9-1-2007

# Atomistic simulation of realistically sized nanodevices using NEMO 3-D - Part I: Models and benchmarks

Gerhard Klimeck
*Purdue University*, gekco@purdue.edu

Shaikh S. Ahmed
*Purdue University*

Hansang Bae
*Purdue University*, baeh@purdue.edu

Neerav Kharche
*Birck Nanotechnology Center and Purdue University*, nkharche@purdue.edu

Rajib Rahman
*Purdue Univ, Sch Elect & Comp Engn*, rrahman@purdue.edu

***See next page for additional authors***

**Authors**
Gerhard Klimeck, Shaikh S. Ahmed, Hansang Bae, Neerav Kharche, Rajib Rahman, Steve Clark, Benjamin Haley, Sunhee Lee, Maxim Naumov, Hoon Ryu, Faisal Saied, Marta Prada, Marek Korkusinski, and Timothy Boykin

# Atomistic Simulation of Realistically Sized Nanodevices Using NEMO 3-D—Part I: Models and Benchmarks

Gerhard Klimeck, *Senior Member, IEEE*, Shaikh Shahid Ahmed, Hansang Bae, Neerav Kharche, Rajib Rahman, Steve Clark, Benjamin Haley, Sunhee Lee, Maxim Naumov, Hoon Ryu, Faisal Saied, Marta Prada, Marek Korkusinski, and Timothy B. Boykin, *Senior Member, IEEE*

*(Invited Paper)*

***Abstract*—Device physics and material science meet at the atomic scale of novel nanostructured semiconductors, and the distinction between new device or new material is blurred. Not only the quantum–mechanical effects in the electronic states of the device but also the granular atomistic representation of the underlying material are important. Approaches based on a continuum representation of the underlying material typically used by device engineers and physicists become invalid. *Ab initio* methods used by material scientists typically do not represent the band gaps and masses precisely enough for device design, or they do not scale to realistically large device sizes. The plethora of geometry, material, and doping configurations in semiconductor devices at the nanoscale suggests that a general nanoelectronic modeling tool is needed. The 3-D NanoElectronic MOdeling (NEMO 3-D) tool has been developed to address these needs. Based on the atomistic valence force field and a variety of nearest neighbor tight-binding models (e.g., $s$, $sp^3s^*$, and $sp^3d^5s^*$), NEMO 3-D enables the computation of strain and electronic structure for more than 64 and 52 million atoms, corresponding to volumes of $(110\ \text{nm})^3$ and $(101\ \text{nm})^3$, respectively. The physical problem may involve very large scale computations, and NEMO 3-D has been designed and optimized to be scalable from single central processing units to large numbers of processors on commodity clusters and supercomputers. NEMO 3-D has been released with an open-source license in 2003 and is continually developed by the Network for Computational Nanotechnology (NCN). A web-based online interactive version for educational purposes is freely available on the NCN portal (http://www.nanoHUB.org). In this paper, theoretical models and essential algorithmic and computational components that have been used in the development and successful deployment of NEMO 3-D are discussed.**

***Index Terms*—Atomistic simulation, Keating model, nanostructures, piezoelectricity, quantum computation, strain, tight binding, valley splitting, 3-D NanoElectronic MOdeling (NEMO 3-D).**

Manuscript received January 3, 2007; revised May 16, 2007. This work was supported in part by the Indiana 21st Century Fund, by the Army Research Office, by the Office of Naval Research, by the Semiconductor Research Corporation, and by the National Science Foundation under Grant EEC-0228390. The review of this paper was arranged by Editor A. Asenou.

G. Klimeck is with the School of Electrical and Computer Engineering and Network for Computational Nanotechnology, Purdue University, West Lafayette, IN 47907 USA, and also with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA (e-mail: gekco@purdue.edu).

S. S. Ahmed, H. Bae, N. Kharche, R. Rahman, B. Haley, S. Lee, H. Ryu, and M. Prada are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907-2035 USA, and also with the Network for Computational Nanotechnology, Purdue University, West Lafayette, IN 47907-2057 USA.

S. Clark and F. Saied are with the Rosen Center for Advanced Computing, Purdue University, West Lafayette, IN 47907 USA.

M. Naumov is with the Department of Computer Science, Purdue University, West Lafayette, IN 47907-2107 USA.

M. Korkusinski is with the Quantum Theory Group, Institute for Microstructural Sciences, National Research Council of Canada, Ottawa, ON K1A 0R6, Canada.

