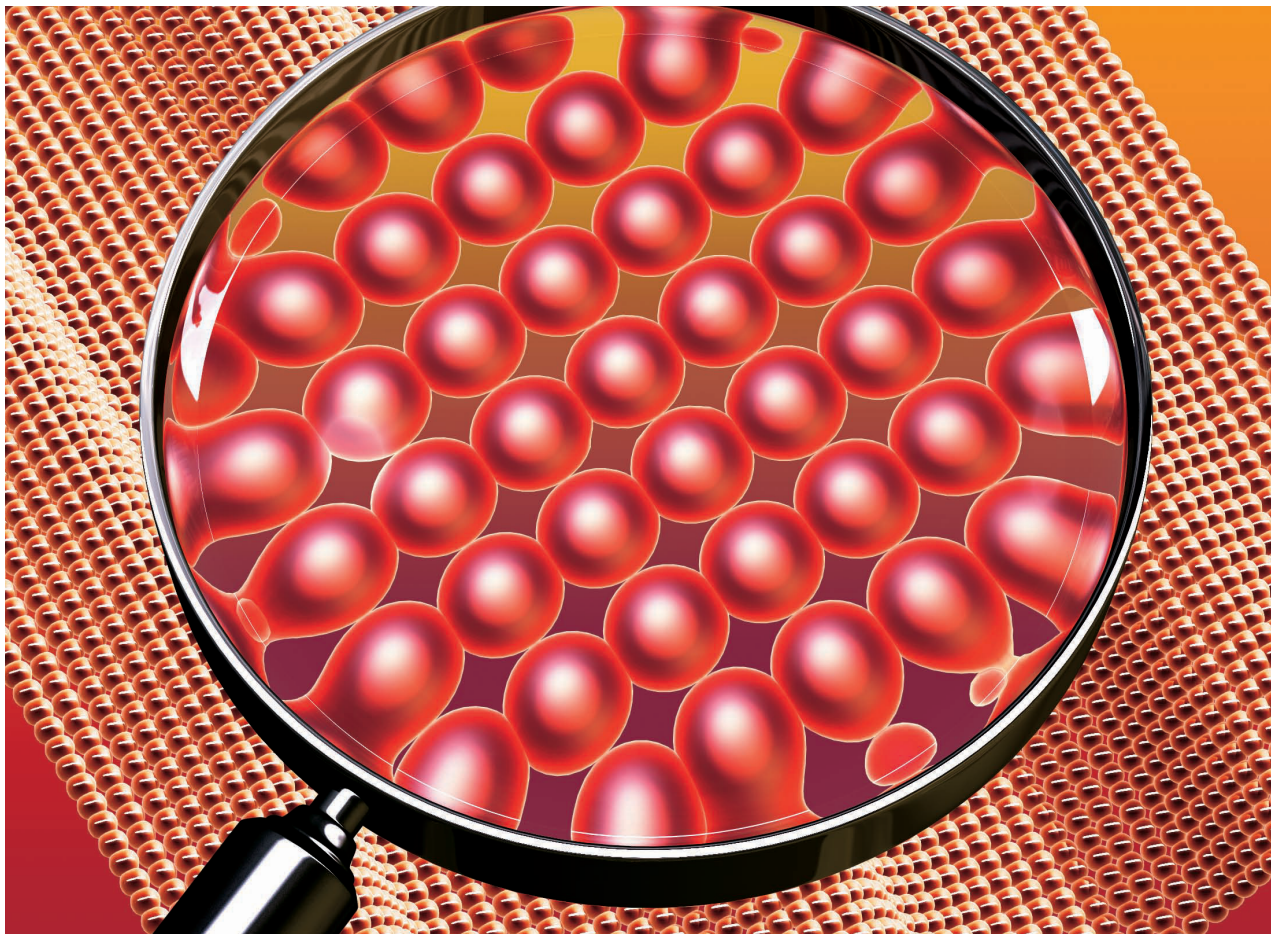
STATEMENT OF PURPOSE

Computing in Science & Engineering aims to support and promote the emerging discipline of computational science and engineering and to foster the use of computers and computational techniques in scientific research and education. Every issue contains broad-interest theme articles, departments, news reports, and editorial comment. Collateral materials such as source code are made available electronically over the Internet. The intended audience comprises physical scientists, engineers, mathematicians, and others who would benefit from computational methodologies. All articles and technical notes in *CiSE* are peer-reviewed.

For more information on these and other computing topics, please visit the IEEE Computer Society Digital Library at www.computer.org/publications/dlib/.

Printed in the USA
with soy-based ink on ANSI/NISO
standard Z39.48-1992 recycled paper.

Cyber-Enabled Simulations in Nanoscale Science and Engineering



1521-9615/10/$26.00 © 2010 IEEE
Copublished by the IEEE CS and the AIP

Alejandro Strachan, Gerhard Klimeck, and Mark Lundstrom
Purdue University

This special issue describes recent progress in atomic- and molecular-level modeling and simulation of nanoscale materials and processes, as well as efforts by the US National Science Foundation's Network for Computational Nanotechnology (NCN) to cyber-enable such simulation tools together with instructional materials and

research seminars. We believe that making advanced simulation tools widely and easily available to the research and education community will significantly enhance the impact of modeling and simulation on nanoscience and nanotechnology. To materialize this vision, NCN established nanoHUB.org, a next-generation Web portal or science gateway that lets users run live, interactive simulations, explore data, and learn—all though a simple Web browser without installing any software or providing compute cycles.

Nanoscale Science and Engineering

Models and simulation codes that describe materials in terms of the fundamental physical laws that govern electrons and atoms are playing an increasingly important role in nanoscale science and engineering. This opportunity is emerging due to a confluence of factors:

- exquisite molecular-level control in device fabrication, along with increasingly difficult experimental characterization;
- characteristic length scales in the nanometer range; and
- advances in first-principles-based simulation techniques and computing power that enable quantitative predictions of realistic materials.

Furthermore, as this special issue's first four articles exemplify, the boundaries between various traditional fields—including computational science, physics, materials, electrical and mechanical engineering, and biology—overlap at the nanometer scales, calling for an unprecedented level of collaboration that nanoHUB.org can facilitate.

Quantum mechanics describes electronic properties and atomic interactions in materials and molecules from first principles. Ab initio methods solve these equations (with some approximations) and require no experimental input. In their article, "Theory and Simulation of Nanostructured Materials for Photovoltaic Applications," Yosuke Kanai, Jeffrey B. Neaton, and Jeffrey C. Grossman discuss state-of-the-art ab initio electronic structure calculations and their application to the field of photovoltaic devices, where simulations are providing insight into current systems and potential new designs.

Unfortunately, ab initio methods remain computationally intensive and restricted to relatively small systems. Two articles in this issue describe semi-empirical methods that enable the description of materials' electronic structure, atomic structure, and thermo-mechanical response in a more computationally efficient way. In "Atomistic Modeling of Realistically Extended Semiconductor Devices with NEMO and OMEN," Gerhard Klimeck and Mathieu Luisier provide an overview of some 15 years of nanoelectronic software developments that let users quantitatively and predictively model realistically scaled devices that can contain millions of atoms. NEMO and OMEN use empirical tight binding that has been parameterized to properly represent the electronic structure of valence electrons for closed systems, such as quantum dots, and open systems, such as nanowires and resonant tunneling diodes. The authors describe the requirements set forth by researchers, developers, and educational users to outline a sequence of simulation tool development.

Yumi Park, Ya Zhou, Janam Jhaveri, and Alejandro Strachan discuss recent developments in molecular dynamics simulations, their use to predict nanoscale heterostructures' atomic structure, and their thermal properties in their article, "Molecular Dynamics Simulations of Strain Engineering and Thermal Transport in Nanostructured Materials." By averaging the role of electrons in atomic bonding and replacing them with interatomic potentials (which can be parameterized from ab initio calculations), MD can describe systems with hundreds of millions of atoms and capture a wide range of thermo-mechanical processes of interest in nanoscience.

In "Simulation of Ion Permeation in Biological Membranes," Reza Toghraee, Kyu-Il Lee, and Umberto Ravaioli discuss the Biology Monte Carlo tool for simulating ion transport across biological membranes and show how mutations affect selectivity. BioMOCA uses Monte Carlo to solve the Boltzmann transport equation—an approach widely used in semiconductor devices—and averages out atoms into a coarse-grained description. This lets BioMOCA achieve temporal time scales beyond what's possible with MD.

The simulation methods and tools that these authors describe (and that are available through nanoHUB.org) aren't the typical Web-form-based scripts, but rather are complete simulation engines that let users fully (and interactively) set up their numerical experiments, explore the results, and ask "what if?" questions.

Simulation Tools for Educators and Researchers

In "HUBzero: A Platform for Dissemination and Collaboration in Computational Science and Engineering," Michael McLennan and Rick Kennell describe the technology necessary to

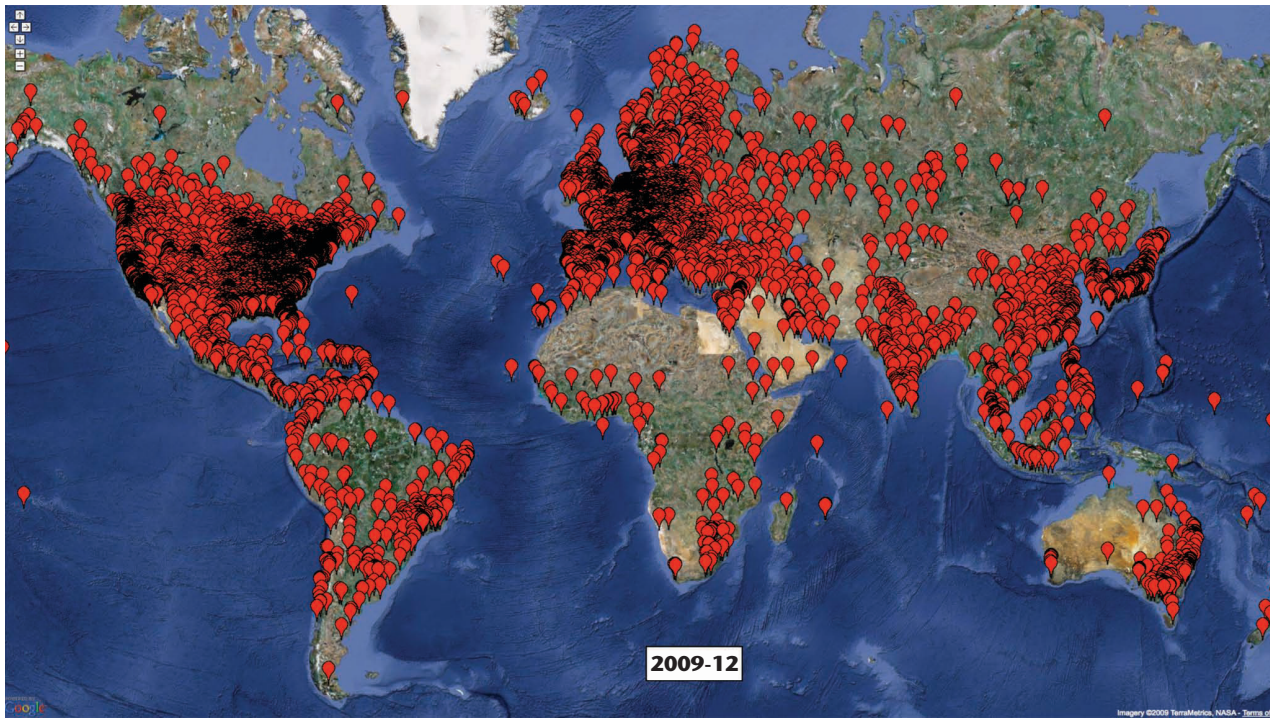


Figure 1. A map of nanoHUB users. Currently, more than 100,000 people in more than 170 countries use nanoHUB, both to create and explore simulations and to participate in courses and seminars.

cyber-enable advanced simulation tools for education and research. They also show how tool developers can effortlessly deploy realistic simulation tools through nanoHUB.org's powerful GUIs rather than having to rewrite the tools for the Web.

The nanoHUB.org community Web site now has more than 1,900 unique resources, 650 contributors, and 160 interactive simulation tools. In 2009 alone, more than 7,400 users have run approximately 400,000 simulations. In addition to online simulations, the site offers more than 40 courses on various nanotopics as well as state-of-the-art research seminars, which have propelled the annual user numbers to more than 100,000 people in 172 countries (see Figure 1 and Figure 2a). nanoHUB embraces Web 2.0 principles, enabling community contributions and facilitating reviews, questions, and wishlists. A virtual economy awards points (nanos) to contributors. Initially, users could barter nanos only for t-shirts. Now, users can barter for an increasing number of services, including tool improvements and job postings (and, possibly in the future, increased disk space and increased simulation throughput). A research study on nanoHUB data showed that the virtual economy has already had positive effects on the question and answer forum's contribution levels.[1]

nanoHUB's slogan is "online simulation and more." The introduction of full classes, tutorials, and research seminars early in the NCN project in 2003 led to a dramatic user growth from the initial 1,000 users (Figure 2a). Deploying interactive simulation tools required significant technology developments. In June 2005, converting from Web-forms-based simulation tools to fully interactive, user-friendly tools gave rise to a steep increase in annual simulation user numbers (see Figure 2b). The positive effect of a powerful GUI can be measured best by looking at monthly usage numbers for the converted tools. As an example, Figures 2c and 2d show the number of monthly interactive users and code downloads for *Schred*, one of the site's most popular tools. As Figure 2d shows, the user increase coincided with a significant decrease in source code downloads. NCN's experience indicates that most users are not interested in downloading/compiling code when an interactive and flexible GUI exists that can be used for exploratory research.

NCN invests significant effort into assessing nanoHUB.org's effectiveness in enabling new research and education. As of May 2009, NCN has identified 430 citations in the scientific literature that referenced nanoHUB.org. Figure 3 shows each paper as a single dot; a line between dots indicates a common author.

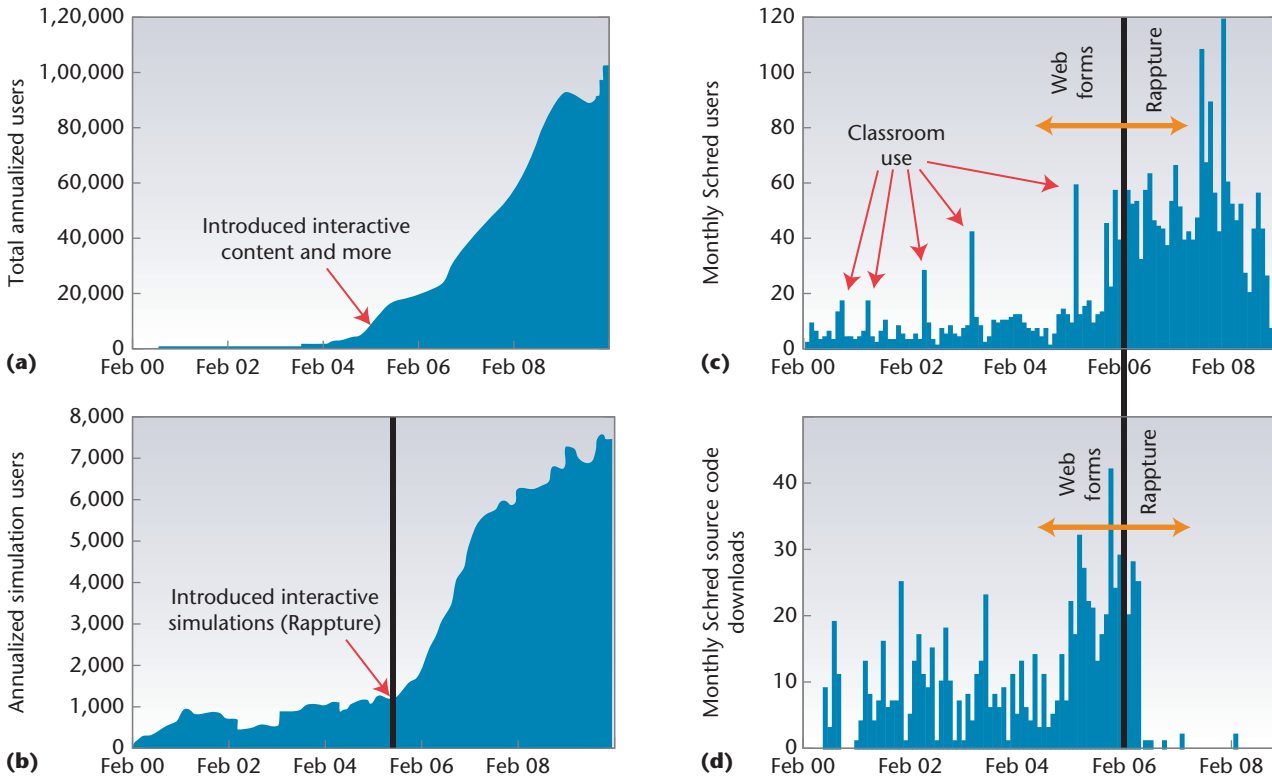


Figure 2. The impact of introducing new nanoHUB concepts, such as nanoHUB online simulation, and user-friendly GUIs to replace Web forms on nanoHUB.org. (a) Total number of annual users over time. This includes simulation and other users. Expanding nanoHUB beyond online simulation via dissemination of interactive research seminars, full classes, and tutorials increased the total number of users to more than 100,000. (b) Total annualized growth in simulation users. As these graphs show, the significant increase in simulation tool users beginning in June 2005 coincided with the deployment of interactive tools with friendly GUIs to replace the traditional Web forms. The introduction of a GUI for the popular Schred tool exemplifies this trend, showing (c) a dramatic increase in users per month and (d) an equally dramatic decrease in source code downloads. Overall, nanoHUB.org user numbers increased by factors of four to five times their pre-GUI levels, while source code downloads all but vanished.

Of these papers, 48 percent are written by NCN-affiliated authors, who form a strong paper-to-paper network within NCN (as shown inside the dashed yellow line in Figure 3). Significantly, nanoHUB.org is also enabling the development of strong networks outside NCN. In fact, more than half the papers are written by authors unaffiliated with NCN.

As the colored dots in Figure 3 show, 79 percent of the papers address aspects of nanoscience, while 16 percent refer to nanoHUB as a cyberinfrastructure. Moreover, a significant number—9 percent—address how nanoHUB can be used in education.[2] Our user registration information indicates that more than 290 classes have utilized nanoHUB at more than 90 institutions (see Figure 4). Because the site is completely open and notification of classroom usage is voluntary, we assume that the actual classroom usage exceeds these numbers. We've identified nanoHUB.org users in the top 50 US universities (per the listing

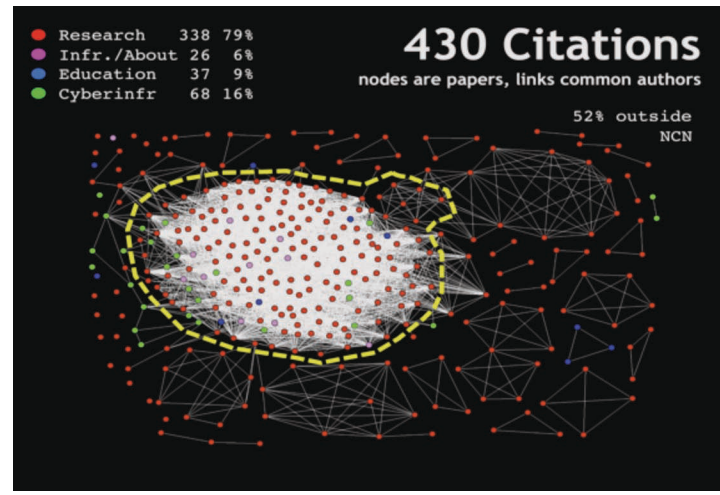


Figure 3. The nanoHUB.org network's impact on research. As of May 2009, nanoHUB.org was cited 430 times in the scientific literature. The dots within the dashed yellow line represent citations by researchers affiliated with the Network for Computational Nanotechnology's nanoHUB.org. More than half of all citations were from unaffiliated authors.

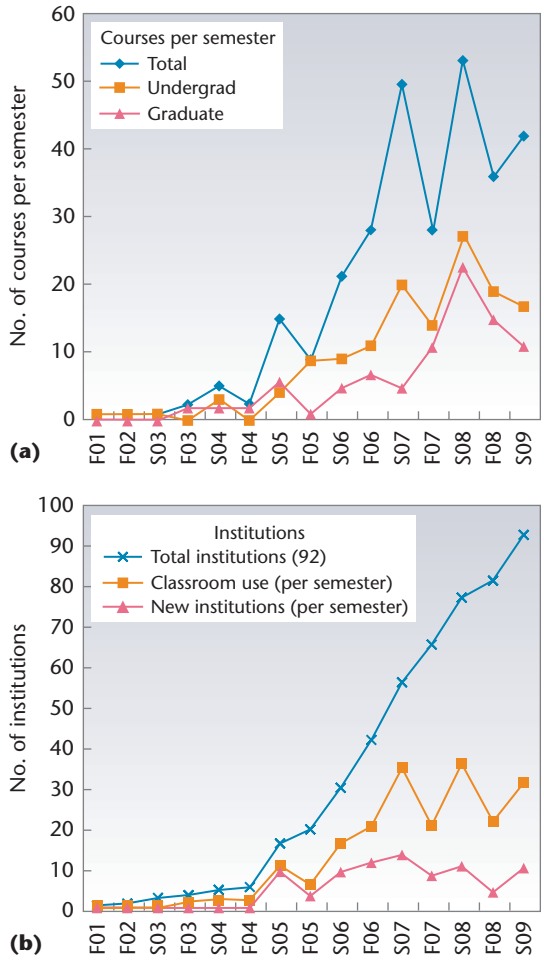


Figure 4. The nanoHUB.org network's impact on formal education. (a) The number of courses using nanoHUB.org. (b) The number of institutions using nanoHUB. nanoHUB was used in a total of 294 courses.

by *US News and World Report*). We've also identified usage in 17 percent of the more than 7,050 US institutions carrying the .edu extension in their domain name. "nano" is a tiny area in science and technology, but nanoHUB is big in many institutions.

Advances in first-principles-based simulation methods and increased computing power are enabling high-performance computational models to provide key insights into nanoscale materials, the phenomena that govern their behavior, and how we can harness them to achieve improved performance or new functionalities. Furthermore, we could significantly enhance the impact that simulation is having on this and other fields if many more investigators had access to the tools and training required to design meaningful simulations. nanoHUB.org is serving this purpose today, and we see it developing into a platform where computational scientists not only make their executables available but also deposit their codes, results, and supplemental data as part of the scientific process.

Through nanoHUB.org, numerous people can now perform online simulations to advance their research, learn and teach, and independently check results in the literature. This *scientific computing in the cloud* will not only significantly enhance the impact of simulation in the fields of nanoscale science and engineering, but also provide an unprecedented opportunity for code and model verification and validation. Such an increased level of scrutiny will help establish a set of best practices widely accepted by developers and simulation users across science and engineering fields. These are key steps necessary for predictive computational science and engineering to achieve its enormous potential and play a more central role in the development of next-generation materials and devices.

Acknowledgments

The US National Science Foundation has funded the Network for Computational Nanotechnology since 2002 under grants EEC-0228390 and EEC-0634750. Middleware development was partially funded by a National Middleware Initiative NSF grant (OCI-0438246) and a Software Development for Cyberinfrastructure NSF grant (OCI-0721680). NSF grant OCI-0749140 funded OMEN development for petascale computing and immediately related nanoHUB applications. We're also grateful to the NSF for supporting materials modeling through grant CMMI-0826356.

References

1. A. Nedossekina, "Evaluating the Effectiveness of nanoHUB Virtual Economy," nanoHUB.org, 2009; http://nanohub.org/resources/6779.
2. G. Klimeck et al., "nanoHUB.org: Advancing Education and Research in Nanotechnology," *Computing in Science & Eng.*, vol. 10, no. 5, 2008, pp. 17–23.

Alejandro Strachan *is an assistant professor of materials engineering at Purdue University, where he serves as the faculty site lead of the Network for Computational Nanotechnology and directs the Center for Predictive Materials Modeling and Simulation. His research interests include the development of predictive*

atomistic and molecular simulation methodologies to describe materials from first principles and their application in MEMS, NEMS, and electronic devices, polymer composites and other molecular solids, and active materials. Strachan has a PhD in physics from the University of Buenos Aires, Argentina. Prior to joining Purdue, he worked at Los Alamos National Laboratory and Caltech. Contact him at strachan@purdue.edu.

***Gerhard Klimeck** is a professor of electrical and computer engineering at Purdue University, where he is the director of the Network for Computational Nanotechnology and leads the development and deployment of Web-based simulation tools hosted on nanoHUB.org. His research interests include modeling nanoelectronic devices, parallel cluster computing, and genetic algorithms. He was previously a principal member of the technical staff at NASA-JPL (where he developed NEMO-3D) and a member of the technical staff at the Texas Instruments Central Research Laboratory (where he developed NEMO-1D). Klimeck has a PhD in electrical engineering from Purdue University. He is a senior member of the IEEE, and a member of the American Physical Society, Eta Kappa Nu, and Tau Beta Pi. Contact him at gekco@purdue.edu.*

***Mark S. Lundstrom** is the founding director of the Network for Computational Nanotechnology and the Don and Carol Scifres Distinguished Professor of Electrical and Computer Engineering at Purdue University. In the 1990s, he cofounded the Punch project, which provided online simulation services for research and education in micro- and nanoelectronics. His research interests are on the physics and ultimate limits of nanoscale electronic devices. Lundstrom has a PhD from Purdue. He is a member of the US Academy of Engineering and a fellow of the IEEE, the American Physical Society, and the American Association for the Advancement of Science. Contact him at lundstro@purdue.edu.*

Theory and Simulation of Nanostructured Materials for Photovoltaic Applications

Quantum mechanical electronic structure calculations are playing an ever-expanding role in advancing nanotechnology as well as in advancing our understanding and design of new functional materials. Recent research utilizing quantum mechanical electronic structure calculations is helping to improve upon our understanding of existing nanomaterials—and predict new nanomaterials—for photovoltaic applications.

Over the past half-century, computation has become an indispensable tool in science. Great advancements in high-performance computational capabilities let us solve complex mathematical equations with ever-increasing accuracy, providing detailed scientific understanding. Supercomputers with teraflops performance are already quite common at many research institutions, and petaflops machines have begun to appear at some US national laboratories. Simultaneously, scientists have expended great effort to take advantage of these increasing computational capabilities by developing new methodologies to tackle outstanding scientific challenges.

In many cases, these challenges call for knowledge of a material's electronic structure—that is, an ability to predict structure-function property relationships from an electron's viewpoint. General frameworks for accurately calculating electronic properties without resorting to empirical or adjustable parameters can benefit a range of scientific disciplines, from biochemistry to materials science, over broad materials classes. Because electrons are quantum mechanical particles, their properties must be calculated by solving the Schrödinger equation, a multidimensional differential equation. However, few cases can be solved exactly, even with today's supercomputers, and researchers typically must use physical and numerical approximations (some of which are described below) to obtain solutions. Thus, the utility of computed materials properties depends on the accuracy of such calculations. Today, viable approximations are available to solve these equations and result in good agreement with experiments. Indeed, in recent decades, such electronic structure calculations have greatly expanded their role in advancing nanotechnology. As a result, computational materials research is playing an important role in not only understanding but also designing novel nanomaterials for photovoltaic (PV) applications.

If solar energy is to become a truly viable alternative to carbon-based fuels for producing electricity, it must also become cost-competitive. Despite substantial advances in efficiency and cost over the past three decades, solar energy from

1521-9615/10/$26.00 © 2010 IEEE
Copublished by the IEEE CS and the AIP

Yosuke Kanai
Lawrence Livermore National Laboratory
Jeffrey B. Neaton
Lawrence Berkeley National Laboratory
Jeffrey C. Grossman
Massachusetts Institute of Technology

conventional silicon-based PV cells is still considerably more expensive per kilowatt-hour than power from coal, gas, or oil. A new generation of PV cells, based on nanoscale materials, has received enormous attention in recent years because of its potential to dramatically reduce the cost of converting solar energy to electricity. These new opportunities arise because of nanostructures' relatively low cost and defect density, and from major advances in our ability to control their optical, electronic, and structural properties.[1,2]

Despite the high degree of basic and applied research in this area, the power conversion efficiencies and long-term stabilities of nanomaterials-based solar cells continue to fall short of what's required for large-scale applications such as commercial or residential power generation. An important reason for this efficiency bottleneck is that many of the fundamental mechanisms and basic attributes of these solar cells are difficult to measure and only partially understood; examples include the efficiency of interfacial charge separation, the impact of impurities on energy-level alignment, optimal contact configuration, and the effects of quantum confinement. These types of properties are particularly challenging to characterize experimentally at the nanoscale, which makes the ability to predict them computationally even more crucial.

Thanks to many significant advances in high-performance computing as well as methodological improvements in electronic structure calculations over the past several decades, it's now possible to investigate material properties derived from atomistic details, and, in some cases, to predict and design new materials with the desired optoelectronic properties. Such calculations, together with experiments, can therefore play a substantial role in this global-scale challenge to both provide basic understanding and greatly accelerate materials design and discovery. In this article, we provide examples of the recent impact of these computations on both our understanding of and ability to predict nanomaterial properties for solar-energy conversion.

Fundamental Processes in PV Materials

Our discussion of theory and computation will focus primarily on atomic-scale quantum mechanical approaches. It's therefore useful to briefly review the essential processes involved in converting sunlight to electricity in solar cells from the standpoint of their fundamental electronic structure (see Figure 1).

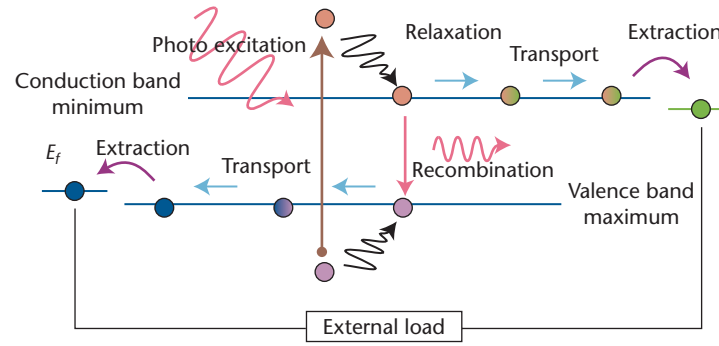


Figure 1. Physical processes involved in photovoltaic operation. Solar photon absorption results in a neutral excitation, which then "relaxes" through a loss of thermal energy to the nuclei via electron-phonon interactions. The excited electron and hole are then separated and the charge is transported to the electrodes.

The first step is solar photon absorption, resulting in a neutral excitation of an electron-hole pair, or *exciton*. From the standard Air Mass 1.5 solar flux spectrum, we note that photon absorption is a highly infrequent event at the nanoscale, even if perfect absorption is assumed. We can thus consider photo excitations to be primarily single-photon absorption processes.

In the next step, the neutral excitation is "relaxed" through a loss of thermal energy to the nuclei via electron-phonon interactions. The exciton binding energy is defined as $E_g^{QP} - E_g^{OPT}$, where E_g^{QP} is the quasi-particle gap $E_g^{QP} = E(N+1) + E(N-1) - 2E(N)$ and E_g^{OPT} is the optical gap $E_g^{OPT} = E^*(N) - E(N)$. If the exciton binding energy is small compared to thermal energy scales, a quasiparticle description is approximately valid, and we can think of this process as one in which the excited electron and hole relax toward the conduction band minimum (CBM) and valence band maximum (VBM), respectively. In bulk systems, such as silicon, the exciton binding energy is small, and this relaxation typically occurs quite rapidly (within picoseconds or faster). However, in many nano- and organic materials, the excited electron and hole are tightly bound, and the exciton binding energy can be significant (on the order of approximately 1 electron volt).

After relaxation, the next step involves dissociating the exciton, separating the electron and hole so they can be transported independently and in opposite directions toward the metal contacts (see Figure 2). Once the excited electron and hole are separated, the next key step is to transport the charge to the electrodes while avoiding recombination of the excited electron and hole. While effective mass theory often successfully describes charge transport in a crystalline material, a

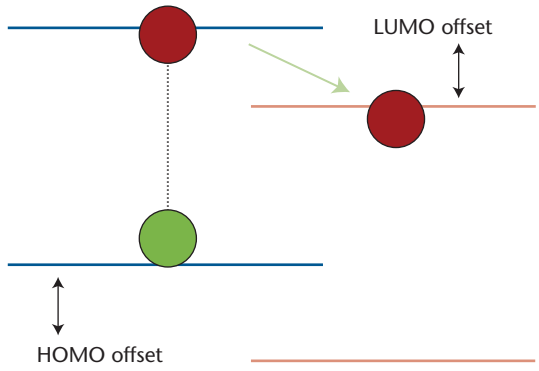


Figure 2. The charge separation process in a Type-II (staggered) energy-level alignment at a heterojunction of two different materials. The exciton, which (strictly speaking) can't be presented in this single-particle description, undergoes dissociation across the interface. The red and white circles represent an excited electron and a hole, respectively. LUMO is the lowest unoccupied molecular orbital state; HOMO is the highest occupied molecular orbital state.

temperature-dependent hopping-type mechanism has been operative in many organic materials and possibly across interfaces between nanomaterials. In some cases, such as the dye-sensitized solar cell (DSSC), transport can be assisted by ions in electrolytes.

The last step is to extract the charges into the corresponding electrodes. In semiconductor physics, concepts such as Schottky barriers are traditionally used to measure how well the contact between the semiconductor and metal follows ideal Ohmic behavior. In terms of electronic structure, how the PV active layer's frontier energy levels (such as CBM/VBM) align with the contacting levels' Fermi energies dictates the current-voltage (IV) characteristic of such interfaces.

Each of these physical processes competes with recombination's deleterious effects, in which the excited electron and hole collapse back to their original ground state. This is of particular concern for solar cells based on nanomaterials where the exciton binding energy is larger than typical thermal energy scales. The large exciton binding energy indicates that the excited electron and hole are strongly attracted to each other. Consequently, the significant spatial overlap of the two results in a high recombination probability as long as the wavefunction symmetry doesn't prohibit such an electronic transition.

Electronic Structure Calculations

The emerging maturity of *first-principles calculations*, whose results don't depend on empirical parameters, has allowed researchers to begin applying them to a broad range of materials science problems. In particular, these methods are becoming indispensable where atomistic details are crucial for understanding and predicting a given material's property-structure relations.

Density Functional Theory

Density functional theory[3] is arguably the most popular approach for investigating materials due to its balance of accuracy and applicability. Within DFT, the highly complex problem of solving the Schrödinger equation for the multidimensional many-body wavefunction is reformulated in terms of the electron density, which is a much simpler function of three dimensions. This reformulation further lets us efficiently utilize a set of single-particle Kohn-Sham (KS) equations to obtain the electronic structure and thus various other materials properties of interest. To obtain the electron density from the orthonormalized KS single-particle wavefunctions as

$$\rho(r) = \sum_{i}^{N} |\psi_i(r)|^2$$

for a N-electron system, we must face the problem of solving a set of coupled KS equations,

$$\left(-\frac{1}{2}\nabla^2 + v_{ext}(r) + v_H(r) + v_{XC}(r)\right)\psi_i(r) = \varepsilon_i\psi_i(r),$$

where $v_{ext}(r)$ is an external potential (such as that arising from the ions), $v_H(r)$ is the electrostatic potential from the electrons' charge density, and $v_{xc}(r)$ is the exchange-correlation potential. In practice, the exact form of this exchange-correlation potential is unknown, and we must approximate it from its known behaviors within some physical limits (such as homogeneous electron density). Because the coupled KS equations' potentials depend on the electron density—which in turn is calculated from solving the KS equations—the process must be solved iteratively. An efficient scheme for iteratively obtaining the "self-consistent" potential is therefore important, and several options exist. For computations, the KS wavefunctions are frequently expanded in terms of planewaves (using Fourier expansion) or physically motivated localized functions such as Gaussians. The convergence of calculated properties with respect to the size of this "basis set" is therefore also important in practice. A crucial computational bottleneck in solving for the KS

single-particle wavefunctions is that of the orthonormality requirement

$$\int dr\psi_i^*(r)\psi_j(r) = \delta_{i,j},$$

which results in the calculation of the integral for all pairs, leading to cubic scaling with the number of electrons $O(N^3)$. While this scaling might sound quite unfavorable, for electronic structure calculations, DFT remains one of the best scaling methods with respect to the number of electrons.

Many research groups today routinely use DFT to investigate the electronic structure and related ground-state properties of materials with up to several thousand electrons. Another DFT advantage is that both nuclei and electron dynamics can be treated within the existing frameworks of first-principles molecular dynamics (FPMD)[4] and time-dependent DFT extensions,[5] respectively. For small systems, we can also combine these approaches to simultaneously address the electronic structure evolution and the nuclei's motion.

DFT calculations' computational aspects have also advanced greatly over the years. In 2006, Francois Gygi and his colleagues showed the impressive performance of DFT-based FPMD with 12,000 electrons over 65,000 computing nodes, successfully taking advantage of the computational capabilities of one of the world's most powerful computers.[6] Achieving such a performance requires not only advanced computational platforms, but also tremendous efforts in optimizing numerical libraries and kernels. DFT's methodological aspects have also greatly advanced: Walter Kohn's nearsightedness principle of electrons motivated research groups to pursue computational approaches that scale as $O(N)$.[7] Such methodological advances are promising as the community pushes forward to deal with materials with ever-increasing complexities alongside of massively parallel supercomputers' steady progress.

Yet, DFT calculations aren't without shortcomings, particularly when calculating material properties away from the electronic ground state. DFT's KS single-particle wavefunctions can, in some cases and depending on the choice of exchange-correlation functional, be good approximations to quasiparticle excited states. Still, in most cases band gaps (energy gaps) are considerably underestimated, and single-particle energies deviate significantly from photoelectron spectroscopy measurements. It's therefore important to gauge the accuracy level attained for a given calculated property, and whenever possible to validate the DFT's performance using available experimental data. While understanding experimental observations and predicting *qualitative* behavior are areas in which DFT can excel, designing novel materials with specifically tailored optoelectronic properties poses a challenge because of the limited experimental comparisons available in such cases. In this regard, researchers are engaged in significant efforts to improve upon traditional DFT calculations for obtaining accurate energy levels and excited-state materials properties.

"Beyond DFT" Approaches

Electronic structure calculations based on many-body perturbation theory (MBPT)[8] and quantum Monte Carlo (QMC)[9] are two examples of first-principles approaches that go beyond static DFT. The two are particularly relevant for obtaining the accurate electronic energy-level alignment needed to understand and predict the excited-state properties such as absorption, charge separation, and charge collection that are important to PV operation.

Many research groups today routinely use DFT to investigate the electronic structure and related ground-state properties of materials with up to several thousand electrons.

MBPT is based on a set of Green's function equations, beginning with a one-electron propagator and considering the electron-hole Green's function for the response. Key ingredients are the electron's self-energy (describing the interaction with its surroundings) and the electron-hole interaction. Within this framework, quasiparticle energies of adding or removing electrons from the system are determined by a Schrödinger-like equation that contains a nonlocal, energy-dependent self-energy, Σ, in place of the exchange-correlation potential as in the KS equation. Formally, the quasiparticle energies ε_i are the solutions of

$$\left[-\frac{1}{2}\nabla^2 + v_{ext}(r) + v_H(r)\right]\psi_i(r) + \int dr'\Sigma(r,r';E)\psi_i(r') = \varepsilon_i\psi_i(r),$$

where $\Sigma(r, r', E)$ is a nonlocal, energy-dependent self-energy operator. A good approximation for Σ is obtained with Lars Hedin's GW approach,[8] in which the self-energy is truncated to first order in a power series in the screened Coulomb interaction, W, and the single-particle Green's function G is determined from the KS eigenstates. We can then express the GW self-energy as

$$\Sigma(r, r'; E) = \frac{i}{2\pi} \int d\omega e^{i\delta\omega} G(r, r'; E - \omega) W(r, r'; \omega),$$

where δ is a positive infinitesimal, $W = \varepsilon^{-1} V$, ε is the dielectric response function, and V_c is the bare Coulomb interaction. First-principles GW calculations provide a balanced treatment of exact exchange and dynamically screened correlation effects starting from a conventional DFT framework, and have been successfully carried out for bulk materials, surfaces, interfaces, and molecules.[10] Distinct approaches for formulating the GW self-energy exist in the literature. In the Hybertsen and Louie approach, we first perform a conventional DFT calculation for the target system. Next, we use the obtained KS wavefunctions as a first approximation to the quasiparticle wavefunctions, and generate the one-electron Green function accordingly. Computing the inverse static dielectric function in terms of the polarizability can be the calculation's most computationally intensive step, often scaling as $O(N^4)$. The dielectric response's frequency dependence is often evaluated using a plasmon-pole model.[8] (We could also generate the fully frequency-dependent dielectric function, though at greater computational cost.) Finally, we evaluate quasiparticle energies from the DFT energies using perturbation theory. Given the added computational cost of computing the inverse dielectric function, present-day algorithms can perform GW calculations for systems consisting of about 50 to 100 atoms.

QMC's great computational advantage is that it essentially scales linearly with the number of processors, thanks to its stochastic nature.