T. B. Boykin is with the Department of Electrical and Computer Engineering, University of Alabama in Huntsville, Huntsville, AL 35899 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2007.902879

## I. Introduction

### A. Emergence of Nanodevices

THE RAPID progress in nanofabrication technologies has led to the emergence of new classes of nanodevices and structures, which are expected to bring about fundamental and revolutionary changes in electronic, photonic, biotechnology, information processing and computation, and medicine industries. These devices demonstrate new capabilities and functionalities, where the *quantum nature* of charge carriers plays an important role in determining the overall device properties and performance. The device sizes have already reached the level of tens of nanometers. In this regime, the *atomistic granularity* of constituent materials cannot be neglected: Effects of atomistic strain, surface roughness, unintentional doping, the underlying crystal symmetries, or distortions of the crystal lattice can have a dramatic impact on the device operation and performance. In effect, the formerly disjoint fields of semiconductor devices and materials science meet at the atomic scale.

A critical facet of the nanodevice development is the creation of simulation tools that can quantitatively explain or even predict experiments. In particular, it would be very desirable to explore the design space before or in conjunction with the (typically time-consuming and expensive) experiments. A general tool that is applicable over a large set of materials and geometries is highly desirable. However, just the tool development itself is not enough. The tool needs to be deployed to the user community so that it can be made more reliable, flexible, and accurate. The main goal of this paper is

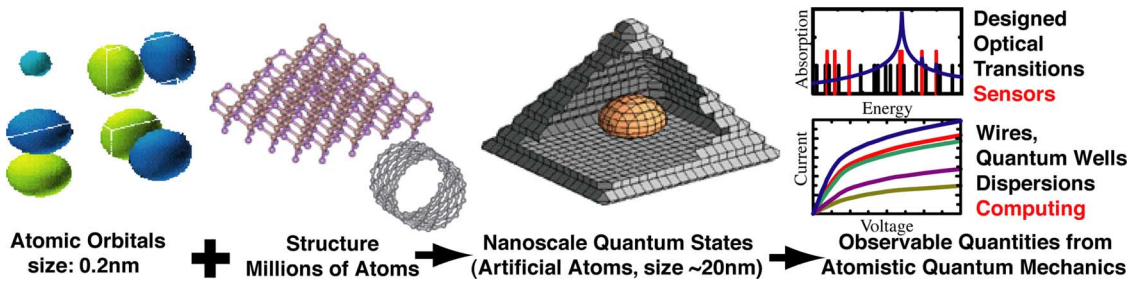



Fig. 1. NEMO 3-D modeling agenda. Map electronic properties of individual atoms into realistic structures containing millions of atoms, and computation of nanoscale QDs that map into real applications.

to describe the theoretical models and the essential algorithmic and computational components that have been used in the development and successful deployment of the 3-D NanoElectronic MOdeling (NEMO 3-D) on http://www.nanoHUB.org. We present some of the new capabilities that have been recently added to NEMO 3-D to make it one of the premier simulation tools for design and analysis of realistically sized nanoelectronic devices and, therefore, to make it a valid tool for the computational nanotechnology community. These recent advances include algorithmic refinements, performance analysis to identify the best computational strategies, and memory-saving measures. The effective scalability of NEMO 3-D code on the BlueGene, an Intel Woodcrest cluster, the Cray XT3, and other Linux clusters is demonstrated. The largest electronic structure calculation with 52 million atoms involved a Hamiltonian matrix over one billion complex degrees of freedom. A comparison is made between the performance of a stored Hamiltonian and the recomputation of the matrix each time it is needed. Through a set of end-to-end calculations, it is shown how the eigenvalues vary as a function of the size of the domain. We describe the state-of-the-art algorithms that have been incorporated in the code, including a very effective Lanczos eigenvalue solver, and present a comparison of the different solvers. While such system sizes of tens of millions of atoms appear huge and wasteful at first sight, we claim here that some physical problems require such large-scale analysis. We recently demonstrated [1] that the analysis of valley splitting in strained Si quantum wells grown on strained SiGe required atomistic analysis of ten million atoms to match the experimental data. The insight that the disorder in the SiGe buffer increases valley splitting in the Si quantum well would probably not be predictable in a continuum effective mass model.

## II. Modeling and Simulation Challenges

The theoretical knowledge of the electronic structure of nanoscale semiconductor devices is the first and essential step toward the interpretation and understanding of the experimental data and reliable device design at the nanometer scale. The following is a list of the modeling and simulation challenges in the design and analysis of realistically sized engineered nanodevices.