QMC approaches are philosophically quite different than MBPT methods for going beyond DFT. QMC effectively solves the many-body Schrödinger equation in an integral form—without resorting to a single-particle-like reformulation—using Monte Carlo techniques. In *diffusion* QMC, one of the most popular QMC forms, a trial many-body wavefunction is evolved toward the true many-body wavefunction using

$$\psi(R, \tau + \delta\tau) = \int dR' G(R, R', \delta\tau) \psi(R', \tau)$$

in imaginary time τ. The many-body Green's function $G(R,R',\delta\tau)$ can be approximated accurately for small $\delta\tau$.[9] For a long enough evolution, the Green's function projects out the true many-body wavefunction from the trial one. In practice, importance sampling is introduced with the trial wavefunction, which determines the nodal surface (where the wavefunction changes its sign), avoiding numerical difficulties related to releasing the nodal surface constraint and potential energy divergence. The trial wavefunction is usually given by a Slater determinant of the single-particle KS wavefunctions from DFT calculations, multiplied by a so-called Jastrow function to include two-body (and beyond if desirable) correlations explicitly into the wavefunction. Even with the fixed-node approximation, QMC provides highly accurate energies.[11] The QMC scaling of $O(N^3)$ is largely due to its evaluation of the Slater determinant. For each of the N electrons, we must evaluate N orbitals for the determinant. Another factor of N comes in because the basis set's size typically scales linearly with N. In some cases, researchers have shown that it's possible to reduce the scaling down to $O(N)$ by applying a unitary transformation to the single-particle wavefunctions such that they're all localized in space (which doesn't change the determinant) using spatially localized basis functions.[12]

QMC calculations' computational cost is generally significant, mostly because of the statistical sampling (inherent to Monte Carlo methods) required to sufficiently reduce the error bar on computed physical quantities. QMC's great computational advantage is that it essentially scales linearly with the number of processors, thanks to its stochastic nature.

Currently, all of these methods demand considerable computation, but rapid advances in methodologies and computational platforms are encouraging for a more routine use of these approaches in the future.

Electronic Structure Calculations of Nanoscale PV materials

Addressing the material challenges for improving PV cells will require collaborative efforts that

span both fundamental and applied scientific disciplines. Understanding the observed phenomena at a fundamental level is one of the keys to a rational approach for predicting and designing novel materials with desired behaviors. These new materials will demand that applied scientific disciplines use highly sophisticated and complex engineering approaches, which in turn require new or deeper understanding. As computational methodologies become more sophisticated and accurate, exciting opportunities exist to effectively bridge this gap: we must genuinely rise to the grand challenge of predicting and designing new materials without resorting to empirical parameters from experiments.

The computational materials science community is well positioned to address such an important challenge, and creating novel nanomaterials for inexpensive and efficient PV cells constitutes an exciting direction with great potential benefits. As we now highlight, our groups employ DFT and excited-state methods to shed light on how fundamental processes in solar cells depend so crucially on the atomistic and chemical nature of the active material's shapes and interfaces.

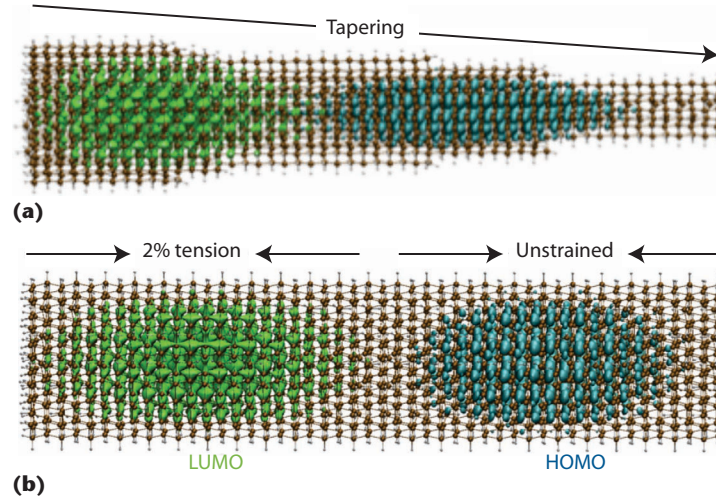


Figure 3. Example applications of DFT calculations aimed at understanding and controlling charge separation at the nanoscale. In these nanowire systems, the conduction and valence bands are shown to be spatially separated in a single nanomaterial by morphology control using (a) tapering and (b) strain.[14,15] The lowest unoccupied molecular orbital (LUMO) and the highest occupied molecular orbital (HOMO) states are mainly located at opposite ends of the tapered silicon nanorod.

Predicting Nanoscale Materials for Charge Separation

As we discussed earlier, efficiently separating charges in the active layer is a key step in converting sunlight into electricity. Because we can tune nanomaterials' electronic structure (such as by quantum confinement and topological symmetries), it's possible to control the intrinsic properties of a *single* nanomaterial without doping the material or introducing an interface with other materials. Thus, an interesting question is whether we could use nanoscale effects to design novel materials for the charge separation processes.

To answer this question, we investigated silicon nanostructures as a proof-of-principle demonstration because they've shown enormous potential in terms of their electronic properties' tuneability. In particular, silicon nanowires provide a physical path for transporting charge carriers, and we can synthesize them routinely and controllably, even well within quantum confinement regimes. Motivated by the experimental observation of tapering in silicon nanowires grown via the vapor-liquid-solid technique, we began to explore how we might modify silicon nanowires' inherent electronic structure by controlling their nanoscale morphology.

A highly efficient implementation of DFT with localized basis functions for describing KS states was employed to investigate nanomaterials consisting of over 10,000 electrons.[13] A key aspect of this work is that broken translational symmetry along the wire axis could affect the conduction band minimum (CBM) and valence band maximum (VBM) electronic states distinctively and might lead them to separate locations in real space. As Figure 3a shows, our DFT calculations predict that the lowest unoccupied molecular orbital (LUMO) and the highest occupied molecular orbital (HOMO) states are mainly located at opposite ends of a small-diameter tapered silicon nanorod.[14] This is because in a narrow tapered silicon nanowire, the quantum confinement strength varies significantly along the wire axis and the surface morphology changes substantially as well, causing different shifts in local near-gap energy levels along the axis. That said, the electronic states deeper into the unoccupied and occupied manifolds are delocalized, and the silicon nanorod's photo excitation would primarily occur across the higher excitation energy states into these levels. Thermal relaxation of the excited electron and hole to the near-gap frontier states naturally separates them spatially along the axis within such a tapered silicon nanorod. Thus, the calculations point to the possibility of spontaneous charge separation induced in a single nanomaterial by controlling the morphology, without the need for doping or interfacing with other materials.

Such a design could also be realized in a number of additional ways, apart from tapering in a single nanomaterial. For example, Figure 3b shows that we can partially strain an untapered (straight) silicon nanowire along the axis to induce CBM and VBM state localization in separate spatial regions. The reason for this effect is that the axial tensile strain lowers a silicon nanowire's frontier energy levels, while axial compressive strain lifts these levels higher due to the wavefunctions' characteristic symmetries.[15] Consequently, in a partially strained silicon nanowire, the strained section's frontier energy levels are simultaneously higher or lower than the unstrained part, effectively forming an ideal *type-II junction* (see Figure 2). Experimentally, researchers have investigated axial strain in silicon nanowires to enhance the charge carrier mobility and have realized partially strained silicon nanowires.

Investigating Electrodes for Charge Collection

For organic PV cells,[16] power conversion efficiency can be critically influenced by the electrode interface between a metal and organic PV active layer. Understanding transport phenomena in metal-molecule contacts is crucial for optimizing a PV operation's current-voltage behavior. Rigorous investigation at the individual molecule level continues to be a fundamental nanoscience challenge; the atomistic details of this interface remain outside even the best experimental probes' resolution. In our second example, we describe how we recently extended and applied first principles DFT-based methods to illuminate this problem, inspired by quantitative experiments.

Researchers often use fullerene derivatives—such as [6,6]-phenyl-C_{61} butyric acid methyl ester—as the electron-transporting material in bulk organic heterojunction solar cells because of fullerene derivatives' high electron mobilities compared to standard semiconducting polymers. However, even for simple atomic interface geometries, the injection barriers aren't well characterized. Indeed, recent scanning tunneling microscopy (STM) experiments have probed C_{60}-based molecular systems adsorbed on metallic surfaces with the goal of improving our understanding of transport and optical properties.[17,18] Although these studies have resulted in remarkable progress, there remains a critical need for a quantitative description of the spectroscopy of this key interface.

Fundamentally, the energetic positions of a molecular adsorbate's frontier molecular orbitals (HOMO/LUMO) relative to a metal surface's Fermi level dictate interfacial electronic, optical, and transport properties. The size of the interface dipole (degree of charge transfer), energy-level broadening and hybridization, and surface polarization are all key factors determining orbital energy alignment. Recent work has shown that the quasiparticle gap of a molecule weakly coupled to a metal surface (given by the difference between molecular HOMO and LUMO energies) was strongly renormalized by the metal surface's polarization response to the added electron or hole, and strongly reduced from its corresponding value in the gas phase.[19] This "image charge effect" results from nonlocal correlations that aren't captured by electronic structure calculations based on DFT with standard exchange-correlation functionals, including the local density approximation and hybrid functionals.

In other work,[20] a new approach based on DFT and the Anderson impurity model (AIM) was developed to calculate accurately charging energies and quasiparticle energy gaps of molecular systems weakly coupled to an environment, and then applied to compute the frontier orbital energies of C_{60} at Au(111)/Ag(100) surfaces, where STM experiments have thoroughly compared the level alignment for metal surfaces at both low temperatures and in ultra-high vacuum conditions.

As Figure 4 shows, this DFT-based method can resolve this apparent difference in level alignment for the two different contacts, Au(111) and Ag(100). In particular, charge-transfer screening significantly affects only the Ag case, where the LUMO considerably overlaps the surface Fermi energy. Such a marked difference in electronic energy-level alignment between C_{60} on Au(111) and Ag(100) provides new quantitative insight into metal-organic interfaces important in organic PV cells.

Quantitative Prediction and Going Beyond DFT

As we've noted, DFT's KS single-particle energies can result in energy gaps and level alignment that differ significantly with spectroscopy and transport measurements. For similar electronic characters' semiconductor interfaces, the resulting errors in energy-level alignment within DFT can be less severe in some cases due to error cancellation.[21] However, for materials of distinctly different electronic structures, the extent of such erroneous shifts could naturally be quite different across the interface, which could result in completely incorrect HOMO and LUMO

offset predictions, in addition to the interface energy gap.

One way to correct KS single-particle energies is to use MBPT.[8] In cases where the KS single-particle states are quite similar to quasiparticle states, we can utilize the KS states to compute the many-body correction that renormalizes the missing electronic many-body effects in this single-particle-like description. We recently applied this method to more complex, hybrid organic-inorganic interfaces (molecular adsorbates on silicon and other semiconductors).[22] In these systems—where DFT errors are typically much larger—our preliminary work shows that MBPT significantly improves the interface's energy-level alignment description compared to DFT.[23] This is a convenient and straightforward approach for extending DFT's applicability, but it's not without computational and methodological issues. As was recently shown, the approach's accuracy could depend significantly on how we evaluate the self-energy operator if subtle differences in the energy-level alignment need resolving.[24] Also, in some approximations, the computational cost doesn't scale as favorably as DFT—$O(N^4)$, rather than $O(N^3)$, where N is the number of electrons in the calculations.[24]

Complementing the MBPT-based approach, we recently proposed an alternative approach for improving DFT's accuracy to predict energy-level alignment using QMC calculations.[25] In this approach, we obtain the many-body correction by relating KS single-particle energies to the derivatives of the system's total energy, which we can obtain accurately using QMC. The advantages of this approach are QMC's high accuracy for a wide range of materials and the method's favorable scaling $O(N^3)$ despite its large computational prefactor.

Electronic many-body effects could be important in many scenarios. For example, to design a desirable Type-II heterojunction using an organic-inorganic interface, we investigated the use of a silicon slab's quantum confinement as a tuning parameter to suitably align the energy levels with a specific organic molecule. Although DFT calculations predict a desirable Type-II heterojunction with the slab thinner than approximately 16 layers, accurately accounting for the missing electronic many-body effects using the QMC scheme completely modifies the prediction, revealing that the interface is actually Type-I (see Figure 5). This is an extreme case where a heterojunction's qualitative optoelectronic character is incorrectly predicted using DFT calculations, exemplified by a heterojunction composed of two electronically distinct materials. Still, such results warrant caution in using DFT for predicting such novel interfaces to obtain targeted optoelectronic behavior. They also provide a glimpse into the necessary developments required for more efficient approaches for obtaining many-body corrections.

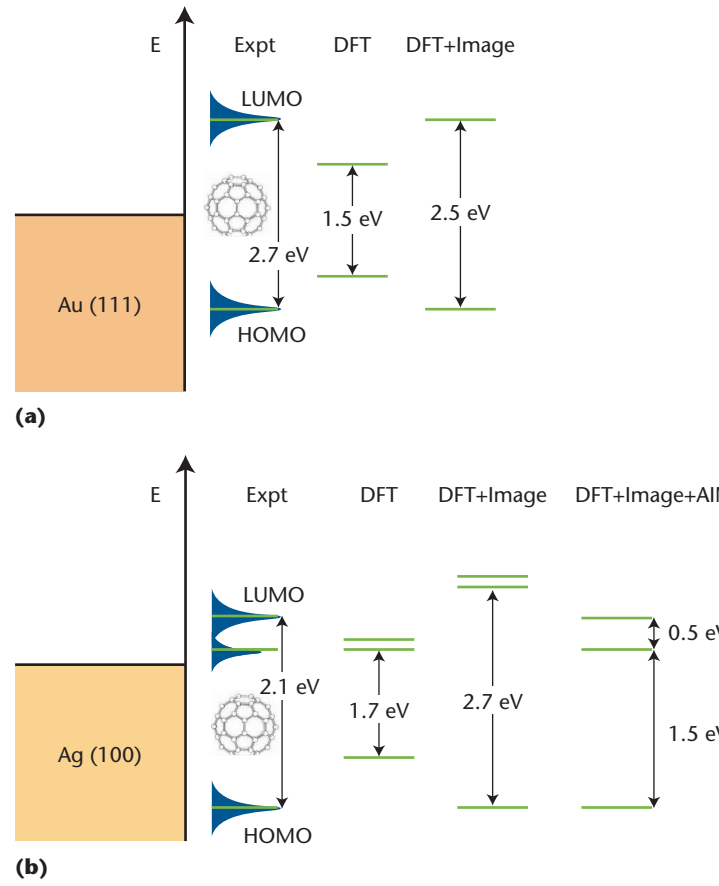


Figure 4. Schematic energy-level diagrams based on recent data[17,20] show the measured frontier orbital-level alignment for C_{60} on (a) Au(111) and (b) Ag(100). Both cases show the standard density functional theory (DFT) results (molecule placed at a relaxed distance) and quasiparticle energies without the Anderson impurity model (AIM). For Ag(100), the results also include the total spectral function with AIM, which accounts for charge-transfer screening. The AIM is essential for good agreement with the experiment in the Ag case, changing the lowest unoccupied molecular orbital (LUMO) position and enhancing the LUMO splitting. HOMO is the highest occupied molecular orbital state; Expt stands for the scanning tunneling microscopy (STM) experiments.

Although many challenges remain, first-principles calculations are beginning to reveal important insights about fundamental processes in nanoscale PV materials. Further development of computational methodologies and researcher collaboration are

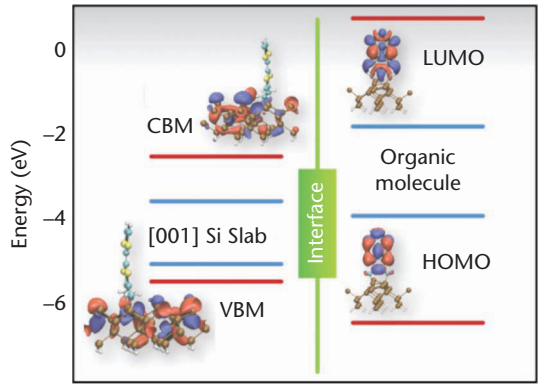


Figure 5. The calculated energy-level alignment of a representative organic-inorganic hybrid interface from single-particle energies of density functional theory (DFT) and also with electronic many-body correction obtained using quantum Monte Carlo (QMC).[25] The blue lines represent DFT levels; red lines represent QMC levels. On the organic side, LUMO is the lowest unoccupied molecular orbital and HOMO is the highest occupied molecular orbital. On the inorganic side, CBM is the conduction band minimum and VBM is the valence band maximum.

likely to foster more detailed understanding and novel designs of optimal nanomaterials for PV applications, and accelerate breakthroughs in PV nanomaterials discovery.

References

1. W. Huynh, J.J. Dittmer, and A.P. Alivisatos, "Hybrid Nanorod-Polymer Solar Cells," *Science*, vol. 295, no. 5564, 2002, pp. 2425–2427.
2. C.J. Brabec, N.S. Sariciftci, and J.C. Hummelen, "Plastic Solar Cells," *Advanced Functional Materials*, vol. 11, no. 1, 2001, pp. 15–26.
3. P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas," *Physical Rev.*, vol. 136, no. 3b, 1964, pp. B864–B871.
4. R. Car and M. Parinello, "Unified Approach for Molecular Dynamics and Density-Functional Theory," *Physical Rev. Letters,* vol. 55, no. 22, 1985, pp. 2471–2474.
5. E. Runge and E.K.U. Gross, "Density-Functional Theory for Time-Dependent Systems," *Physical Rev. Letters*, vol. 53, no. 12, 1984, pp. 997–1000.
6. F. Gygi et al., "Large-Scale First-Principles Molecular Dynamics Simulations on the Blue Gene/L Platform Using the Qbox Code," *Proc. Int'l Conf. High-Performance Networking and Computing,* IEEE CS Press, 2005, pp. 24.
7. W. Kohn, "Density Functional and Density Matrix Method Scaling Linearly with the Number of Atoms," *Physical Rev.*, vol. 76, no. 17, 1996, pp. 3168–3171.
8. G. Onida, L. Reining, and A. Rubio, "Electronic Excitations: Density-Functional Versus Many-Body Green's-Function Approaches," *Rev. Modern Physics*, vol. 74, no. 2, 2002, pp. 601–659.
9. W.M. Foulkes et al., "Quantum Monte Carlo Simulations of Solids," *Rev. Modern Physics*, vol. 73, no. 1, 2001, pp. 33–83.
10. M.S. Hybertsen and S.G. Louie, "Electron Correlation in Semiconductors and Insulators: Band Gaps and Quasiparticle Energies," *Physical Rev.*, vol. 34, no. 8, 1986, pp. 5390–5413.
11. J.C. Grossman, "Benchmark Quantum Monte Carlo Calculations," *J. Chemical Physics*, vol. 117, no. 1434, 2002; doi:10.1063/1.1487829.
12. A. Williamson, R. Hood, and J. Grossman, "Linear-Scaling Quantum Monte Carlo Calculations," *Physical Rev. Letters*, vol. 87, no. 24, 2001, http://prola.aps.org/pdf/PRL/v87/i24/e246406.
13. J.M. Soler et al., "The Siesta Method for ab initio Order-N Materials Simulation," *J. Physics: Condensed Matter*, vol. 14, no. 11, 2002, pp. 2745–2779.
14. Z. Wu, J.B. Neaton, and J.C. Grossman, "Quantum Confinement and Electronic Properties of Tapered Silicon Nanowires," *Physical Rev. Letters*, vol. 100, no. 24, 2008; doi: 10.1103/PhysRevLett.100.246804.
15. Z. Wu, J.B. Neaton, and J.C. Grossman, "Charge Separation in Strained Silicon Nanowires," *Nano Letters*, vol. 9, no. 6, 2009, pp. 2418–2422.
16. S.R. Forrest, "The Limits to Organic Photovoltaic Cell Efficiency," *MRS Bulletin*, vol. 30, no. 1, 2005, pp. 28–32.
17. H. Park et al., "Nanomechanical Oscillations in a Single-C_{60} Transistor," *Nature*, vol. 407, 2000, pp. 57–60.
18. A. Kahn, N. Koch, and W. Gao, "Electronic Structure and Electrical Properties of Interfaces between Metals and Conjugated Molecular Films," *J. Polymer Science B*, vol. 41, no. 21, 2003, pp. 2529–2548.
19. J.B. Neaton et al., "Renormalization of Molecular Electronic Levels at Metal-Molecule Interfaces," *Physical Rev. Letters*, vol. 97, no. 21, 2006; doi: 10.1103/PhysRevLett.97.216405.
20. J. Sau et al., "Electronic Energy Levels of Weakly Coupled Nanostructures: C60-Metal Interfaces," *Physical Rev. Letters*, vol. 101, no. 2, 2008; doi: 10.1103/PhysRevLett.101.026804.
21. X. Zhu and S.G. Louie, "Quasiparticle Band Structure of Thirteen Semiconductors and Insulators," *Physical Rev. B*, 1991, vol. 43, no. 17, pp. 14142–14156.
22. G.M. Rignanese, X. Blase, and S.G. Louie, "Quasiparticle Effects on Tunneling Currents: A Study of C_2H_4 Adsorbed on the Si(001)-(2×1) Surface," *Physical Rev. Letters*, vol. 86, no. 10, 2001; doi: 10.1103/PhysRevLett.86.2110.

23. S.Y. Quek et al., "Negative Differential Resistance in Transport through Organic Molecules on Silicon," *Physical Rev. Letters*, vol. 98, no. 6, 2007; doi: 10.1103/PhysRevLett.98.066807.
24. R. Shaltaf et al., "Band Offsets at the Si/SiO2 Interface from Many-Body Perturbation Theory," *Physical Rev. Letters*, vol. 100, no. 18, 2008; doi: 10.1103/PhysRevLett.100.186401.
25. Z. Wu, Y. Kanai, and J.C. Grossman, "Quantum Monte Carlo Calculations of the Energy-Level Alignment at Hybrid Interfaces: Role of Many-Body Effects," *Physical Rev. B*, vol. 79, no. 20, 2009; doi: 10.1103/PhysRevB.79.201309.

Yosuke Kanai *is a Lawrence Fellow at Lawrence Livermore National Laboratory in the Condensed Matter and Materials Division's Quantum Simulations Group. His research interests include computational materials chemistry and physics. Kanai has a PhD in theoretical chemistry from Princeton University and was a Berkeley Nanosciences and Nanoengineering Institute Postdoctoral Scholar at the University of California, Berkeley. Contact him at ykanai@llnl.gov.*

Jeffrey B. Neaton *is director of the Theory of Nanostructured Materials Facility at Lawrence Berkeley National Laboratory's Molecular Foundry. His research interests include the study of electronic excitations and transport processes in various nanostructures, including single-molecule junctions and other inorganic-organic interfaces; inorganic nanostructures; transition metal oxides; and novel light-harvesting nanostructures for solar energy applications. Neaton has a PhD in physics from Cornell University. He received the US Department of Energy's Presidential Early-Career Award for Scientists and Engineers in 2009. Contact him at jbneaton@lbl.gov.*

Jeffrey C. Grossman *is a professor of materials science and engineering at the Massachusetts Institute of Technology, where his work focuses on materials with applications in energy conversion and storage. His research interests also include problems related to surface phenomena, nanomechanical phenomena, and synthesis and assembly. Grossman has a PhD in theoretical physics from the University of Illinois and was a Lawrence Fellow at the Lawrence Livermore National Laboratory. Contact him at jcg@mit.edu.*

Atomistic Modeling of Realistically Extended Semiconductor Devices with NEMO and OMEN

Researchers have continually developed the Nanoelectronic Modeling (NEMO) toolset over the past 15 years to provide insight into nanoscale semiconductor devices that are dominated by quantum mechanical effects. The ability to represent realistically large devices on an atomistic basis has been the key element in matching experimental data and guiding experiments. The resulting insights led to the creation of OMEN, a new simulation engine.

In 1965, Moore's law predicted an exponential cost reduction with an exponential increase in the number of components per integrated circuit. Indeed, over the past 45 years, we've obtained an amazing increase in computational capabilities through the dramatic size reduction of individual transistor components. Despite predictions of insurmountable technical difficulties, sheer economic drivers have now created a global US$260 billion semiconductor industry. There is, however, a fundamental limit that we can't overcome: atoms aren't divisible, so downscaling must stop in the realm of countable atoms.

Some material layers in commercial devices have now reached the thickness limit of a few atoms. Lateral dimensions are now at 20 to 30 nanometers (80 to 120 atoms) and device geometries are no longer flat planar, but rather 3D objects. New materials have entered the device design realm to reduce leakage currents through thin layers and to deform the active transistor material through strain engineering, which improves transistor characteristics. The problems are large enough that silicon industry is considering new transistor materials, such as carbon and even the eternal materials of the future: Gallium arsenide (GaAs) and Indium arsenide (InAs). The design space's size is suddenly exploding because the detailed local atom arrangements have become critical and new materials have been added in novel 3D geometries.

Modeling and simulation might offer ways to explore options before actual experimentation. However, it's probably fair to say that we can't use any of the typical commercial semiconductor device design tools to explore the 3D atomistically defined search space. Most tools are based on continuum material assumptions and therefore ignore the mere existence of the atomic granularity. Most tools also, at very best, patch quantum mechanical effects into the simulation concept through perturbative treatments. What we need, however, is a fundamentally quantum mechanical carrier transport model built on an atomistic material description.

A suite of tools that can model realistically extended nanoelectronic devices such as 3D quantum dots, ultra-thin-body transistors, nanowires, carbon nanotubes, and graphene sheets at an

1521-9615/10/$26.00 © 2010 IEEE
Copublished by the IEEE CS and the AIP

Gerhard Klimeck and Mathieu Luisier
Purdue University

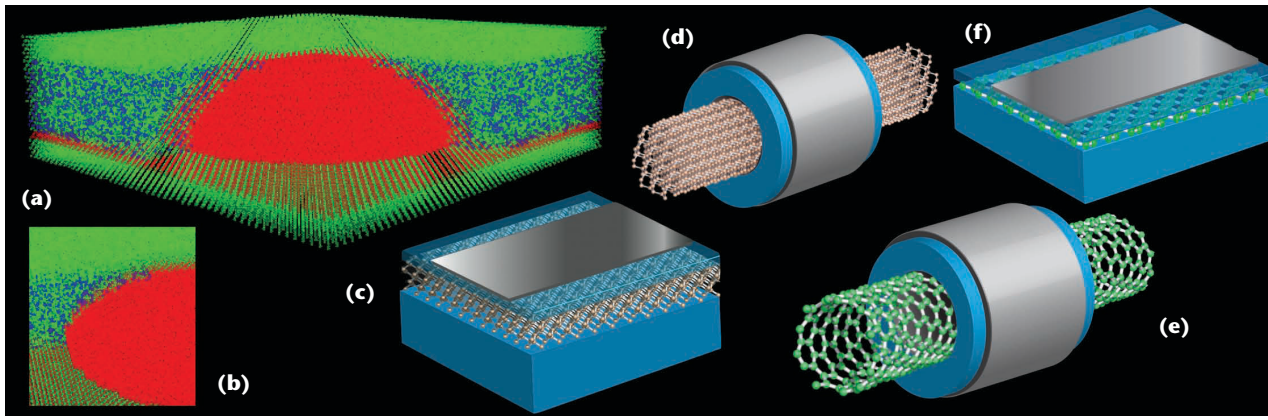


Figure 1. Nanoelectronic device geometries. (a) The multimillion-atom Nanoelectronic Modeling 3D (NEMO-3D) simulation geometry of an Indium arsenide (InAs) quantum dot on a Gallium arsenide (GaAs) substrate, capped by an alloy Indium aluminum arsenide (InAlAs) strain-reduction layer. (As atoms were omitted for clarity's sake.) (b) A zoom view of (a) from a different angle. Among the several 2D and 3D device geometries for OMEN transport simulations are (c) an ultra-scaled thin body transistor, (d) a gate all-around nanowire, (e) a gate-all-around carbon nanotube, and (f) a top-gated graphene sheet.

atomistic resolution might help narrow the search and optimization space, reduce the cost of new technology developments, and reveal new device operation concepts (see Figure 1 for examples).

Device downscaling has also increased costs for experimental determination of transistor designs, as well as for overall manufacturing; constructing a new fabrication line is approaching US$10 billion. Quantitative modeling and simulation at the atomistic scale could enable the exploration of the design space for high-performance, fault-tolerant, manufacturable devices. Modeling and simulation might therefore be the key to keeping Moore's law valid for a few more years.

NEMO-1D: The First Industrial Quantum Transport Simulator

The first analog and digital nanoelectronic devices to operate at room temperature with nanometer-scale material variations were resonant tunneling diodes (RTDs). Indeed, these devices require a quantum mechanical understanding and can't be modeled with semiclassical approaches. The Texas Instruments Central Research Lab assembled a team of theorists, computational scientists, software engineers, and—last, but not least—experimentalists to create an industrial-strength modeling tool that could drive RTD design. The primary challenge was to increase the peak-to-valley current ratio to reduce the "off current" in possible digital and analog circuits. Theoretically, this amounted to understanding the valley current's physical origin (see Figure 2).

In 1997, toward the project's end, the modeling was guiding quantitative, predictive simulation that agreed with experiments (see Figure 2b),[1] with the bottom line being that we can't dramatically reduce the off current in high-performance, high-current-density devices in standard RTDs. Researchers eventually patented design alternatives for low-power memory cells.[2] While general interest in RTDs has subsided since then, researchers achieved critical insights resulting from modeling carrier transport at the nanometer scale.

The Nanoelectronic Modeling 1D (NEMO-1D) team developed new boundary conditions that enabled treatment of extended device contact regions in which strong scattering and thermalization of carriers and electrostatic control are critical, while quantum mechanically confined states still rule over the carrier injection into the central device. Most RTD community members had suspected that incoherent scattering in the central device region was the valley current's key element. That effectively turned out to be incorrect. The critical element was actually to understand where the resonances are in energy and how they're coupled to the contacts. To enable quantitative modeling of carrier transport in room-temperature high-performance, high-current-density RTDs, we needed an atomistic representation of the device layers.

Setting the Path for General Nanoelectronic Device Simulation

While the academic and industrial interest in RTDs has subsided, we've gained critical insights into carrier transport at the nanometer scale. As we now describe in more detail, the development of NEMO-1D has set the model, user, and

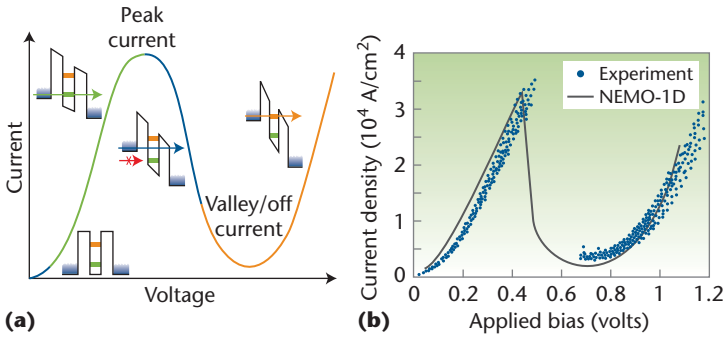


Figure 2. Resonant tunneling diode. (a) A conceptual sketch of a double-barrier structure under various bias conditions, leading to a current turn-on and turn-off with increasing bias. Controlling the peak and the valley (off) current and their relative size is critical to transistor action and the amount of energy consumed when the device is off. (b) Experimental data from 12 different current and volt (I-V) curves overlapping theoretical predictions of Nanoelectronic Modeling 1D (NEMO-1D).

developer requirements for the various simulation engines that followed.

Model Requirements

Among the requirements, models

- must *not* be based on continuum material description, but must include atomistic granularity and crystal symmetries;
- need basic mechanical atom–atom interactions (strain and phonons);
- need full-band electronic structure representation for the central valence and conduction bands (as opposed to expansions around central symmetry points like gamma);
- must represent bulk material properties, such as band gaps to within a few millielectron volts (meV) and effective masses to within a few percent, because band structure engineering works in the tuning band edges domain to within a few tens of meV;
- must have transferability from bulk to the nanometer scale;
- must be computationally lightweight enough to represent realistically extended devices (realistic structures imply tens of millions of atoms in quantum dot structures and around 100,000 atoms in ultra-scaled field effect transistors and nanowires);
- must model devices of a finite extent in realistic environments, which precludes the assumption of infinite periodic structures surrounded by a vacuum;
- must be able to include open boundary conditions, not just closed systems;
- must be able to represent atomistic disorder without additional computational burden or user interference—computation time and model tuning should not increase with degrees of disorder; and
- must have an atomistic model that can represent atomistic disorder explicitly—rather than in a statistical average way—as each device is indeed different from the next.

These requirements guided the choice of basis sets in subsequent developments, both with the NEMO team and other researchers around the world. "Exact" band gaps and masses, millions of atoms, and finite extent preclude most typical *ab initio* models. In fact, you could argue that for typical semiconductor devices, these might be unnecessary anyhow, because device simulation doesn't need to establish the existence and formation of bonds. If the bonds varied in time, the semiconductor device would be unstable and plagued by noise. However, the empirical tight-binding approach that we selected meets the requirements and needs stated earlier.

User Requirements

We envisioned various classes of users, including computational scientists, experimentalists, educators and students, and NEMO developers.

Computational scientists typically push the model development and validation. They usually have no real compute-time requirements and do whatever it takes to solve the model. Simulations are often limited to a few cases to establish the existence of solutions or to gain fundamental insight. These users are willing to move data and restart files all over the place and perform data analysis in ad hoc interfaces.

Experimentalists generally know little about the model's details, but have realistic problems to solve rapidly. They need to understand and develop device concepts, explore many different designs, and ask "What if?" questions rapidly. The primary exploration times are seconds to minutes, while secondary execution times might be the lunch hour or a few overnight simulations. With this in mind, developers must integrate data management, visualization, and export into notebooks into the overall tool experience. The software must not demand that these users intricately understand the underlying theory. Computational science experiments that add more and more basis functions to test the model's validity are inappropriate for this class of users.

As with the experimentalists, *educators and students* need a rapid simulation turnaround, but

they also need an even simpler access method. Use might be anywhere in the world and in the classroom, without any license and installation issues. In fact, neither teachers nor students are typically allowed to install additional software on classroom or laboratory computers. These users also need reuse of previous simulation results or set-up before classtime and simple results reporting and interpretation.

Developer Requirements

NEMO developers included theorists, computational scientists, algorithm developers, software engineers, and user interface designers. They must be able to work together in various software project phases. These developers have quite different requirements in their usual workflow. At one end of the spectrum, they might need code to change on a daily basis; at the other end, they might need to freeze code for several months while they test it and roll it out to users. They need a dynamical I/O design that limits the exposure of new models to users and enables developers to add models rapidly.

From NEMO-1D to OMEN

Several researchers identified the fundamental transport methodology as the non-equilibrium Green's function (NEGF) approach and established it as such in this field. Given our model requirements, we used empirical orthogonal tight-binding (ETB) approaches to represent the semiconductor materials. We implemented different basis sets, ranging from simple effective mass (single "s" orbital) to a suite of more comprehensive models, such as sp3s* and sp3d5s*, typically in nearest-neighbor representations.

ETB is based on the symmetry-formalized interaction of valence electrons on neighboring atomic sites. It ignores the core electrons and is therefore not a total energy Hamiltonian. However, we can fit ETB to match experimental and more fundamental theoretical band structure properties[3] when we improve the model to match general strain behaviors.[4,5]

NEMO-3D Development at NASA JPL

In 1998, the NASA Jet Propulsion Laboratory had a strong experimental technology group interested in developing optical detectors and lasers. Experimentalists were growing self-assembled quantum dots and needed to guide the efforts with modeling and simulation. The advent of Beowulf-based cluster computing in the JPL High Performance Computing Group made dedicated parallel compute power available to engineers. Transport simulations through 3D-resolved structures with millions of atoms were completely unfeasible; the work's focal point was to create an electronic structure simulator that could compute the confined conduction and valence band states of realistically extended quantum dot systems fully atomistically.[4] Strain is a crucial element in the self-assembled quantum dot system and is modeled through a classical ball-on-a-spring model. The code is designed to be parallel with a variety of different memory and compute time trade-off capabilities. Million-atom electronic structure simulations were first demonstrated in the year 2000. In 2003, NEMO-3D was released as open source and Purdue University continued its development.[6]

OMEN Development at ETH and Purdue

In 2004 and 2005, it became clear that the computing power needed to perform atomistic transport simulations for extended devices would soon be available. Approaches to 3D transport in nanowires using a simple effective mass model were underway already and being deployed on nanoHUB.org. During his PhD work at the Swiss Federal Institute of Technology (ETH), Mathieu Luisier worked with Purdue to develop several completely new Matlab prototypes and then a C++ foundation for a general 3D quantum transport approach.[7] Initially, the simulators didn't include incoherent scattering, and the compute time was reduced using a wavefunction approach. In 2008, Luisier began developing the OMEN transport tool at Purdue; the first implementations that included incoherent scattering were achieved in 2009.[8]

Input/Output Approaches and Interface Designs

Developing rapidly evolving software and deploying it quickly to a user base has its own interesting requirements. For instance, it's easier to build GUIs when the I/O is carved in stone and doesn't change. However, theory and algorithm developers need to add new models and algorithm parameters rapidly. For the three large software projects described here, we experimented with different I/O handling mechanisms.

NEMO-1D uses a mixed static and dynamic GUI design approach. A GUI developer implemented a set of static windows filled with dynamically defined data structure objects. We can dynamically expand device design and material data windows with new design descriptions and

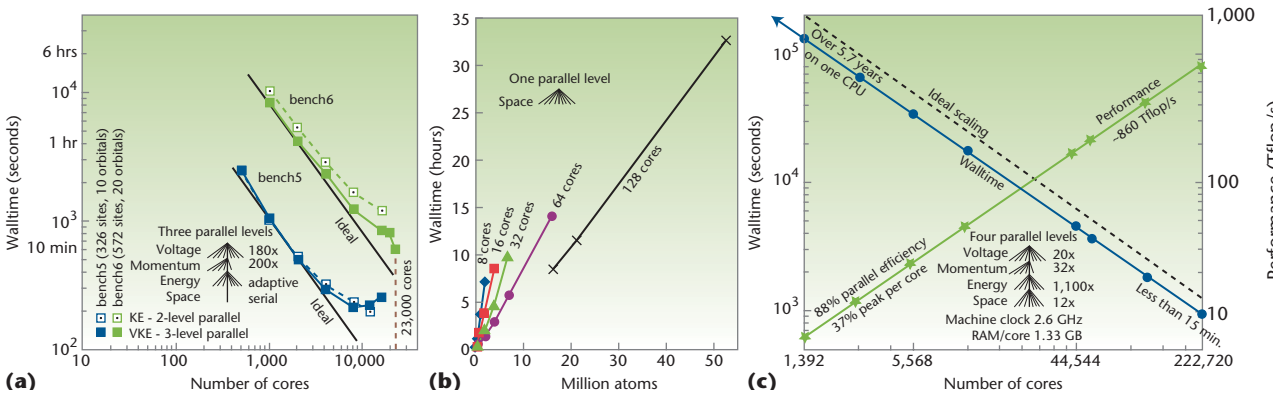


Figure 3. End-to-end performance results. (a) Parallel scaling of Nanoelectronic Modeling 1D (NEMO-1D) for a hole-based resonant tunneling diode on the Oak Ridge National Laboratory (ORNL) Cray XT3/4. Trilevel parallelism in current (I) and the double integral in momentum (k) and energy (E) results in very good scaling. (b) Demonstration of an end-to-end performance benchmark of NEMO-3D for an Indium arsenide (InAs) quantum dot embedded in Gallium arsenide (GaAs). The GaAs buffer is increased to increase the simulation domain with little effect on the central confined states. As the buffer is increased, the same states can still be computed; we've demonstrated a system size of 52 million atoms. (c) OMEN's parallel scaling results for an InAs High-Electron Mobility Transistor (HEMT) device[11,12] on ORNL's Jaguar.

material models. We structured simulation and algorithm parameters in parent-child-related dependency trees that we can dynamically configure at compile time. This dynamic design helped us decouple the static GUI work of the software engineer from the dynamic requirements of the algorithm and theory developers. We defined all I/O in C data structures; translators convert these into GUI components or batch input deck items. Another key element in NEMO-1D I/O is the ability to define all material parameters in an externally scriptable interface using LEX (a lexical analyzer) and YACC (yet another compiler compiler).

NEMO-3D uses a variant of the NEMO-1D approach except that all I/O items are defined in XML. The C++ data structures and a Java GUI are created dynamically from the XML.

OMEN handles I/O at the core based on LEX and YACC, with Tcl bindings to a metalanguage. Rappture-based drivers let us create GUIs for rapid deployment on nanoHUB.org.

Sample Modeling Results and Impact

We've extensively used NEMO-1D, NEMO-3D, OMEN, and related concepts to analyze nanoelectronic devices, as documented in over 200 peer-reviewed publications. With the deployment of these tool concepts on nanoHUB.org, more researchers are already using these tools as well. In the following, we highlight a few of the fundamental tool capabilities in terms of both physical content and numerical performance.