### A. Full 3-D Atomistic Representation

The lack of *spatial symmetry* in the overall geometry of the nanodevices usually requires explicit 3-D representation. For example, Stranski–Krastanov growth techniques tend to produce self-assembled InGaAs/GaAs quantum dots (QDs) [2]–[5] with cylindrical-like shape symmetry, e.g., disks, truncated cones, domes, or pyramids [6]. These geometries are generally not perfect geometric objects since they are subject to interface interdiffusion and discretization on an atomic lattice. There is no such thing as a round disk on a crystal lattice. The underlying crystal symmetry imposes immediate restrictions on the realistic geometry and influences the quantum mechanics. Continuum methods such as effective mass [7] and $\mathbf{k}\cdot\mathbf{p}$ [8], [9] typically ignore such crystal symmetry and atomistic resolution. The required simulation domain sizes of about one million atoms prevent the usage of *ab initio* methods. Empirical methods that eliminate enough unnecessary details of core electrons but are finely tuned to describe the atomistically dependent behavior of valence and conduction electrons are needed. The current state-of-the-art leaves two choices: 1) pseudopotentials [10] and 2) tight binding [11]. Both methods have their advantages and disadvantages. Pseudopotentials use plane waves as a fundamental basis choice. Realistic nanostructures contain high-frequency features such as alloy disorder or heterointerfaces. This means that the basis needs to be adjusted (by an expert) for every different device, which limits the potential impact for nonexpert users. Numerical implementations of pseudopotential calculations typically require a Fourier transform between real and momentum space, which demand full matrix manipulations and full transposes. This typically requires high-bandwidth communication capability (i.e., extremely expensive) parallel machines, which limit the practical dissemination of the software to end users with limited compute resources. Tight binding is a local basis representation, which naturally deals with finite device sizes, alloy disorder, and heterointerfaces, and it results in very sparse matrices. The requirements of storage and processor communication are therefore minimal as compared to pseudopotentials, and actual implementations perform extremely well on cheap clusters [11]. Tight binding has the disadvantage that it is based on empirical fitting, and the community continues to raise the issue on the fundamental applicability of tight binding. The NEMO team has spent a significant effort to expand and document the tight-binding capabilities with respect to handling of strain [12], electromagnetic fields [13], and Coulomb matrix elements [14] and fit them to well-known and accepted bulk parameters [11], [15], [16]. With tight binding, the NEMO team was able early on to match experimentally verified high-bias current–voltage curves of resonant tunneling [17], [18]

that could not get modeled by either effective mass (due to the lack of physics) or pseudopotential methods (due to the lack of open-boundary conditions). We continue to learn about the tight-binding method capabilities and are in the process of benchmarking it against more fundamental *ab-initio* approaches and pseudopotential approaches. Our current Si/Ge parameterization is described in [19] and [20]. Fig. 1 depicts a range of phenomena that represent new challenges presented by new trends in nanoelectronics and lays out the NEMO 3-D modeling agenda.

### B. Atomistic Strain

Strain that originates from the assembly of lattice-mismatched semiconductors strongly modifies the energy spectrum of the system. In the case of the InAs/GaAs QDs, this mismatch is around 7% and leads to a strong *long-range* strain field within and is very wide reaching (typically $\sim$25 nm) around each QD [21]. Si/Ge core/shell structured nanowires are another example of strain-dominated atom arrangements [22], and Si-based quantum-well quantum-computing architectures rely on strain for state separation [23]. The strain can be atomistically inhomogeneous, involving not only biaxial components but also nonnegligible shear components. Strain strongly influences the core and barrier material band structures, modifies the energy bandgaps, and lifts the heavy hole–light hole degeneracy at the zone center. In the nanoscale regime, the classical harmonic linear/continuum elasticity model for strain is inadequate, and device simulations must include the fundamental quantum character of charge carriers and the long-distance atomistic strain effects with proper boundary conditions on equal footing [24], [25].