NEMO-1D

NEMO-1D set the standard for quantitative RTD simulation. Figure 2b shows a prototypical comparison between NEMO-1D simulations and experimental data. Researchers have modeled high-performance, high-current-density RTDs with quantitative agreement with experiments.[1] Also, simulations at low temperature demonstrate NEMO-1D's capability to quantitatively model the phonon echo in the valley current.[9]

NEMO-1D can require a significant computation time, and we developed a trilevel parallelism around voltage points and a double integral over momentum k and energy E on a relatively small Beowulf cluster.[10] In 2007, in preparation for the petascale computing initiative, we demonstrated the parallel scalability of the NEMO-1D code to 23,000 processors on a Cray XT4 (see Figure 3a).[11]

NEMO-3D

NEMO-3D can compute the electronic structure in typical semiconductor systems in the Zincblende Group III-V semiconductor and Silicon-Germanium (Si/Ge) material systems. We can apply closed and periodic boundary conditions in various dimensions, such that we can also consider 2D and 1D devices. We demonstrated end-to-end calculations of 52-million-atom systems (see Figure 3b);[6] 52-million atoms correspond to a cubic simulation domain of roughly $(101\ \text{nm})^3$; a laterally extended domain of $230 \times 230 \times 20\ \text{nm}^3$; or a nanowire geometry of $50 \times 50 \times 425\ \text{nm}^3$. This capability lets us model realistic structures for embedded quantum dot stacks, strained quantum wells, and disordered

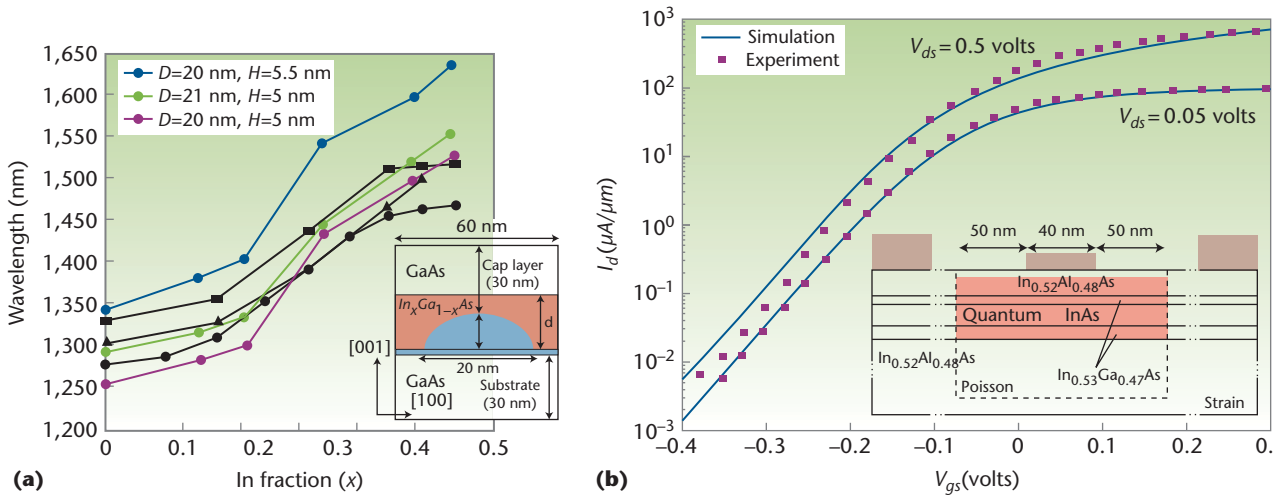


Figure 4. Sample results from NEMO-3D and OMEN. (a) Self-assembled Indium arsenide (InAs) quantum dots grown on Gallium arsenide (GaAs) capped by an $In_xGa_{1-x}As$ strain-reducing capping layer (see insert). Variation of the Indium concentration x results in a nonlinear variation in the optical emission wavelength. Black lines are from experiments and colored lines are from three different NEMO-3D simulations.[15, 16] (b) Current voltage characteristics computed in OMEN for an advanced InAs/InGaAs high-electron mobility transistor (HEMT).[12, 17]

wires. Some of the NEMO-3D usage cases include the modeling of valley splitting in tilted Si quantum wells on disordered SiGe[13] and the metrology of single impurities in Si-FinFETs' (Silicon-fin field effect transistors') modeling of single impurities,[14] where we achieved agreement of multimillion-atom electronic structure simulations with experimental data without any adjustments to previously published material parameters.

Here, we highlight the modeling of InAs self-assembled quantum dots grown on GaAs and selectively capped with $In_xGa_{1-x}As$, where the Indium mol fraction x is varied from 0 to 40 percent to achieve optical activity of the InAs quantum dots at the optical communication wavelength of 1.5 micrometers.[15] Figure 2a and the insert of Figure 4a show a nominally 20-nm wide and 5-nm tall dome-shaped quantum dot embedded by the random alloy. Figure 4a shows three experimental curves under different growth conditions, where each point on the curve represents a different device.[15, 16] We

- represented a total volume of $60 \times 60 \times 66$ nm^3 with 9-million-atoms,
- chose the nominal quantum dot size as given in the experiment,
- didn't modify any of the previously published atomistic material parameters, and
- varied only the Indium fraction.

To our surprise, the NEMO-3D simulations overlaid the experiments rather interesting nonlinear behavior. Careful analysis shows that two critical atomistic details are important to achieve such quantitative agreement. First, the InGaAs alloy's bonds must be distributed bimodally and retain their In-As and Ga-As bond lengths rather than build an average (InGa)-As bond length. Second, the quantum dot must change its shape.[15] The requirement to model the bimodal bond-length distribution precludes use of a continuum material representation. NEMO-3D provided a virtual microscope, which enabled us to understand nontrivial mechanical and electronic structure interactions that cannot be measured experimentally.

OMEN

OMEN solves the coupled Poisson and Schrödinger equations on an atomistic basis with open boundary conditions. Atoms can be arranged with each other through bonds in arbitrary 3D geometries (see Figures 1c through 1f). The electron and current density must be computed in a 3D spatial domain for different electron energies and momenta, which are accumulated in a double integral. This must be performed for many bias points in a current-voltage characteristic. Spatial decomposition, a double integral in energy and momentum, and multiple bias points offer the opportunity for four levels of parallelism. Indeed, OMEN has been scaled to more than 220,000 cores[11] on the world's second fastest computer. Such large-scale parallelism lets us reduce compute time from several years on a single CPU to approximately 15 minutes on a parallel machine. On a daily basis,

OMEN researchers are using 2,000 to 8,000 cores on parallel computers—sponsored by the US National Science Foundation and Department of Energy—to explore nanoelectronic device concepts and optimizations.

OMEN does, however, have serious limits that must be overcome. Coherent transport calculations are limited by the cross section, in which transport is modeled atomistically. The limiting number of atoms in repeated cross-section cells is around 2,000 atoms and therefore limits the cross sections to around 80 nm^2. When we introduce incoherent scattering, all electron degrees of freedom in energy and momentum are coupled and the computational cost increases by a factor of 200 to 1,000. This increase in computational cost is required because the pure wavefunction approach must be replaced by the NEGF approach at a computational cost of about 10× and for each bias point we need 20 to 100 iterations in the self-consistent Born approximation loop to couple all degrees of freedom. Finally, the treatment of surfaces (open, passivated, and relaxed), the amorphous dielectrics, and the strongly polar-bonded semiconductors remain problems with the empirical tight-binding approach.

Researchers have used OMEN to

- investigate the influence of interface roughness on the threshold voltage of triple-gate Si nanowires,[18]
- simulate the performances of *n*- and *p*-doped ultra-thin-body field-effect transistors with different crystal orientations,[19]
- reproduce experimental data for realistically extended InAs high-electron mobility transistors,[12, 17]
- study the properties of single- and double-gate ultra-thin body and gate-all-around nanowire InAs tunneling field-effect transistors,[20] and
- determine the limitations of graphene-based tunneling FETs.[21]

In addition, researchers have used OMEN to study the energy loss mechanisms through electron-phonon scattering and their impact on nanowire transistors.[8]

Broad Impact on nanoHUB.org

Use of the NEMO and OMEN tools isn't restricted to just an elite few; in fact, they're used as engines for five nanoHUB tools—Quantum Dot Lab, Bandstructure Lab, OMENwire, OMENFET, 1Dhetero, and RTDnegf—that have served more than 4,000 users in more than 60,000 simulation runs. Tool execution time varies from a few seconds on a single virtual machine to several hours on a parallel computer with 256 cores. The tools are cited in the scientific literature more than 40 times.

The NEMO and OMEN tool suite brings together material and device-modeling capabilities at the atomic resolution to impact realistically large devices. The codes perform well on serial and parallel computers, deliver results that explain and guide experiments, and are in the hands of real users with problems to solve. CiSE

Acknowledgments

NEMO and OMEN developments have been ongoing for more than 15 years and have involved many professionals, postdocs, and students who've spent countless hours working on the codes. The list of people would be too long to provide here, but the following list of critical references is part of our acknowledgments. In our work, we used nanoHUB.org's computational resources, Oak Ridge National Laboratory's leadership computing, and the TeraGrid resources Ranger and Kraken. Work at the Network for Computational Nanotechnology is supported by grants from the US National Science Foundation, the Semiconductor Research Corporation, Nanoelectronics Research Initiative, Focus Center Research Program, the Army Research Office (ARO), the State of Indiana, and Purdue University.

References

1. R.C. Bowen et al., "Quantitative Resonant Tunneling Diode Simulation," *J. Applied Physics*, vol. 81, no. 7, 1997, pp. 3207–3213, doi:10.1063/1.364151.
2. P. Van der Wagt and G. Klimeck, *Method and System for Generating a Memory Cell*, US Patent 6,667,490 to Raytheon TI Systems, Patent and Trademark Office, 2003.
3. G. Klimeck et al., "sp3s* Tight-Binding Parameters for Transport Simulations in Compound Semiconductors," *Superlattices and Microstructures*, vol. 27, 2000, pp. 519–524, doi:10.1006/spmi.2000.0862.
4. G. Klimeck et al., "Development of a Nanoelectronic 3-D (NEMO-3-D) Simulator for Multimillion Atom Simulations and Its Application to Alloyed Quantum Dots," *Computer Modeling in Eng. and Science* (CMES), vol. 3, no. 5, 2002, pp. 601–642.
5. T.B. Boykin et al., "Diagonal Parameter Shifts Due to Nearest-Neighbor Displacements in Empirical Tight-Binding Theory," *Physics Rev. B*, vol. 66, no. 12, 2002; doi:10.1103/PhysRevB.66.125207.

6. G. Klimeck et al., "Atomistic Simulation of Realistically Sized Nanodevices Using NEMO-3-D: Part I–Models and Benchmarks," *IEEE Trans. Electron Devices*, vol. 54, no. 9, 2007, pp. 2079–2089, doi:10.1109/TED.2007.902879.
7. M. Luisier et al., "Atomistic Simulation of Nanowires in the sp3d5s* Tight-Binding Formalism: From Boundary Conditions to Strain Calculations," *Physics Rev. B*, vol. 74, no. 20, 2006, doi:10.1103/PhysRevB.74.205323.
8. M. Luisier and G. Klimeck, "Atomistic Full-Band Simulations of Silicon Nanowire Transistors: Effects of Electron-Phonon Scattering," *Physics Rev. B*, vol. 94, no. 22, 2009; doi:10.1063/1.3140505.
9. R.K. Lake et al., "Interface Roughness and Polar Optical Phonon Scattering and the Valley Current in Resonant Tunneling Devices," *Superlattices and Microstructures*, vol. 20, 1996, pp. A163–A164, doi:10.1088/0268-1242/13/8A/046p.279.
10. G. Klimeck, "Parallelization of the Nanoelectronic Modeling Tool (NEMO-1-D) on a Beowulf Cluster," *J. Computational Electronics*, vol. 1, nos. 1–2, 2002, pp. 75–79, doi:10.1023/A:1020767811814.
11. B.P. Haley et al., "Advancing Nanoelectronic Device Modeling through Peta-Scale Computing and Deployment on nanoHUB," *J. Physics Conf. Series*, vol. 180, no. 1, 2009, doi:10.1088/1742-6596/180/1/012075.
12. M. Luisier et al., "Full-Band and Atomistic Simulation of Realistic 40 nm InAs HEMT," *Proc. IEEE Int'l Electronic Devices Meeting,* IEEE Press, 2008, pp. 1–4, doi:10.1109/IEDM.2008.4796842.
13. N. Kharche et al., "Valley-Splitting in Strained Silicon Quantum Wells Modeled with 2 Degree Miscuts, Step Disorder, and Alloy Disorder," *Applied Physics Letters,* vol. 84, no 9, 2004, pp. 115–117, doi:10.1063/1.2591432.
14. G.P. Lansbergen et al., "Gate Induced Quantum Confinement Transition of a Single Dopant Atom in a Si FinFET," *Nature Physics*, vol. 4, 2008, pp. 656–661, doi:10.138/nphys994.
15. M. Usman et al., "Moving Towards Nano-TCAD Through Multi-Million Atom Quantum Dot Simulations Matching Experimental Data," *IEEE Trans. Nanotechnology*, vol. 8, no. 3, 2009, pp. 330–344, doi:10.1109/TNANO.2008.2011900.
16. J. Tatebayashi, M. Nishioka, and Y. Arakawa, "Over 1.5 μm Light Emission from InAs Quantum Dots Embedded in InGaAs Strain-Reducing Layer Grown by Metalorganic Chemical Vapor Deposition," *Applied Physics Letters*, vol. 78, no. 22, 2001, p. 3469, doi:10.1063/1.1375842.
17. D. Kim and J. del Alamo, "30 nm E-mode InAs PHEMTs for THz and Future Logic Applications," *IEEE Int'l Electronic Device Meeting*, IEEE Press, 2008, pp. 719–722.
18. M. Luisier, A. Schenk, and W. Fichtner, "Atomistic Treatment of Interface Roughness in Si Nanowire Transistors with Different Channel Orientations," *Applied Physics Letters*, vol. 90, no. 10, 2007, doi:10.1063/1.2711275.
19. M. Luisier and G. Klimeck, "Full-Band and Atomistic Simulation of *N*- and *P*-Doped Double-Gate Mosfets for the 22 nm Technology Node," *Proc. Int'l Conf. Simulation Semiconductor Processes and Devices* (SISPAD), IEEE Press, 2008, pp. 1–4, doi:10.1109/SISPAD.2008.4648226.
20. M. Luisier and G. Klimeck, "Atomistic, Full-Band Design Study of InAs Band-to-Band Tunneling Field-Effect Transistors," *IEEE Electron Device Letters*, vol. 30, 2009, pp. 602–604, doi:10.1109/LED.2009.2020442.
21. M. Luisier and G. Klimeck, "Performance Analysis of Statistical Samples of Graphene Nanoribbon Tunneling Transistors with Line Edge Roughness," *Applied Physics Letters,* vol. 94, no. 22, 2009; doi:10.1063/1.3140505.

Gerhard Klimeck *is a professor of electrical and computer engineering at Purdue University, where he is director of the Network for Computational Nanotechnology and leads the development and deployment of Web-based simulation tools hosted on nanoHUB.org. His research interests include modeling nanoelectronic devices, parallel cluster computing, and genetic algorithms. He was previously a principal member of the technical staff at JPL (where he developed NEMO-3D) and a member of the technical staff at the Texas Instruments Central Research Laboratory (where he developed NEMO-1D). Klimeck has a PhD in electrical engineering from Purdue University. He is a senior member of the IEEE and a member of the American Physical Society, Eta Kappa Nu, and Tau Beta Pi. Contact him at gekco@purdue.edu.*

Mathieu Luisier *is a research assistant professor at the Network for Computational Nanotechnology, Purdue University. His research interests are in quantum transport in nanoscale devices, such as multigate and tunneling field effect transistors; parallel numerical algorithms for large computer applications; and the development of next-generation computer-aided design (CAD) tools. Luisier has a PhD in electrical engineering from the Swiss Federal Institute of Technology, Zurich; his thesis on atomistic and full-band simulation of nanowire transistors eventually evolved into OMEN. Contact him at mluisier@purdue.edu.*

Molecular Dynamics Simulations of Strain Engineering and Thermal Transport in Nanostructured Materials

Given the large surface-to-volume ratio of nanoscale and nanostructured materials and devices, their performance is often dominated by processes occurring at free surfaces or interfaces. By connecting a material's atomic structure and thermo-mechanical response, molecular dynamics is helping researchers better understand and quantify these processes.

Understanding why and how material properties change when we reduce their characteristic size to the nanoscale is the central challenge of nanoscience. Harnessing such changes to improve technological performance often requires the design, optimization, and fabrication of materials at molecular scales. Molecular dynamics determines the temporal evolution of each atom in a material, making MD a powerful technique to predict the atomic structure of complex nanoscale materials and characterize their thermomechanical properties. Furthermore, the development of interatomic potentials based on ab initio calculations—rather than experimental data—opens the possibility of predicting the structure and properties of new, previously unfabricated materials, bringing the dream of computational materials design one step closer to reality.

Here, we discuss MD's use in nanoscience using two examples. We first predict the atomic structure of nanoscale semiconductor heterostructures for electronic applications, focusing on how we can use interfaces and free surfaces to engineer a complex state of strain (a change in the lattice parameter) in a device to improve its electronic properties. The second example deals with thermal conduction in nanowires, where subcontinuum processes become important.

If educators, students, and researchers had easy access to both simulation tools and the background material required to design meaningful computer experiments, the impact of atomistic simulations on nanoscience and technology could increase significantly. Thus, we also describe efforts by the US National Science Foundation's Network for Computational Nanotechnology to make such simulations available online through nanoHUB.org.

1521-9615/10/$26.00 © 2010 IEEE
Copublished by the IEEE CS and the AIP

Yumi Park, Ya Zhou, Janam Jhaveri, and Alejandro Strachan
Purdue University

Molecular Dynamics

MD numerically solves the classical equations of motion for individual atoms in a material. We can write the equations of motion in Hamilton form as

$$\dot{r}_i = \frac{p_i}{m_i}$$
$$\dot{p}_i = F_i,$$

where the index i denotes atoms, the dots represent time derivative, and r_i, p_i, and F_i represent the position, linear momentum, and force acting on atom i, respectively. Absent external forces, we obtain the atomic forces as gradients of an interatomic potential that represents the material's total potential energy:

$$F_i = -\nabla_{r_i} V(r_1, r_2, ..., r_N).$$

The MD simulations' accuracy largely depends on the accuracy of this energy expression.

Atomic Interactions

Quantum mechanics uses first principles to describe atomic interactions in terms of their electronic structure. However, such simulations remain computationally intensive and MD with ab initio forces is restricted to relatively small systems. Thus, we use interatomic potentials that average out electrons for large-scale MD simulations. These potentials use functional forms believed to capture the materials' physics and are parameterized either from experiments or *ab initio* calculations. The simplest possible expression for an interatomic potential is a sum of pairwise terms that depend on the interatomic distance. However, such two-body potentials—where the strength of two atoms' interaction is independent of their environment—can only describe simple nonbond interactions like those between noble gas atoms. Many-body interactions are necessary to describe metallic systems, and researchers often use the embedded atom model (EAM) or modified embedded atom model (MEAM) for such simulations. In cases where interactions exhibit a significant covalent character (as in semiconductors), three-body (or angular) terms are necessary to stabilize the open structures (such as diamond, zincblende, and wurtzite).

For the Si/Ge simulation, we use the reactive force-field ReaxFF,[1,2] which has an energy expression based on the concept of partial bond order. For thermal transport calculations, we use the Stillinger-Webber potential,[3] which uses two- and three-body interactions.

Limitations and Approximations

There are two key approximations in MD that we must account for to design meaningful simulations and interpret their results. One relates to the accuracy of the energy expression used to calculate atomic forces. This is particularly important when the simulations explore structures and processes that deviate significantly from those used to parameterize the interatomic potential. In such cases, high-accuracy ab initio simulations should be used to check and tune the predictions' accuracy. The second approximation is more fundamental and stems from the use of classical (not quantum) equations of motion. Newton's laws accurately describe atomic dynamics, except for light atoms and low temperatures.

To illustrate the origin of the deviation between quantum mechanics and classical dynamics, let's consider an atom in a quadratic potential energy well with a temperature (proportional to kinetic energy) that is slowly reduced. As its temperature decreases, the atom's motion will be confined to a smaller region around the potential energy minimum, and the magnitude of its velocities will also decrease. That is, both potential and kinetic energy will decrease with the temperature. Classically, at $T = 0$ K, the kinetic energy would be zero and the atom would sit at the bottom of the potential energy well. However, Heisenberg's uncertainty principle states that it's impossible to know both a particle's position and velocity with infinite precision. Consequently, it's impossible to remove all the energy from the system; the excitation energy that remains at $T = 0$ K is called zero-point energy. Such quantum effects become important at finite, material-dependent temperatures that depend on the characteristic frequency of atomic vibrations. A material's Debye temperature indicates the beginning of quantum effects, so researchers should carefully analyze MD predictions below or near this temperature.

For readers interested in learning more, there are excellent books on MD simulations[4,5] and nanoHub.org offers a lecture series and additional references on the topic.[6]

Si/Ge Heterostructures and Strain Engineering

Undergraduate-level physics and chemistry tells us that the electronic structure of molecules and solids governs their atomic structure. This is, however, a two-way relationship; changes in atomic structure modify electronic properties. For example, when crystalline Si—a semiconductor with a bandgap of approximately 1.12 electron volts (eV) at room temperature—is molten, it becomes metallic. Similarly, changing the lattice parameter of a semiconductor (straining) affects its electronic properties. This approach is widely used to tailor semiconductors' electronic properties for specific applications.

Understanding strained heterostructures, such as thin-body Si/SiGe materials, is important for the performance of metal-oxide-semiconductor field-effect transistors (MOSFETs). Biaxial tensile strain in Si channels leads to improved electron and hole mobilities[7] in *n*- and *p*-MOSFETs. Ge is attractive due to its large hole mobility and a uniaxial strain state is promising for *p*-type devices.[8] Although achieving a biaxial strain state in planar configurations is relatively straightforward using epitaxial integration of materials with different lattice parameters, achieving uniaxial strain is more difficult. Nanopatterning is emerging as a promising avenue for asymmetric strain relaxation,[8] but the experimental characterization of strain in nanoscale materials remains challenging and only average values are typically obtained.[8] Given its atomic-level resolution, MD simulation can play an important role in this area, providing detailed information such as local strain variations and the role of surface processes for target materials.[9]

MD simulations can help us explore nanopatterning as an avenue to achieve general multiaxial strain states in semiconductor heterostructures. When a thin film is grown on a substrate with similar crystal structure and lattice parameter, the deposited material will often deform to match the substrate lattice. Such structures are called *epitaxial* (from the Greek *epi*, meaning above, and *taxis*, meaning ordered). If both substrate and film crystal structures are cubic, the film will be in a state of biaxial strain, with the in-plane strain given by the difference in lattice parameters between film and substrate,

$$\varepsilon_{iP} = (a_s - a_f)/a_f,$$

and out-of-plane strain governed by Poisson's relation,

$$\varepsilon_{oP} = -2 \cdot \varepsilon_{iP}(1-v)/v.$$

Figure 1a shows an atomic snapshot of a Si/Ge/Si epitaxial heterostructure of interest for high-speed channel materials in microelectronics. These model structures are designed to represent films grown on a Si_xGe_{1-x} substrate, and the bottom Si layer's in-plane strain is 2 percent (positive strain means tension). Ge has a larger lattice parameter than Si (about 4 percent higher), so its in-plane strain is −2.09 percent, that is, Ge is in compression. Both materials are biaxially strained. As we mentioned earlier, relaxing the Ge's strain in one of the in-plane directions would result in improved transport properties desirable for some applications. As Figure 1b shows, if we cut a narrow bar out of the thin film, we might expect strain relaxation in the transverse direction (normal to the new free surfaces) due to surface relaxation as Ge expands into the vacuum. The other in-plane direction (coming out of the page in Figure 1b) maintains its original strain because the bars are long in that direction.

We used MD simulations with an accurate interatomic potentials derived from first principles to relax Si/Ge/Si nanobars with various geometries and characterize how the height of the Ge section and the bar width affect strain relaxation in Ge. Figure 1c shows our results: lateral strain relaxation in Ge increases with decreasing bar width and increasing Ge height. The MD predictions, with no adjustable parameters, are in agreement with available experiments.[9] Our simulations also show that, for such nanoscale samples, details of the atomic state surfaces affect the amount of strain relaxation. For example, surface reconstruction, which undercoordinated surface atoms rearrange to create additional bonds and minimize energy, can increase relaxation by as much as 15 percent. This observation stresses the importance of atomic-level simulations to describe nanoscale materials.

Phonon Thermal Conduction

Thermal management in nanoscale materials is a significant challenge in various fields, including electronics and energy-conversion devices. Size effects on thermal transport arise as a result of increased scattering when heat carriers' mean free path or their wavelength become comparable to the device's characteristic size. This leads to a decrease in thermal conductivity—an unfortunate fact in many applications (including microelectronics), where good thermal conduction is desirable to dissipate an increasing power density. On the other hand, in thermoelectrics (solid-state devices that convert thermal energy to electricity), researchers can engineer size effects and boundary scattering to reduce thermal conductivity and improve device performance. In fact, reducing phonon thermal conductivity has become one of the most promising avenues to improve the figure of merit of thermoelectric devices.[10]

Continuum-based heat transport modeling in realistic nanoscale devices remains challenging because atomic details of their surfaces and internal defects play a dominant role. For example, a thin amorphous surface layer in Si nanowires can lead

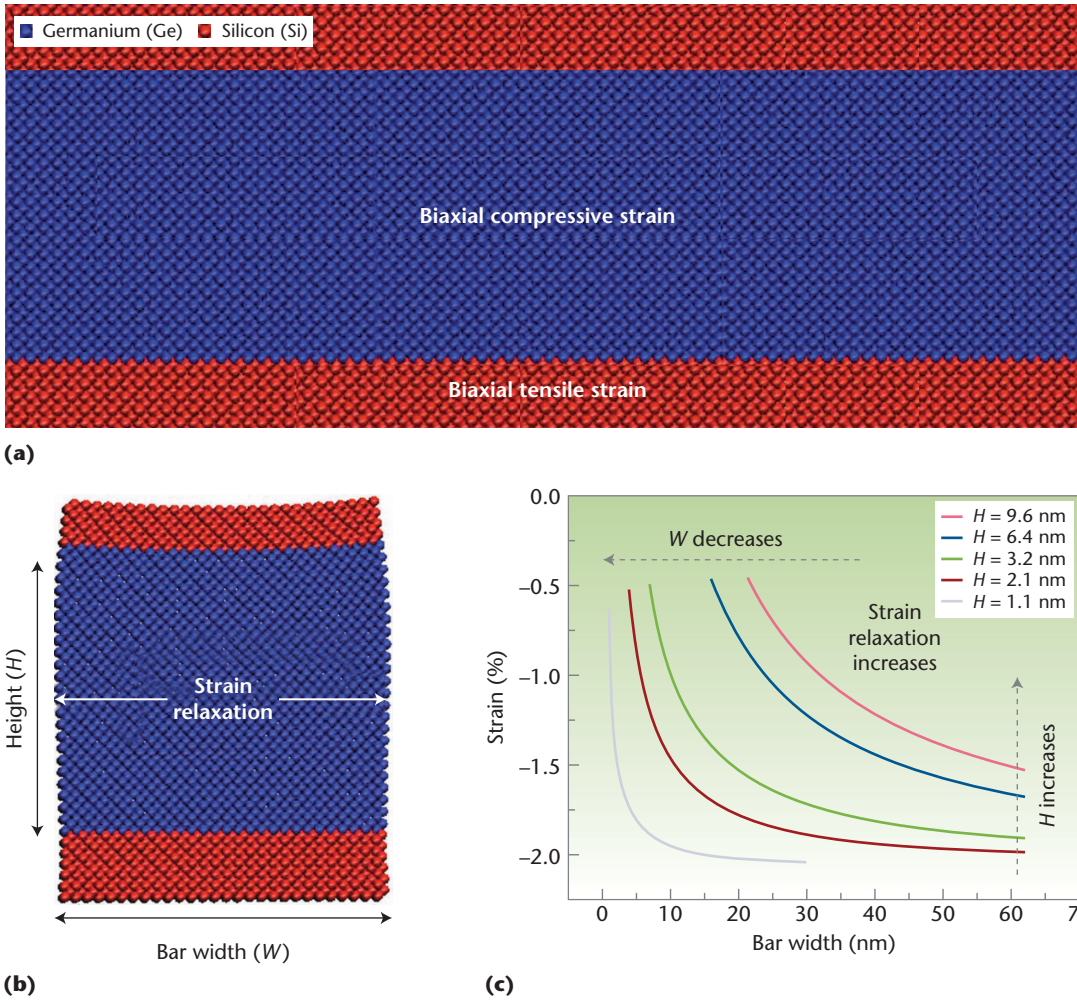


Figure 1. An atomistic model of the Si/Ge/Si heterostructures. (a) An epitaxial structure with biaxial strain, (b) a nanobar showing lateral strain relaxation, and (c) the average transverse strain of Si/Ge/Si nanobars as functions of bar width and height of the Ge section.

to a 100-fold reduction in thermal conductivity from the bulk value,[11] and the physics that leads to this effect is not captured by the Boltzmann transport equation. In this context, MD is a powerful tool for understanding and quantifying nanoscale thermal transport because it explicitly describes phonons, free surfaces, and interfaces.

In a material subjected to a temperature gradient, thermal conduction occurs to restore the thermodynamic equilibrium. Thermal conductivity κ is a materials property that relates heat flux $\vec{J}$ and temperature gradient by $\vec{J} = -\kappa\vec{\nabla}T$, where the negative sign indicates an energy flow from hot to cold. In general, κ is a second-rank tensor, but in isotropic and cubic materials, symmetry reduces it to a scalar.

Here, we calculate thermal conductivity using Mullër-Plathe's nonequilibrium method,[12] where a known heat flux is imposed on the system by periodically swapping atomic velocities in predetermined cold and hot regions. The MD-described material responds by transporting energy from the hot to the cold regions and a temperature gradient develops. After a steady state is reached, which requires between 50 picoseconds and approximately 1 nanosecond of simulation, depending on system size,[12,13] we use the average heat flux and temperature gradient to compute thermal conductivity.

We characterize size effects on phonon thermal conduction in silicon nanowires. We describe atomic interactions using the Stillinger-Weber potential, which is widely applied in the study of bulk and nanoscale Si structures. Non-equilibrium MD simulations using this potential predict bulk thermal conductivity within 30 percent of experiments.[14] According to kinetic theory, we can write phonon thermal conductivity in terms

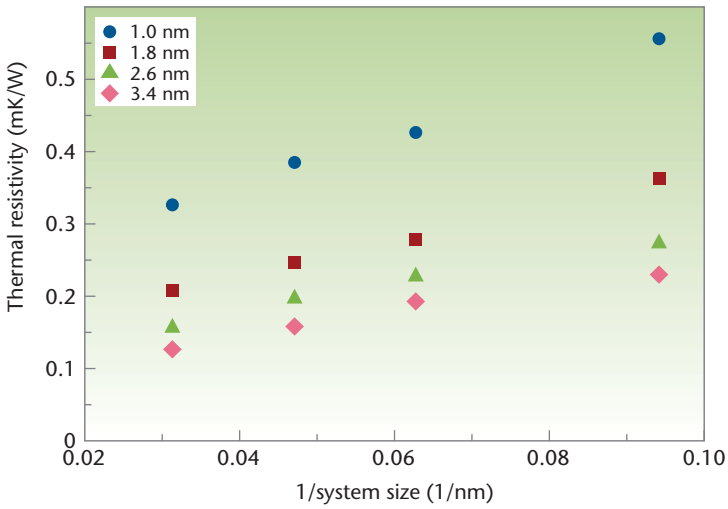


Figure 2. The inverse of thermal conductivity for wires of different diameters as a function of inverse length for $T = 300$ K. Size effects on thermal transport can be observed as thermal resistivity is linearly correlated to the nanowires' inverse length.

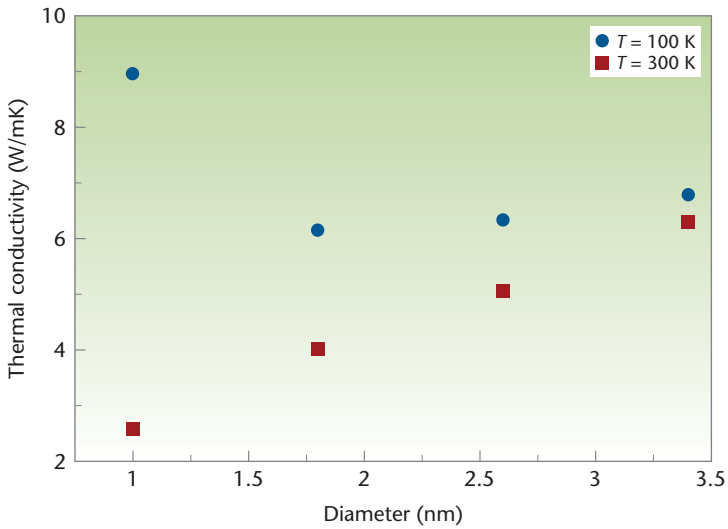


Figure 3. Thermal conductivity as a function of nanowire diameter at $T = 100$ K and 300 K. At $T = 300$ K, thermal conductivity decreases monotonically with decreasing nanowire diamater. At $T = 100$ K, diameter reduction leads to increased thermal conductivity in nanowires with diameters smaller than 1.8 nm.

of the phonon mean free path l, mean velocity v, and the lattice-specific heat-per-unit volume C_v:

$$\kappa = \frac{1}{3} C_v v l.$$

In nanoscale materials, surface and interface scattering reduces the effective phonon mean free path. Thus, nanowires' thermal conductivity depends on their length, radius, and surface details. Figure 2 shows the inverse of thermal conductivity (thermal resistivity) for wires of different diameters as a function of inverse length (defined as the distance between the hot and cold regions where heat flux is imposed) for $T = 300$ K. The MD simulations predict a linear relationship between thermal resistivity, indicating that l can be expressed as

$$\left(l_{\infty}^{-1} + 2L^{-1}\right)^{-1},$$

where l_{∞} is the phonon mean free path in finitely long nanowires and L is the separation between the hot and cold regions.[13]

We'd also expect thermal conductivity to decrease with decreasing nanowire diameter due to enhanced surface scattering. Figure 2 shows this general trend, which has previously been observed theoretically[15] and experimentally.[16] However, as Inna Ponomareva and her colleagues predicted,[14] this is not always the case. Figure 3 shows thermal conductivity of 21.3-nanometer long nanowires as a function of their diameter for $T = 100$ K and 300 K. At $T = 100$ K, thermal conductivity decreases as the nanowire diameter is reduced to 1.8 nm, but further reduction leads to a significant increase. Ponomareva and her colleagues attribute this behavior to confinement, but its origin is not fully understood. Interestingly, our simulations show that the increase in conductivity for small-diameter wires isn't present when the temperature is increased to 300 K, where we observe a monotonic decrease with decreasing diameter. Such results from molecular simulations provide important insight into the fundamental limits of miniaturization of electronic and energy conversion devices and could play an important role in guiding future experiments.

Because MD simulations are classical in nature, we must carefully analyze the effects of quantum dynamics. For Si nanowires, Giulia Galli and her colleagues found that quantum corrects amount to less than 2 percent for thermal conductivity,[11] so we don't expect the trend we observed in Figures 2 and 3 to result from classical artifacts.

As our examples show, when we account for MD's inherent limitations and approximations in designing and interpreting simulations, the technique is a powerful tool for uncovering and quantifying the underlying atomic-level processes that govern materials' behavior.

Molecular modeling will play an increasingly important role in the design and optimization of new materials and devices. It's therefore important for new generations of engineers and scientists to become familiar with its concepts, approximations, and applicability. To achieve this and increase the impact of MD simulations in research and education, we've deployed the nanoMaterials simulation toolkit (https://nanohub.org/tools/matsimtk) so that users from around the world can perform online MD simulations at nanoHUB.org.[17] The tool consists of an easy-to-use graphical user interface around our group's research MD code.

As Figure 4 shows, people can use nanoMaterials to analyze their results graphically and interact with molecular models. We use the tool for teaching, training new group members, and increasingly, for research. Along with the tool, we've deployed educational materials[6] and tutorials[18] to help new users learn the fundamentals of MD simulations, design and interpret simple computational experiments, and begin creating MD simulations using the nanoMaterials tool.

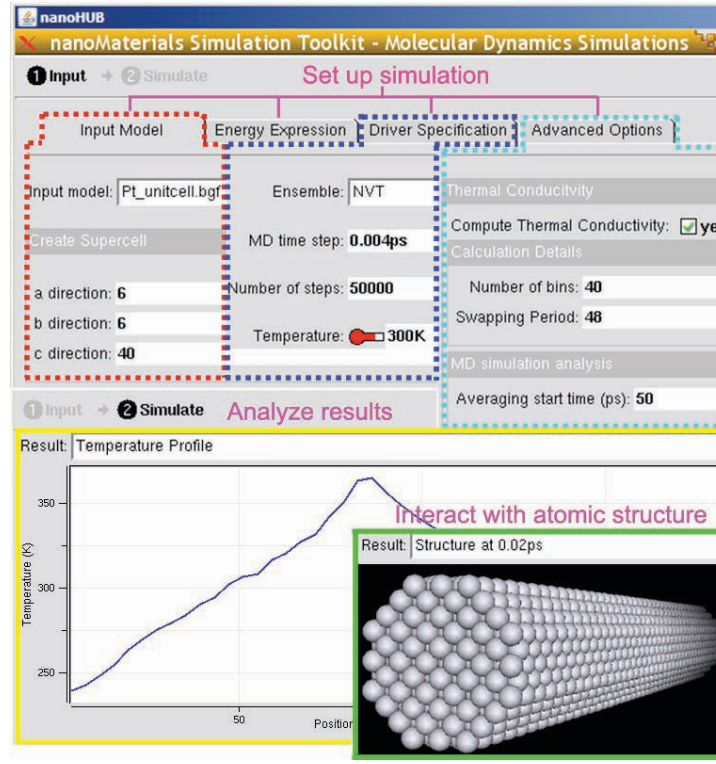


Figure 4. Thermal conduction simulation using the nanoMaterials simulation toolkit. The tool lets users set up their own simulations and analyze the results graphically.

Acknowledgments

This work was partially supported by the US National Science Foundation's Network for Computational Nanotechnology and by the Microelectronics Advanced Research Corporation and its Focus Center on Materials, Structures and Devices. Computational resources were provided by nanoHUB.org and Purdue University.

References

1. A.C.T. van Duin et al., "ReaxFF: A Reactive Force Field for Hydrocarbons," *J. Physical Chemistry A*, vol. 105, no. 41, 2001, pp. 9396–9409.
2. A.C.T. van Duin et al., "ReaxFF$_{SiO}$ Reactive Force Field for Silicon and Silicon Oxide Systems," *J. Physical Chemistry A*, vol. 107, no. 19, 2003, pp. 3803–3811.
3. F. Stillinger and T. Weber, "Computer Simulation of Local Order in Condensed Phases of Silicon," *Physics Rev. B*, vol. 31, no. 8, 1985, pp. 5262–5271.
4. M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids*, Oxford Univ. Press, 1989.
5. D. Frenkel and B. Smit, "Understanding Molecular Simulation," *Algorithms to Applications (Computational Science Series)*, 2nd ed., vol. 1, Academic Press, 2001.
6. A. Strachan, "MSE 597G An Introduction to Molecular Dynamics," 2008; http://nanohub.org/resources/5838.
7. M.L. Lee et al., "Strained Si, SiGe, and Ge Channels for High-Mobility Metal-Oxide-Semiconductor Field-Effect Transistors," *J. Applied Physics*, vol. 97, no. 1, 2005, doi:10.1063/1.1819976.
8. P. Hashemi et al., "Asymmetric Strain in Nanoscale Patterned Strained-Si/Strained-Ge/Strained-Si Heterostructures on Insulator," *Applied Physics Letters*, vol. 91, no. 8, 2007, doi:10.1063/1.2772775.
9. Y. Park et al., "Strain Relaxation in Si/Ge/Si Nanoscale Bars from Molecular Dynamics Simulations," *J. Applied Physics*, vol. 106, no. 3, 2009, doi:10.1063/1.3168424.
10. A.I. Hochbaum et al., "Enhanced Thermoelectric Performance of Rough Silicon Nanowires," *Nature*, vol. 451, no. 10, 2008, pp. 163–168.
11. D. Donadio and G. Galli, "Atomistic Simulations of Heat Transport in Silicon Nanowires," *Physics Rev. Letters*, vol. 102, no. 19, 2009, doi:10.1103/PhysRevLett.102.195901.
12. F. Mullër-Plathe, "A Simple Nonequilibrium Molecular Dynamics Method for Calculating the Thermal Conductivity," *J. Chemical Physics*, vol. 106, no. 14, 1997, pp. 6082–6085.
13. Y. Zhou, B. Anglin, and A. Strachan, "Phonon Thermal Conductivity in Nanolaminated Composite Metals via Molecular Dynamics," *J. Chemical Physics*, vol. 127, no. 18, 2007, doi:10.1063/1.2802366.
14. I. Ponomareva, D. Srivastava, and M. Menon, "Thermal Conductivity in Thin Silicon Nanowires: Phonon Confinement Effect," *Nano Letters*, vol. 7, no. 5, 2007, pp. 1155–1159.

15. N. Mingo et al., "Predicting the Thermal Conductivity of Si and Ge Nanowires," *Nano Letters*, vol. 3, no. 12, 2003, pp. 1713–1716.
16. D. Li et al., "Thermal Conductivity of Individual Silicon Nanowires," *Applied Physics Letters*, vol. 83, no. 14, 2003, pp. 2934–2936.
17. A. Strachan et al., "Nano-Materials Simulation Toolkit," nanoHUB.org, 2009, http://nanohub.org/resources/1692.
18. A. Strachan, "Materials Strength: Does Size Matter? nanoMaterials Simulation Toolkit Tutorial," nanoHUB.org, 2007; http://nanohub.org/resources/2322.