### C. Piezoelectric Field

A variety of advanced materials that are of interest is piezoelectric, such as GaAs, InAs, and GaN. Any spatial distortions in nanostructures made of these materials will create significant piezoelectric fields, which will significantly modify the electrostatic potential landscape. Recent spectroscopic analyses of self-assembled QDs demonstrate polarized transitions between confined hole and electron levels [6]. While the continuum models (effective mass or $\mathbf{k}\cdot\mathbf{p}$) can reliably predict aspects of the single-particle energy states, they fail to capture the observed nondegeneracy and optical polarization anisotropy of the excited energy states in the (001) plane. These methods fail because they use a confinement potential that is assumed to have only the *shape symmetry* of the nanostructure and they ignore the underlying crystal symmetry. However, experimentally noticeable is the fact that the true symmetry is lower than the assumed continuum symmetry because of the following: 1) underlying crystalline symmetry; 2) atomistic strain relaxation; and 3) piezoelectric field. For example, in the case of pyramid-shaped QDs with square bases, continuum models treat the underlying material in $C_{4\nu}$ symmetry, whereas the atomistic representation lowers the crystal symmetry to $C_{2\nu}$. Piezoelectric potential originating from the nonzero shear component of the strain field must be taken into account to properly model the associated symmetry breaking and the introduction of a global shift in the energy spectra of the system.

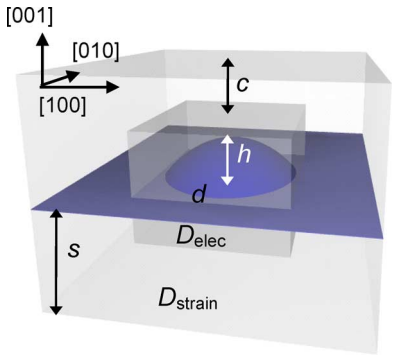



Fig. 2. Simulated dome-shaped InAs/GaAs QD. Two simulation domains are shown. $D_{\text{elec}}$ is the central domain for electronic structure calculation, and $D_{\text{strain}}$ is the larger/outer domain for strain calculation. $s$ is the substrate height, $c$ is the cap layer thickness, $h$ is the dot height, and $d$ is the dot diameter.

## III. NEMO 3-D Simulation Package

### A. Basic Features

NEMO 3-D [11], [26]–[29] bridges the gap between the large-size classical semiconductor device models and the molecular-level modeling. This package currently allows calculating the single-particle electronic states and optical response of various semiconductor structures, including bulk materials, QDs, quantum wires, quantum wells, and nanocrystals. NEMO 3-D includes spin in its fundamental atomistic tight-binding representation. Spin is therefore not added in as an afterthought into the theory, but spin–spin interactions are naturally included in the Hamiltonian. Effects of interaction with external electromagnetic fields are also included [11], [30], [13]. This paper focuses on the design and performance of NEMO 3-D illustrated in the case of InAs QDs embedded in a GaAs barrier material. A schematic view of the sample is presented in Fig. 2. The QD is positioned on a 0.6-nm-thick wetting layer (dark region). The simulation of strain is carried out in the large computational box $D_{\text{strain}}$, whereas the electronic structure computation is restricted to the smaller domain $D_{\text{elec}}$. In Part II of this paper, it has been shown that under the assumptions of realistic boundary conditions, strain is long ranged and penetrates around 25 nm into the dot substrate, thus stressing the need for using large substrate thickness in the simulations. NEMO 3-D enables the computation of strain and electronic structure in an atomistic basis for more than 64 and 52 million atoms, corresponding to volumes of $(110\text{ nm})^3$ and $(101\text{ nm})^3$, respectively. These volumes can be spread out arbitrarily over thin-layer geometry. For example, if a thin layer of 15-nm height is considered, the corresponding widths in the $x$–$y$ plane correspond to 298 nm for strain calculations and 262 nm for electronic structure calculations. No other atomistic tool can currently handle such volumes needed for realistic device simulations. NEMO 3-D runs on serial and parallel platforms, local cluster computers, as well as the National Science Foundation (NSF) Teragrid.

### B. Components and Models

The NEMO 3-D program flow consists of four main components.

*1) Geometry Construction:* The first part is the geometry constructor, whose purpose is to represent the treated nanostructure in atomistic detail in the memory of the computer. Each atom is assigned three single-precision numbers representing its coordinates and also stored are its type (atomic number in short integer), the information whether the atom is on the surface or in the interior of the sample (important later on in electronic calculations), what kind of computation it will take part of (strain only or strain and electronic), and what its nearest neighbor relation in a unit cell is. The arrays holding this structural information are initialized for all atoms on all central processing units (CPUs); that is, the complete information on the structure is available on each CPU. By default, most of this information can be stored in short integer arrays or as single-bit arrays, which does not require significant memory. This serial memory allocation of the atom positions, however, becomes significant for very large systems, which must be treated in parallel. A compile option exists in the code to use a parallelized atom position storage scheme, which limits some output capabilities but provides significant memory savings.

*2) Strain:* The materials making up the QD nanostructure may differ in their lattice constants; for the InAs/GaAs system, this difference is on the order of 7%. This lattice mismatch leads to the appearance of strain: Atoms throughout the sample are displaced from their bulk positions. Knowledge of equilibrium atomic positions is crucial for the subsequent calculation of QD's electronic properties, which makes the computation of strain a necessary step in realistic simulations of these nanostructures.

NEMO 3-D computes strain field using an atomistic valence force field (VFF) method [31] with the Keating potential. In this approach, the total elastic energy of the sample is computed as a sum of bond-stretching and bond-bending contributions from each atom. The local strain energy at atom $i$ is given by a phenomenological formula

$$E_i = \frac{3}{8}\sum_j \left[ \frac{\alpha_{\rm ij}}{2d_{\rm ij}^2}\left(R_{\rm ij}^2 - d_{\rm ij}^2\right)^2 + \sum_{k>j}^{n} \frac{\sqrt{\beta_{\rm ij}\beta_{\rm ik}}}{d_{\rm ij}d_{\rm ik}}\left(\vec{R}_{\rm ij}\cdot\vec{R}_{\rm ik} - \vec{d}_{\rm ij}\cdot\vec{d}_{\rm ik}\right)^2 \right] \quad (1)$$

where the sum is carried out over the $n$ nearest neighbors $j$ of atom $i$, $\vec{d}_{\rm ij}$ and $\vec{R}_{\rm ij}$ are the bulk and actual (distorted) distances between neighbor atoms, respectively, and $\alpha_{\rm ij}$ and $\beta_{\rm ij}$ are empirical material-dependent elastic parameters. The equilibrium atomic positions are found by minimizing the total elastic energy of the system. Several other strain potentials [24], [25] are also implemented in NEMO 3-D. While they modify some of the strain details, they roughly have the same computational efficiency.

*3) Electronic Structure:* The single-particle energies and wave functions are calculated using an empirical nearest neighbor tight-binding model. The underlying idea of this approach is the selection of a basis consisting of atomic orbitals (such as $s$, $p$, $d$, and $s^*$) centered on each atom. These orbitals are further treated as a basis set for the Hamiltonian, which assumes the following form:

$$\hat{H} = \sum_i \varepsilon_i^{(\nu)} c_{i,\nu}^{+} c_{i,\nu} + \sum_{i,\nu,\mu} t_i^{(\nu\mu)} c_{i,\nu}^{+} c_{i,\mu} + \sum_{i,j,\nu,\mu} t_{\rm ij}^{(\nu\mu)} c_{i,\nu}^{+} c_{j,\mu} \quad (2)$$

where $c_{i,\nu}^{+}$ ($c_{i,\nu}$) is the creation (annihilation) operator of an electron on the orbital $\nu$ that is localized on atom $i$. In (2), the first term describes the on-site orbital terms that are found on the diagonal of the Hamiltonian matrix. The second term describes coupling between different orbitals that are localized on the same atom (only the spin–orbit coupling between $p$ orbitals), and the third term describes coupling between different orbitals on different atoms. The restriction in the summation of the last term is that atoms $i$ and $j$ are nearest neighbors.

The characteristic parameters $\varepsilon$ and $t$ are treated as empirical fitting parameters for each constituent material and bond type. They are usually expressed in terms of energy constants of $\sigma$ and $\pi$ bonds between the atomic orbitals. For example, for a simple cubic lattice, the interaction between the $s$ orbital that is localized on atom $i$ at the origin and the orbital $p_x$ that is localized on atom $j$ with coordinate $\vec{d}_{\rm ij} = a\hat{x}$ with respect to atom $i$ would simply be expressed as $t_{\rm ij}^{(s,p_x)} = V_{{\rm sp}\sigma}$. Most of the systems under consideration, however, crystallize in the zinc-blende lattice, which means that the distance between the nearest neighbors is described by a 3-D vector $\vec{d}_{\rm ij} = l\hat{x} + m\hat{y} + n\hat{z}$, with $l$, $m$, and $n$ as the directional cosines. These cosines rescale the interaction constants, so that the element describing the interaction of the orbitals $s$ and $p_x$ is $t_{\rm ij}^{(s,p_x)} = lV_{{\rm sp}\sigma}$. The parameterization of all bonds using analytical forms of directional cosines for various tight-binding models is given in [32]. NEMO 3-D provides the user with choices of the $sp^3d^5s^*$, $sp^3s^*$, and single $s$-orbital models with and without spin, in zincblende, wurzite, and simple cubic lattices.