Yumi Park *is a PhD student in materials science and engineering at Purdue University. Her research interests include molecular dynamics modeling to understand mechanical behavior of semiconductor heterostructures for microelectronics application. Park has a MS in broadband information and communication from Inje University, South Korea. Contact her at park201@purdue.edu.*

Ya Zhou *is a PhD student in material science and engineering at Purdue University. Her research interests include molecular dynamics simulation of thermal conduction in nanostructured materials. Zhou has an MS in materials science and engineering from Tsinghua University, China. Contact her at zhou11@purdue.edu.*

Janam Jhaveri *is an undergraduate student in the School of Electrical and Computer Engineering at Purdue University. His research interests include nanoscale semiconductor devices and materials, focusing on the thermal conductivity of silicon nanowires. Contact him at jjhaver@purdue.edu.*

Alejandro Strachan *is an assistant professor of materials engineering at Purdue University. His research interests include the development of predictive atomistic and molecular simulation methodologies to describe materials from first principles and their use in novel areas to help solve current challenges; specific areas of focus include microelectromechanical and nanoelectromechanical systems and electronic devices, polymer composites and other molecular solids, and shape memory and active materials. Strachan has a PhD in physics from the University of Buenos Aires, Argentina. Contact him at strachan@purdue.edu.*

Simulation of Ion Permeation in Biological Membranes

As part of nature's solution for regulating biological environments, ion channels are particularly interesting to device engineers seeking to understand how nanoscale molecular systems realize device-like functions, such as biosensing of organic analytes. By attaching molecular adaptors inside genetically engineered ion channels, it's possible to further enhance this biosensor functionality.

From a physiological point of view, ion channels regulate ion transport to maintain the internal ion composition vital for cell survival and functionality. Because these channels are in charge of electrical signaling in the nervous system, malfunctioning ion channels are associated with many diseases; studying and understanding different aspects of these proteins is therefore of great pharmaceutical importance.

There are many types of ion channels, which are generally classified according to their ion selectivity, gating mechanism, or sequence similarity.[1] Different species' device-like functions—such as gating/switching and selectivity/discrimination—have motivated the engineering community to design novel biodevices.[2] Furthermore, modern biology offers the tools to reengineer channels to alter their functions.[3] DNA detection[4] is especially interesting; in particular, single molecule detection is a highly desirable capability and could revolutionize many nanoscale electronic and electromechanical applications.[5] In general, we need a hierarchy of simulation methodologies to study different aspects of a biological system like ion channels.

We used BioMOCA—a Boltzmann Transport Monte Carlo (BTMC) code developed at the University of Illinois at Urbana-Champaign—to simulate ion transport in an electrolyte environment through ion channels or nanopores embedded in membranes.[5] Our approach is based on BTMC and particle-particle-particle-mesh (P^3M) methodology,[6] which splits the electrostatic forces into short- and long-range components. Ions are the only particles in motion, moving according to Newtonian physics, and we dampen them using frequent scattering events with water molecules. To prevent ions from overlapping each other, we've used the Lennard-Jones 6-12 potential to mimic the ionic core repulsion.

Here, we describe our work using the Streptococcus bacteria as an example. To determine the diffusion coefficient in the narrow pore, we used the 1D hydrodynamic drag-wall model.[7] Also, to accelerate solution of the Poisson equation, we employed the grid-focusing method. As a case study, we simulated the mutant $(M113F)_6(M113C\text{-}D8RL2)_1$-$\beta$-CD and compared our results with the available experimental data.

Example Case: Streptococcus

α-Hemolysin, a pore-forming toxin produced by bacteria Streptococcus, lyses red blood cells and is involved in many human diseases.[8] Upon pore formation, ions and small molecules such as amino acids flow out of the cell, while water from

1521-9615/10/$26.00 © 2010 IEEE
Copublished by the IEEE CS and the AIP

Reza Toghraee, Kyu-Il Lee, and Umberto Ravaioli
University of Illinois at Urbana-Champaign

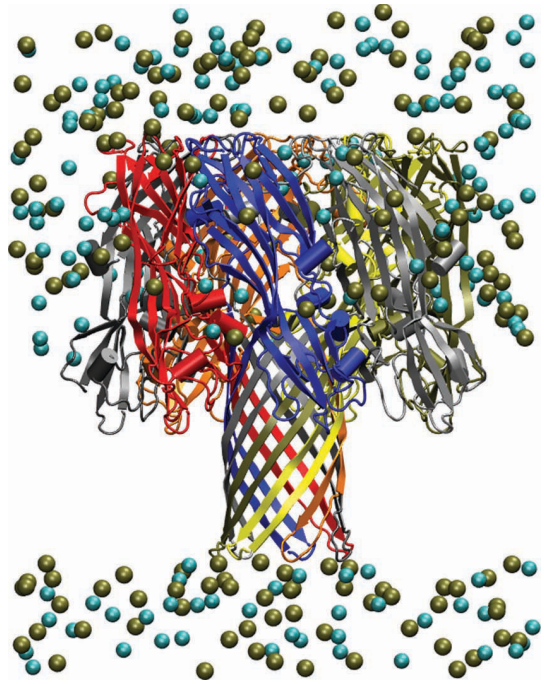


Figure 1. A visualization of the toxic protein α-Hemolysin, a heptamer (each identical monomer is colored differently). This snapshot, created using Visual Molecular Dynamics (VMD), shows randomly distributed ions (with 1 Molar salt concentration on each side) and the ion channel itself.

surrounding tissue enters.[9] α-HL has a mushroom-shaped heptamer about 100 Å in height and outer channel diameter, with pore diameter ranging from 15 to 46 Å.[8] Although α-HL's wild type (WT) channel shows weak anion selectivity,[10] mutations of this protein channel could alter its selectivity. Along with its wide aqueous pore and ability to insert small molecules (adaptors) in its lumen, α-HL has become an interesting protein channel for biosensor manufacturing.[11] Hagan Bayley and colleagues have used an engineered version of the channel with molecular adapters—such as β-cyclodextrins—inserted in its lumen to sense many molecule classes.[11] The two engineered covalent complexes—$(M113F)_6(M113C\text{-}D8RL2)_1$-$\beta$-CD and $(M113N)_6(T117C\text{-}D8RL3)_1$-$\beta$-CD—have been the subject of detailed study.

The size of the α-HL channel (more than 32,000 atoms) makes running full atomic molecular dynamics (MD) simulations extremely expensive, especially because simulating ion permeations requires relatively long trajectories on the timescale of hundreds of nanoseconds. Consequently, coarse-grained approaches to study ion permeation and selectivity through this channel are very appealing.

Methodology

Figure 1 shows the α-HL channel, along with a snapshot of ion trajectories generated by BioMOCA and visualized using Visual Molecular Dynamics (VMD).

BioMOCA Simulations

To compare computations properly with previous experimental results,[11] we set the pH to 8.0 to compute the protonation states of the mutant $(M113F)_6(M113C\text{-}D8RL2)_1$-$\beta$-CD. We selected default protein partial charges from Charmm Param 22 force fields, but modified them as indicated by pKa calculations. We used a Gaussian package to compute the partial charges of β-CD and the linking part, including the cysteine residue in α-HL.

The net charge on each mutant is $+7\,e$. We generated a $120 \times 120 \times 150$ mesh box with $\Delta = 1.0$ Å spacing to represent the system, along with other coarse/fine focused grids to compute the electrostatic potential efficiently. We assigned the three distinct regions—lipid, protein, and electrolyte (baths and channel pore)—dielectric coefficients of $\varepsilon = 5$, 2, and 80, respectively. We embedded the channels in a virtual membrane of 30 Å thickness and positioned the channel along the z-axis with the membrane's center at $z = 0$. To generate and prepare the structures, we used BioMOCA Suite, the GUI Web version of BioMOCA at nanoHUB.com.

At either end of the simulation box, two buffer regions maintain a constant ion density and extend inwards from the Dirichlet planes for a distance of 25 Å. The total ion scattering rate in bulk with bulk diffusivity for K^+ and Cl^- ions is

$$\lambda_{K^+} \cong 3.23 \times 10^{13} s^{-1},\ \lambda_{C^-1} \cong 3.46 \times 10^{13} s^{-1}.$$

Yet the high ion-water scattering rates necessitate a short-range ion trajectory integration time step (that is, $\Delta t = 10$ femtoseconds). However, the ion motion is highly damped, so the mesh nodes' assigned charge density changes relatively slowly with regard to the ion integration time step. Still, we might slightly relax the frequency of successive Poisson equation solutions, which account for most of the computational cost. For the simulations we present here, the Poisson solver runs every 100 time steps (that is, one picosecond in our simulations). To further accelerate Poisson equation solution, we implemented the grid-focusing multigrid method in BioMOCA. Although this complicates system setup, it's quite rewarding in reducing the computational cost.

Grid-Focusing Method

Ion channel simulations require large bath regions for accurate screening treatment. Such bath regions make the Poisson equation's mesh domain large and lead to either many grid points with fine mesh resolution or a few grid points with very coarse discretization. A coarse mesh is sufficient for describing baths using the P^3M scheme.[5] However, the channel domain requires a fine resolution because of the highly charged nature of these regions and the presence of spatially varying dielectric regions.

We developed a grid-focusing scheme in BioMOCA that lets us satisfy the requirements of both a large bath region and a fine grid resolution in channel in a computationally efficient way.[12] For the α-HL channel, we defined two discretized meshes. The first grid is a coarse mesh spanning the entire problem, with $40 \times 40 \times 50$ grid points and a uniform mesh spacing of 3 Å. The second (fine) grid has $30 \times 30 \times 60$ nodes and a mesh spacing of 1 Å, which we define along the channel's narrow selectivity section (see Figure 2).

First, we solve the Poisson equation on the coarse grid spanning the whole system using the finite difference method, and then use this solution to set Dirichlet boundary conditions on the fine mesh box using linear interpolation.[12] We then solve the Poisson equation again, but only on the fine grid to further refine the region's electrostatic potential values.

The computational cost of solving the Poisson equation using the preconditioned conjugate gradient method depends superlinearly—$(O(\sqrt{\kappa}N)$, with $(\kappa \geq 1)$—on the total number of grid points. To benchmark the speed up, we set up the same system with a single mesh of $120 \times 120 \times 150$ nodes and 1 Å mesh spacing. We did the benchmarking on a Xeon 5450 series processor. As Table 1 shows, the simulations are conveyed significantly faster—in this case, an order of magnitude faster—when using the grid-focusing method. Also, as we reduce the number of grids, the computational cost is reduced about 15 times.

Space-Dependent Diffusion Coefficient

It's widely accepted that the ions and water molecules don't have the same mobility or diffusivity in confined regions as they do in bulk. In fact, confined regions are more likely to lessen the ions' effective mobility in ion channels. In reduced particle methods that assume the channel water is an implicit continuum background, we need a mean ion mobility to reveal how ions might diffuse due to local electrostatic forces and random events. In transport Monte Carlo simulations, the total scattering rate (λ), is assumed to result only from ion-water interactions; it's related to ion diffusivity with

$$\lambda = \frac{kT}{mD}, \tag{3}$$

where m is the ion's mass and D is its diffusion constant. As the equation indicates, reduced diffusivity of ions inside the channel's lumen leads to an increase in scattering events.

Given α-HL's relatively wide pore size, we initially considered bulk diffusivity everywhere in the baths and channel—which, as expected, led to a higher ion flow than experiments predicted.[11] Next, we used the 1D wall-drag model, where the diffusion coefficient is dependent only on the

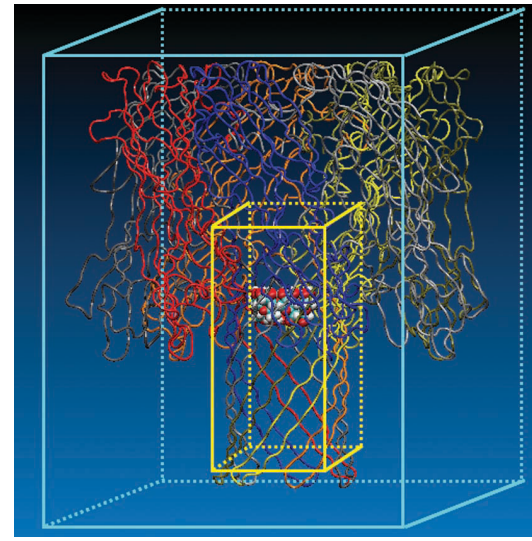


Figure 2. The focused grid scheme. The larger box (cyan) spans the entire region with the coarse mesh, while the smaller box (yellow) focuses on the channel's narrow part with fine meshing.

Table 1. Case study benchmarking with single grid and multigrid schemes.

	Meshes	Number of grid points	CPU-hour per 1 ns simulations
Multigrid scheme	$40 \times 40 \times 50$ + $30 \times 30 \times 60$	134,000	0: 23′: 30″
Single-grid scheme	$120 \times 120 \times 150$	2,160,000	5: 36′: 50″

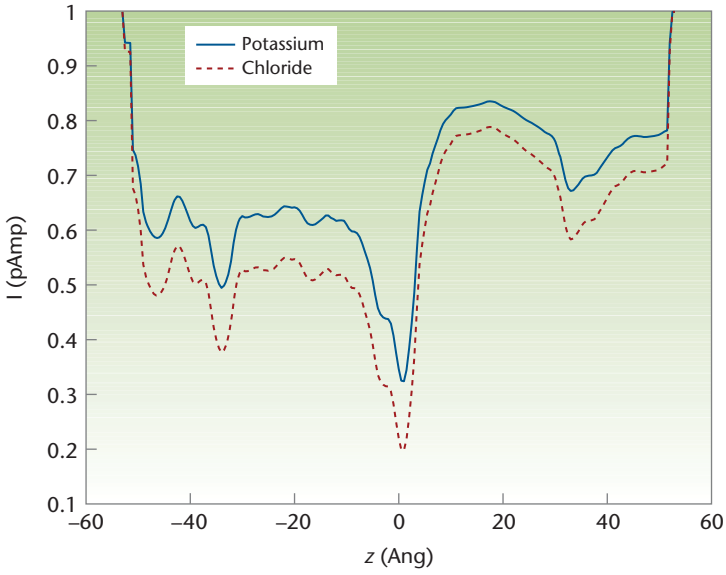


Figure 3. The relative diffusivity for each ion species along the α-HL channel. On the z-axis, 0 is at the protein center of mass. Compared to potassium, chloride has a smaller relative diffusion coefficient along the pore due its larger radius.

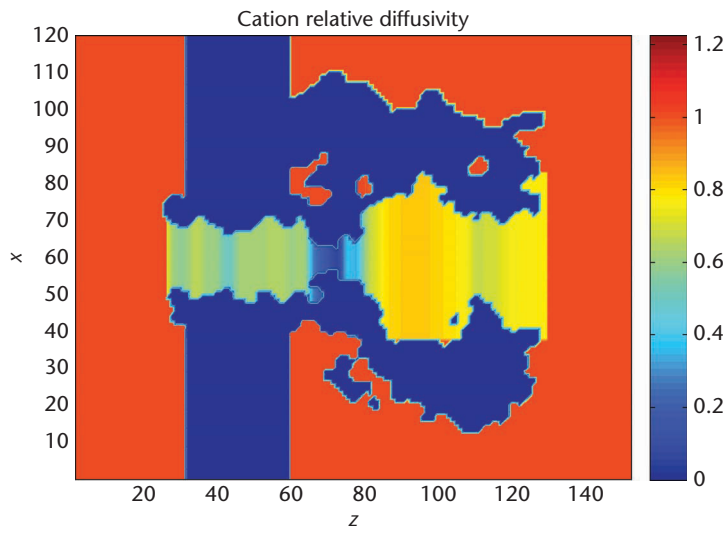


Figure 4. A 2D snapshot of modified diffusivity implementation inside the channel. The color red, which covers all regions outside the channel, has a value of 1, representing bulk diffusivity/mobility.

channel's geometric restriction.[7] We call the ratio of the diffusion coefficient inside the channel with respect to that of bulk *the scaling factor* (kz) and formulate it in an analytical expression, such as

$$k_z = \frac{1 - 2.1050\beta + 2.0865\beta^3 - 1.7068\beta^5 + 0.72603\beta^6}{1 - 0.75857\beta^5},$$

where the z direction is along the channel lumen and β (function of z) denotes the ratio of the ion species and the pore radius. The fractions here come from the best fit to the existing tabulated values.[7] The equation's derivation assumes

- uncharged hard-sphere mobile ions,
- uncharged cylindrical hard pore (channel),
- continuum solvent background (water), and
- 1D ion transport along the pore center axis (z-axis).

We experimentally determine the bulk diffusion coefficient itself in extremely diluted conditions. Figure 3 shows the scaling factor for both potassium and chloride ions along the mutant α-HL with β-cyclodextrin. Figure 4 shows how we implemented such a scaling factor inside the channel.

Results

We carried out two sets of simulations: one used bulk diffusivity for each ion species in all regions, and one used modified diffusivity inside the pore. For each bias point, five 100-nanosecond simulations were conveyed and averaged out, which is equivalent to running one longer 500 ns simulation per bias point.

Figure 5 shows the current-voltage (IV) curves for symmetric 1,000 mille-Molar potassium chloride (KCl) salt concentrations and compares the simulation results with the available experimental data.[11] Using bulk diffusivity inside the channel produces huge currents (dashed blue line) compared to experiments (solid green line). However, we achieve good agreements with experiments using the modified diffusivity (dashed red line), which seems to be diverging from the experiments in larger negative biases.

Although Monte Carlo-based approaches have limitations, they allow channel simulations on biologically related time scales, thus letting us study an ion channel's current flow and ion selectivity. The availability of such faster simulation approaches is valuable to making comparisons and predictions between different mutants before performing expensive experiments.

The results we obtained from BioMOCA are consistent with available experimental results. In general, we need a hierarchy of different approaches—such as Monte Carlo and MD—to study different aspects of ion channels and other biological systems.

Acknowledgments

The US National Institutes of Health Roadmap National Center for the Design of Biomimetic Nanoconductors supported this work under its Nanomedicine Development Center Program. nanoHUB also supported our work.

References

1. B. Hille, *Ionic Channels of Excitable Membranes*, Sinauer Assoc., 2001.
2. B. Eisenberg, "Ion Channels as Devices," *J. Computational Electronics*, vol. 2, nos. 2–4, 2003, pp. 245–249.
3. J.D. Watson et al., *Molecular Biology of the Gene*, 5th Ed., Benjamin/Cummings, 2003.
4. H. Bayley and P.S. Cremer, "Stochastic Sensors Inspired by Biology," *Nature*, vol. 4, no. 413, 2001, pp. 226–230.
5. T.A. van der Straaten, G. Kathawala, and U. Ravaioli, "Device Engineering Approaches to the Simulation of Charge Transport in Biological Ion Channels," *J. Computing and Theoretical Nanoscience*, vol. 3, no. 1, 2006, pp. 1–21.
6. R. Hockney and J. Eastwood, *Computer Simulation Using Particles*, McGraw-Hill, 1981.
7. P.L. Paine and P. Scherr, "Drag Coefficients for the Movement of Rigid Spheres through Liquid-Filled Cylindrical Pores," *Biophysical J*, vol. 15, no. 10, 1975, pp. 1087–1095.
8. L. Song et al., "Structure of Staphylococcal Alpha-Hemolysin, a Heptameric Transmembrane Pore," *Science*, vol. 274, no. 5294, 1996, pp. 1859–1869.
9. V. Guillet et al., "Crystal Structure of Leucotoxin S Component: New Insight into the Staphylococcal β-Barrel Pore-Forming Toxins," *J. Biological Chemistry*, vol. 279, 2004, pp. 41028–41037.
10. J.J. Kasianowicz and S.M. Bezrukov, "Protonation Dynamics of the Alpha-Toxin Ion Channel from Spectral Analysis of pH-Dependent Current Fluctutations, *Biophysical J.*, vol. 69, no. 1, 1995, pp. 94–105.
11. H. Wu et al., "Protein Nanopores with Covalently Attached Molecular Adapters," *Am. Chemical Soc.*, vol. 129, no. 51, 2007, pp. 16142–16148.
12. G. Kathawala and U. Ravaioli, *Monte Carlo Simulation of Nanostructures: Semiconductor Devices to Ion Channels*, doctoral dissertation, Dept. Electrical and Computer Eng., Univ. Illinois, Urbana-Champaign, 2005.