Additional complications arise in strained structures, where the atomic positions deviate from the ideal (bulk) crystal lattice [33]. The presence of strain leads to distortions not only of bond directions but also of bond lengths. In this case, the discussed interaction constant $t_{\rm ij}^{(s,p_x)} = l'V_{{\rm sp}\sigma}(d/d_0)^{\eta(sp\sigma)}$, where the new directional cosine $l'$ can be obtained analytically from the relaxed atom positions, but the bond-stretch exponent $\eta(sp\sigma)$ needs to be fitted to the experimental data. The energy constants parameterizing the on-site interaction change as well due to bond renormalization [11], [12].

The 20-band nearest neighbor tight-binding model is thus parameterized by 34 energy constants and 33 strain parameters, which need to be established by fitting the computed electronic properties of materials to those measured experimentally. This is done by considering bulk semiconductor crystals (such as GaAs or InAs) under strain. The summation in the Hamiltonian for these systems is done over the primitive crystallographic

TABLE I
PERFORMANCE ON $P$ PROCESSORS: time (IN SECONDS), NUMBER OF MATRIX VECTOR MULTIPLICATIONS (# mvs), MEMORY (mem.), AND NUMBER OF CORRECT EIGENVALUES TIMES THEIR MULTIPLICITY (# eigs) FOR LANCZOS, TRACEMIN, AND PARPACK $k$ EIGENVALUE SOLVERS IN NEMO 3-D SOFTWARE PACKAGE

| | LANCZOS | | | | TRACEMIN | | | | PARPACK | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *P* | *time* | *# mvs* | *mem.* | *# eigs* | *Time* | *# mvs* | *mem.* | *# eigs* | *time* | *# mvs* | *mem.* | *# eigs* |
| 1 | 197.2 | 7000 | 55.15 | 14 | 14298.0 | 390000 | 227.15 | 14 | 2155.9 | 11700 | 177.83 | 11 |
| 2 | 121.9 | 7000 | 28.11 | 14 | 8341.7 | 400000 | 114.28 | 14 | 579 | 8300 | 89.58 | 11 |
| 3 | 89.0 | 7000 | 18.95 | 14 | 5480.0 | 390000 | 76.20 | 14 | 412 | 8900 | 59.85 | 11 |
| 4 | 75.3 | 7000 | 14.49 | 14 | 4346.9 | 400000 | 57.66 | 14 | 912.8 | 26900 | 45.35 | 12 |

unit cell only. The model makes it possible to compute the band structure of the semiconductor throughout the entire Brillouin zone. For the purpose of the fitting procedure, however, only the band energies and effective masses at high symmetry points are targeted, and the tight-binding parameters are adjusted until a set of values closely reproducing these target values is found. Search for optimal parameterization is done using a genetic algorithm, as described in detail in [11] and [23]. Once it is known for each material constituting the QD, a full atomistic calculation of the single-particle energy spectrum is carried out on samples composed of millions of atoms. No further material properties are adjusted for the nanostructure once they are defined as basic bulk material properties.

*4) Postprocessing of QD Eigenstates:* From the single-particle eigenstates, various physical properties can be calculated in NEMO 3-D, such as optical matrix elements [34], Coulomb and exchange matrix elements [14], and approximate single-cell band structures from supercell band structure [35], [36], [37].

### C. Algorithmic and Numerical Aspects

*1) Parallel Implementation:* The complexity and generality of physical models in NEMO 3-D can place high demands on computational resources. For example, in the 20-band electronic calculation, the discrete Hamiltonian matrix is on the order of 20 times the number of atoms. Thus, in a computation with 20 million atoms, the matrix is on the order of 400 million. Computations of that size can be handled because of the parallelized design of the package. NEMO 3-D is implemented in ANSI C, C++ with MPI used for message passing, which ensures its portability to all major high-performance computing platforms and allows for an efficient use of distributed memory and parallel execution mechanisms.

Although the strain and electronic parts of the computation are algorithmically different, the key element in both is the sparse matrix–vector multiplication. This allows the use of the same memory distribution model in both phases. The computational domain is divided into vertical slabs. All atoms from the same slab are assigned to a single CPU, so if all nearest neighbors of an atom belong to its slab, no inter-CPU communication is necessary. The interatomic couplings are then fully contained in one of the diagonal blocks of the matrix. On the other hand, if an atom is positioned on the interface between slabs, it will couple to atoms belonging both to its own and the neighboring slab. This coupling is described by the off-diagonal blocks of the matrix. Its proper handling requires inter-CPU communication. However, due to the first nearest neighbor character of the strain and electronic models, the messages need to be passed only between pairs of CPUs corresponding to adjacent domains—even if the slabs are one atomic layer thick. Full-duplex communication patterns are implemented such that all interprocessor communications can be performed in two steps [11].

*2) Core Algorithms and Memory Requirements:* In the strain computation, the positions of the atoms are computed to minimize the total elastic strain energy. The total elastic energy in the VFF approach has only one global minimum, and its functional form in atomic coordinates is quartic. The conjugate gradient minimization algorithm in this case is well behaved and stable. The total elastic energy operator is never stored in its matrix form, but the interatomic couplings are computed on the fly. Therefore, the only data structures allocated in this phase are the vectors necessary for the conjugate gradient. The implementation used in NEMO 3-D requires six vectors, each of the total size of $3\times$ the number of atoms (to store atomic coordinates, gradients, and intermediate data); however, all those vectors are divided into slabs and distributed among CPUs, as previously discussed. The final atom position vectors are, by default, stored on all the CPUs for some technical output details. They can be distributed to the various CPUs at compile time, resulting in reduced output capabilities.

The electronic computation involves a very large eigenvector computation (matrices on the order of hundreds of millions or even billions). The algorithms/solvers available in NEMO 3-D include the PARPACK library [38], a custom implementation of the Lanczos method [39], the spectrum folding method [40], and the Tracemin [41]. The research group is also working on implementation of Lanczos with deflation, block Lanczos, and Jacobi–Davidson [42] methods.

The Lanczos algorithm employed here is not restarted, and the Lanczos vectors are not reorthogonalized. Moreover, the spectrum of the matrix has a gap, which lies in the interior of the spectrum. Typically, a small set of eigenvalues is sought, immediately above and below the gap. The corresponding eigenstates are electron and hole wave functions, assuming effectively nonzero values only inside and in the immediate vicinity of the dot. Also, in the absence of the external magnetic field, the eigenvalues are repeated, which reflects the spin degeneracy of electronic states. The advantage of the Lanczos algorithm is that it is fast, whereas the disadvantage is that it does not find the multiplicity and can potentially miss eigenvalues. Some comparisons have shown that the Lanczos method is faster

TABLE II
SPECTRUM OF THE EIGENVALUES AROUND 0 (WITH CORRECT MULTIPLICITY OF 2) AND EIGENVALUE MULTIPLICITY OBTAINED BY THE LANCZOS, TRACEMIN, AND PARPACK EIGENVALUE SOLVERS. THE NUMBER OF SEARCHED EIGENVALUES WAS KEPT CONSTANT FOR THESE THREE METHODS

| *Eigenvalues* | LANCZOS | TRACEMIN | PARPACK |
|---|---|---|---|
| -1.72338200E-01 | 1 | | 1 |
| -1.66029400E-01 | 1 | | 1 |
| -1.59010400E-01 | 1 | | 1 |
| -1.47522100E-01 | 1 | | 1 |
| -1.38917800E-01 | 1 | | 1 |
| -1.17807000E-01 | 1 | | 1 |
| -1.01703700E-01 | 1 | 2 | 2 |
| -7.80348200E-02 | 1 | 2 | 2 |
| -5.17194400E-02 | 1 | 2 | 1 |
| -2.81959000E-03 | 1 | 2 | |
| 3.93045300E-02 | 1 | 2 | |
| 7.66237500E-02 | 1 | 2 | |
| 1.16104700E-01 | 1 | 2 | |
| 1.57112300E+00 | 1 | | |

by a factor of 10 for the NEMO 3-D matrix than PARPACK. Tracemin algorithm finds the correct spectrum of degenerate eigenvalues but is slower than Lanczos. PARPACK has been found to be less reliable for this problem, taking more time than Lanczos and missing some of the eigenvalues and their multiplicity. Tables I and II give a comparison of Lanczos, PARPACK, and Tracemin (the *number* of searched eigenvalues was kept constant). The majority of the memory allocated in the electronic calculation in Lanczos is taken up by the Hamiltonian matrix. This matrix is very large but is typically very sparse; this property is explicitly accounted for in the memory allocation scheme. All matrix entries are, in general, complex and are stored in single precision. The code has an option not to store the Hamiltonian matrix but to recompute it each time it needs to be applied to a vector. In the Lanczos method, this is required once in each iteration. The PARPACK and Tracemin algorithms require the allocation of a significant number of vectors as a workspace, which is comparable to or larger than the Hamiltonian matrix. This additional memory need may require a matrix recompute for memory savings.