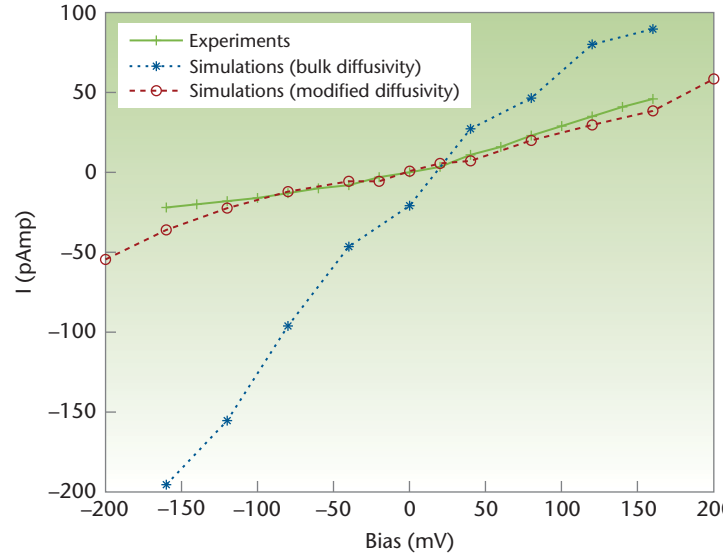


Figure 5. The current-voltage (IV) curves for the total currents. The solid green line is extracted from experimental data; the dashed red is for simulations using the modified diffusivity; and the dashed blue line is for simulations with bulk diffusivity in all regions including inside the channel. There's good agreement between the experiments and simulations with modified diffusivity of ions, while the current overshoots strongly with bulk diffusivity.

Reza Toghraee *is a doctoral student at the University of Illinois at Urbana-Champaign, where he's a researcher in the Computational Multiscale Nanosystems Group. His research interests include studying conduction phenomena in ion channels using Monte Carlo methods. Toghraee has an MS in microwave and optoelectronics communications from Sharif University of Technology, Tehran. Contact him at rtoghra2@illinois.edu.*

Umberto Ravaioli *is a professor in the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign, where he is senior assistant dean of Undergraduate Programs in the College of Engineering. His research interests include particle simulation, with an emphasis on Monte Carlo techniques and applications from ultrascaled MOSFETs to electrothermal simulation and charge transport in biological systems. Ravaioli has a PhD in electrical engineering from Arizona State University. He is a fellow of the IEEE and the Institute of Physics. Contact him at ravaioli@illinois.edu.*

Kyu Il Lee *is a postdoctoral researcher at the Center for Bioinformatics, University of Kansas. He was previously a postdoctoral research associate at Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign. His research interests include physical modeling and simulation of ion channel transport. Lee has a PhD in electrical engineering from Seoul National University. Contact him at macjr@ku.edu.*

HUBzero: A Platform for Dissemination and Collaboration in Computational Science and Engineering

The HUBzero cyberinfrastructure lets scientific researchers work together online to develop simulation and modeling tools. Other researchers can then access the resulting tools using an ordinary Web browser and launch simulation runs on the national Grid infrastructure, without having to download or compile any code.

What if researchers could access and share scientific simulation and modeling tools as easily as YouTube.com videos? That's the underlying premise for the HUBzero Platform for Scientific Collaboration, a cyberinfrastructure developed at Purdue University. We created HUBzero to support nanoHUB.org, an online community for the Network for Computational Nanotechnology, which the US National Science Foundation has funded since 2002 to connect the theorists who develop simulation tools with the experimentalists and educators who might use them. Since 2007, HUBzero use has expanded to support a growing number of hubs for pharmaceutical engineering (pharmaHUB.org), heat transfer (thermalHUB.org), microelectromechanical systems (memsHUB.org), healthcare (IndianaCTSI.org), cancer care (cceHUB.org), and engineering education (globalHUB.org).

1521-9615/10/$26.00 © 2010 IEEE
Copublished by the IEEE CS and the AIP

Michael McLennan and Rick Kennell
Purdue University

Although they serve different communities, the hubs all support collaborative development and dissemination of scientific models running in an infrastructure that leverages a "cloud" of computing resources. As other articles in this issue show, such an infrastructure can have an impact on scientific discovery. (Other work has shown this impact in relation to education and the community at large.[1–2]) Here, we describe the infrastructure itself and provide an overview of the model publication process.

Middleware for Hosted Execution

Each tool published on a hub looks much like a journal paper, with a title, a list of authors, an abstract, and even a list of references, along with instructions on how to cite the resource as a reference in other publications. In addition to the overview information, each tool page contains a prominent button to launch a live session. Clicking the button brings up an interactive GUI for the tool right within the browser (see Figure 1).

On the surface, the simulation tools look like simple Java applets embedded within the browser, but they're actually running on a cluster of execution hosts housed near the Web server and projected to the user's browser using virtual network

computing (VNC).[3] Each tool runs in a restricted lightweight virtual environment implemented using OpenVZ (www.openvz.org), which carefully controls access to file systems, networking, and other server processes. Each user has a unique home directory with conventional ownership, access controls, and quota limitations. Tools are run with the rights and privileges of the particular user rather than a shared execution account. This lets the hub precisely regulate and account for each tool's resource usage.

Tools running on the hub expect a typical X11 Window System environment. This graphical session is created by running, in each container, a special X server that also acts as a VNC server. Because it takes several seconds to reliably create a container and initialize the X server, the hub prestarts a pool of containers and X servers that await user requests for new tools. When a user requests a tool, the hub selects an eligible container, starts the tool, and sends the user's browser instructions on how to connect to it. This typically reduces the perceived tool startup time to about two seconds.

The HUBzero middleware controls the tool container's network operations. For example, it authenticates and routes incoming VNC viewer and file transfer requests from browsers to the correct container. The middleware also monitors the start time and duration of each connection for accounting purposes. For example, a tool continues to run even while not being viewed, which lets users close their browsers and later return to their tools, accessing them through their My Hub pages. Any tool that hasn't had a viewer connection after a configurable duration (typically 24 hours) is marked inactive and terminated.

Outgoing connections from each tool container to other resources are limited to only locally trusted services. Because tools run in a controlled environment on the hub's execution cluster rather than the user's desktop, they can be more easily authorized for secure access to high-performance visualization facilities as well as execution on remote resources. Tool containers, for example, can be configured to route jobs through a "submit" server that acts as a secure proxy through which the hub can direct jobs to national Grid resources such as the TeraGrid,[4] Open Science Grid,[5] and Purdue's DiaGrid.[6]

This system delivers substantial computing power to thousands of end users without requiring that they log into a head node or fuss with any proxy certificates. In 2008, approximately 6,700 users launched more than 380,000 simulation jobs using 129 tools available on nanoHUB.org. Figure 2 shows the amount of use for this diverse

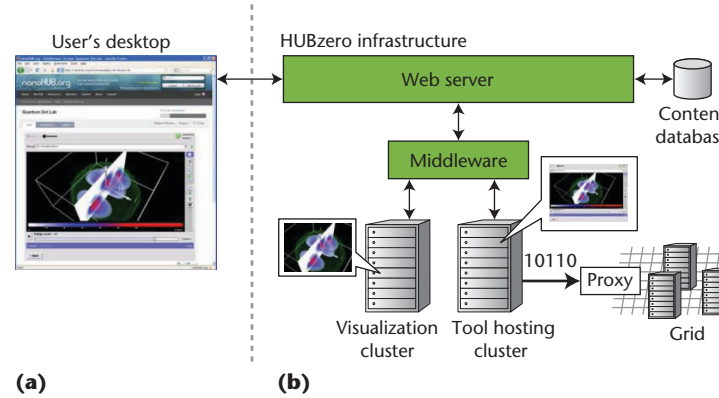


Figure 1. HUBzero's architecture. (a) Users access interactive, graphical tools via an applet in their Web browser. (b) The tools run on a cluster at Purdue University, where jobs can be dispatched to more powerful computers in the national Grid infrastructure.

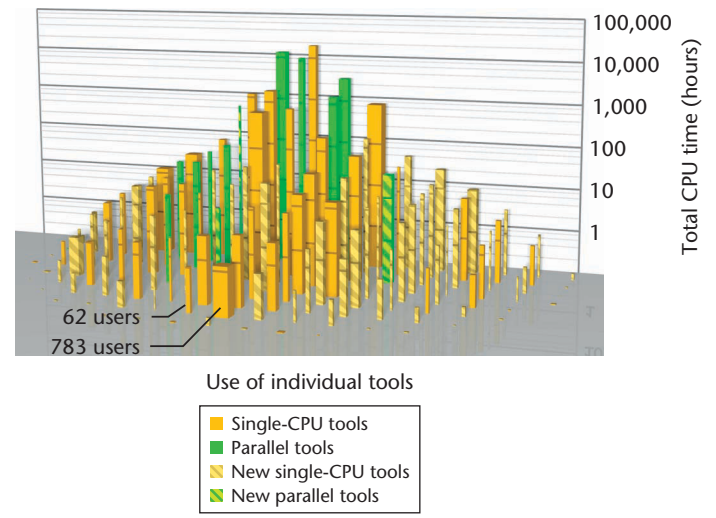


Figure 2. Total CPU time for all nanoHUB.org tools in 2008. Each bar represents one of the 129 tools. All together, the tools consumed a total of 56,930 CPU-hours. Most of the tools were accessed by hundreds of users, which combined, represent a community of 6,700 unique users. Some of the tools were run hundreds of times, on average, by each user and delivered thousands of CPU hours of computation. Orange bars are tools that run on a single CPU, and green bars are parallelized tools that use multiple CPUs for each run. The striped bars represent tools introduced midyear, which therefore have lower total CPU time.

set of tools. Each bar represents one tool, with a base proportional to the total number of users and a height indicating the total CPU time for all runs in 2008.

As Figure 2 shows, even the most computationally intensive, parallelized tools have hundreds of users. Other tools execute in only a fraction of second, and yet—with 783 users and 14,628 simulation runs, for example—add up to more than an hour of total CPU time for the entire year. More than half of the bars (indicated with a striped

pattern) are tools that were introduced during 2008; because they have less than a full year of use, each has lower amounts of total CPU time. One such tool, online for only two weeks, still managed to attract 16 users and accumulate more than 700 CPU hours of computation.

Tool Development and Deployment

The tools on each hub come not from the core development team but from hundreds of other researchers scattered throughout the world. HUBzero supports the workflow for all of these developers and has a content management system for tool publication. Developers enter the system by filling out a Web form to register their tools. They then receive a software project area similar to that on SourceForge.net, complete with a Subversion source code repository and a wiki area for project notes.

At that point, all development team members receive access to a special tool that we call a "workspace," which is a Linux desktop running in our secure execution environment, accessed via a Web browser (like any other hub tool). Tool developers launch a workspace; check out their project's directory skeleton; and then begin editing files, compiling their code, and testing their tool. Each workspace runs in the same execution environment as the published tools, so developers can access the same visualization servers and Grid resources for testing.

Once the tool is working properly and changes have been committed, the developers click the "My code has been uploaded" link on their tool status page and the tool is staged for final testing. Once the tool has been tested, developers either approve or update the tool by clicking one of two links:

- "My tool is working properly. I approve it."
- "I've fixed my code. Please install the latest updates."

When approving a tool for publication, developers can specify an open source license, in which case the hub automatically publishes an archive file containing the release's source code. If they choose to keep their source code private, other users can still access the live tool via the hub, but the source code bundle isn't published.

Our experience with nanoHUB.org has shown that HUBzero can scale to support hundreds of independent tool development teams, each publishing, modifying, and republishing their tool dozens of times per year. All of this can be supported by a single part-time person on the hub support team to help with tool compilation, installation, and staging.

User Interfaces for Simulation and Visualization

If a tool already has a GUI that runs under Linux, it can be deployed as-is in a matter of hours. If not, tool developers can use HUBzero's Rappture toolkit to create a GUI with little effort.[7] Rappture reads an XML description of the tool's inputs and outputs and then automatically generates a GUI.

Figure 3 shows an example of the XML description for a simple tool that takes two numbers and computes a curve representing the Fermi-Dirac distribution as a function of energy. The `<input>` section describes each number, including the physical units, minimum and maximum allowed values, and default value. The `<output>` section describes the resulting `<curve>` object. The Rappture library supports approximately two dozen objects—including numbers, Boolean values, curves, meshes, scalar/vector fields, and molecules—which can be used to represent each tool's inputs and outputs.

The XML file's `<tool>` section describes the code for the tool itself. When the user clicks the GUI's Simulate button, Rappture will execute the command specified in the `<tool>` section, feeding it an XML description of the input values and receiving from it an XML description of the overall run. The example in Figure 3 launches a Python script `fermi.py`. The input and output values are accessed within the Python script via an API. Rappture also supports an API for other languages, such as C/C++, Fortran, Matlab, Perl, Ruby, and Tcl, so it can accommodate various modeling codes.

The results from each run are loaded back into the GUI and displayed in a specialized viewer created for each output type. Viewers for molecules, scalar and vector fields, and other complex types automatically connect to servers in our visualization cluster and deliver hardware-accelerated 3D data views.[8] Results from multiple runs can be overlaid and compared within each viewer.

Collaboration and Support

HUBzero sites also provide ways for colleagues to work together. For example, because of the unique way the HUBzero middleware hosts tool sessions, a single session can be shared among any number of people. A box beneath each tool session lets the user enter login names for one or more colleagues;

```
<?xml version="1.0"?>
<run>
  <tool>
    <about>Computes Fermi-Dirac distributions.</about>
    <command>python @tool/fermi.py @driver</command>
  </tool>
  <input>
    <number id="temperature">
      <about> <label>Ambient temperature</label> </about>
      <units>K</units>
      <min>0K</min> <max>500K</max>
      <default>300K</default>
    </number>
    <number id="Ef">
      <about> <label>Fermi Level</label> </about>
      <units>eV</units>
      <default>0eV</default>
    </number>
  </input>
  <output>
    <curve id="f12">
      <about> <label>Fermi-Dirac Factor</label> </about>
      <xaxis> <label>Fermi-Dirac Factor</label> </xaxis>
      <yaxis> <label>Energy</label> <units>eV</units> </yaxis>
    </curve>
  </output>
</run>
```

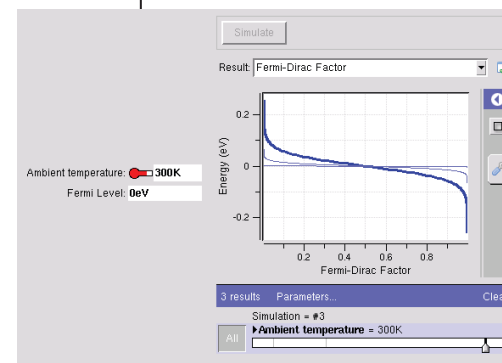


Figure 3. An example XML tool description. The Rappture toolkit generates a GUI based on an XML description for a tool's inputs and outputs. In this case, the `<input>` section describes two `<number>` parameters, and the output section describes the resulting `<curve>`.

clicking the Share button grants them instant session access. A group of people can look at the same session at the same time and discuss ideas over the phone or instant messaging. If some of the people aren't online or available, they can access the session later from their My Hub page and follow up at their convenience.

As people are using the tools, questions arise and sometimes things go wrong. HUBzero supports many ways for users to find help and help one another. Users can post questions in a community forum modeled after Amazon.com's Askville or Yahoo! Answers. Users can tag questions with tool names and other important concepts so that other users can find questions matching their interests and post answers. In doing so, they earn reward points that can be used to purchase t-shirts and other hub store merchandise, post a bounty for their own questions, or purchase other premium hub services.

When something goes wrong, users can file trouble reports by clicking on the Help! button in the upper-right corner of any page. When a user submits a trouble ticket from the page containing a tool, it's sent directly to the tool's development team. Members of the development team will see the trouble report on the My Tickets list on their own customized My Hub pages. Tickets can be updated with comments, screen shots, and other attachments, and they can be transferred to other development teams (including the HUBzero core team). All parties involved with the ticket are updated as the ticket is investigated and resolved.

In practice, we've found that many tickets aren't really problems but are actually requests for new features. HUBzero supports a wish list capability for collecting, prioritizing, and acting on such requests. Users can post an idea to the wish list associated with each tool or to the general list associated with the hub itself. If users feel strongly about a wish, they can offer a bounty of reward points, thereby giving the development team incentive to implement the change. Other users can provide their input by giving a "thumbs up" or "thumbs down" to the idea or by adding their own reward points to the bounty.

When each development team views a wish on their list, they see additional controls for gauging both the importance (from rubbish to crucial) and the effort (from a few hours to several months). The consensus from the team determines a wish's relative priority; those judged important and requiring little effort appear at the top of the list, while those that are unimportant and

time-consuming go to the bottom. The development team can also add threaded comments to the wish, develop an implementation plan, and assign the wish to a particular team member. Once that team member grants the wish, he or she earns 80 percent of the bounty; the remaining 20 percent is divided among other team members that offered input. This system helps each hub team organize and prioritize its development tasks.

HUBzero's unique blend of simulation power and social networking seems to resonate across engineering and science communities. As hub use continues to grow, we plan to develop new capabilities to connect related content so that tools published on one hub can be easily found on all others. We also plan to improve tool interconnection, so that one tool's output can be used as input to another, letting developers solve larger problems by connecting a series of models from independent authors. Along the way, we hope to change the practice of simulation and modeling so that

- simulation experiments are as carefully controlled and reproducible as laboratory experiments, and
- the publication of live tools confers as much professional recognition as the publication of journal papers about the underlying science.

In so doing, we hope to elevate computing to become the true third pillar of science.[9]

Acknowledgments

HUBzero is a trademark of the Purdue Research Foundation. HUBzero's development has been supported by the US National Science Foundation through awards EEC-0228390, EEC-0634750, OCI-0438246, and OCI-0721680 and by Purdue University. We're grateful for the hard work and dedication of our development team: Chris Camisa, Joe Cychosz, Ann Christine Catlin, Steve Clark, Ben Haley, George Howlett, Derrick Kearney, Nick Kisseberth, Joshua Morris, Alissa Nedossekina, Shawn Rice, John Rosheck, Beth Schroeder, Swaroop Shivarajapura, Steve Snyder, and Jeff Turkstra. We also appreciate thoughtful comments from Mark Lundstrom and Gerhard Klimeck during the article's preparation.

References

1. G. Klimeck et al., "nanoHUB.org: Advancing Education and Research in Nanotechnology," *Computing in Science & Eng.,* vol. 10, no. 5, 2008, pp. 17–23.
2. N. Wilkins-Diehr et al., "TeraGrid Science Gateways and Their Impact on Science," *Computer,* vol. 41, no. 11, 2008, pp. 32–41.
3. T. Richardson et al., "Virtual Network Computing," *IEEE Internet Computing,* vol. 2, no. 1, 1998, pp. 33–38.
4. D.A. Reed, "Grids, the TeraGrid, and Beyond," *Computer,* vol. 36, no. 1, 2003, pp. 62–68.
5. R. Pordes et al., "The Open Science Grid," *J. Phys Conf. Series,* vol. 78, no. 1, 2007; www.iop.org/EJ/article/1742-6596/78/1/012057/jpconf7_78_012057.pdf.
6. P.M. Smith, T.J. Hacker, and C.X. Song, "Implementing an Industrial-Strength Academic Cyberinfrastructure at Purdue University," *Proc. IEEE Int'l Conf. Parallel and Distributed Processing* (IPDPS 2008), IEEE CS Press, 2008, pp. 1–7.
7. M. McLennan, "Introducing the Rappture Toolkit," HUBzero, 21 June 2009; http://hubzero.org/resources/45.
8. W. Qiao et al., "Hub-Based Simulation and Graphics Hardware Accelerated Visualization For Nanotechnology Applications," *IEEE Trans. Visualization and Computer Graphics,* vol. 12, no. 5, 2006, pp. 1061–1068.
9. M. Lundstrom and G. Klimeck, "The NCN: Science, Simulation, and Cyber Services," *IEEE Conf. Emerging Technologies—Nanoelectronics,* IEEE CS Press, 2006, pp. 496–500.

Michael McLennan *is a senior research scientist at Purdue University's Rosen Center for Advanced Computing, where he is director of the HUBzero Platform for Scientific Collaboration. He developed [incr Tcl], an object-oriented extension of Tcl, and is coauthor of two books,* Effective Tcl/Tk Programming *(Addison-Wesley, 1997) and* Tcl/Tk Tools *(O'Reilly Media, 1997). McLennan has a PhD in electrical engineering from Purdue University, where he was an SRC Graduate Fellow. He is a member of IEEE, Eta Kappa Nu, and Tau Beta Pi. Contact him at mmclennan@purdue.edu.*

Rick Kennell *is a senior software engineer at Purdue University's Rosen Center for Advanced Computing. He developed the middleware used to host virtual tool sessions and simulations for HUBzero.org and its derivative hubs. Contact him at kennell@purdue.edu.*

POWER YOUR CONNECTIONS

HUB

Simulation tools

As close as your web browser

For research and education

Collaborate with thousands of users worldwide

Create your own HUB

Available as Open Source at

HUBzero.org

Supported by the HUBzero™ Consortium

PURDUE UNIVERSITY

INDIANA UNIVERSITY

THE UNIVERSITY of WISCONSIN

CLEMSON UNIVERSITY