Fig. 3 shows the memory requirements for the two main phases of the code (strain and electronic structure calculations). It shows how the number of atoms that can be treated grows as a function of the number of CPUs for a fixed amount of memory per CPU. The number of atoms can be intuitively characterized by the length of one side of a cube that would contain that many atoms. This length is shown in Fig. 3, on the vertical axis on the right side of each plot. This figure shows that for a given amount of memory per CPU in the strain calculation [shown in Fig. 3(a)], the number of atoms that can be handled levels off after a certain CPU count, whereas for the electronic structure calculation [shown in Fig. 3(b)], the number of atoms that can be treated in NEMO 3-D continues to grow for larger CPU counts. The unfavorable memory scaling in the strain calculation is due to the allocation of all the atom positions on a single CPU. Distribution of this memory is possible at compile

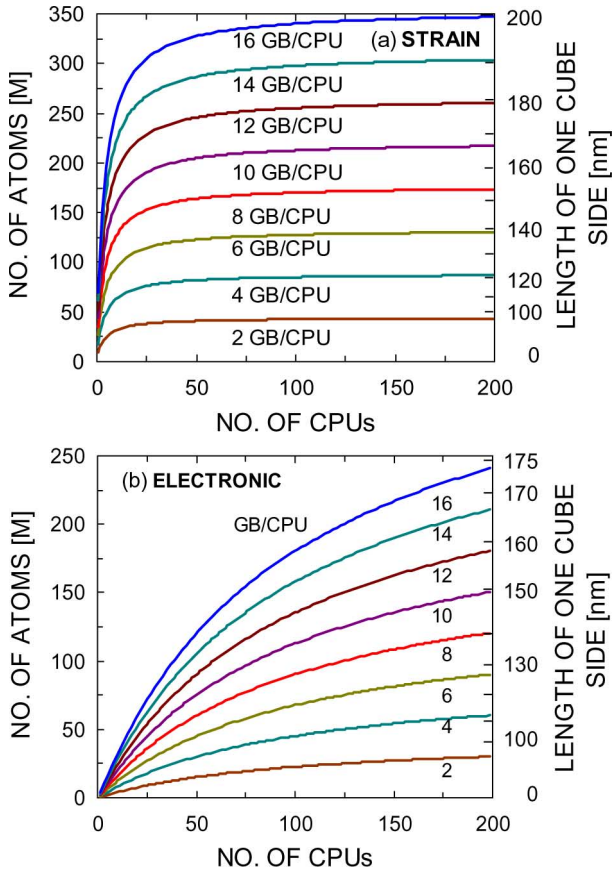



Fig. 3. Number of atoms that can be treated as a function of the number of CPUs for different amounts of memory per CPU. The plot on top is for the strain calculation, and the one on the bottom is for the electronic structure calculation. The vertical axis on the right side of each plot gives the equivalent length in nanometers of one side of the cube that would contain the given number of atoms.

time but has limited output capability. The strain calculations have so far never been memory limited. NEMO 3-D is typically size limited in the electronic structure calculation.

*3) Scaling:* Out of the two phases of NEMO 3-D, the strain calculation is algorithmically and computationally simpler. The Lanczos diagonalization of the Hamiltonian matrix, on the other hand, is much more challenging computationally.

To investigate the performance of NEMO 3-D package, computation was performed in a single dome-shaped InAs QD nanostructure embedded in a GaAs barrier material, as shown in Fig. 2. The high-performance computing (HPC) platforms used in the performance studies are shown in Table III. These include three Linux clusters at the Rosen Center for Advanced Computing (RCAC), Purdue University (PU), with Intel processors (32-bit Xeon, 64-bit Xeon, and dual-core Woodcrest). The PU/Woodcrest cluster has two dual-core chips per node. The other three platforms are a BlueGene at the Rensselaer Polytechnic Institute, the Cray XT3 at the Pittsburgh Supercomputing Center, and the SGI Altix at the National Center for Supercomputing Applications (NCSA). The processors on the Altix are Intel Itanium 2 processors; on the BlueGene, they are IBM PowerPC's, whereas the Cray XT3 has AMD Opterons. These three platforms have proprietary interconnects that are of higher performance than Gigabit Ethernet for the three Linux

TABLE III
SPECIFICATIONS FOR THE HPC PLATFORMS USED IN THE PERFORMANCE COMPARISONS

| *Platform* | *Type* | *CPU* | *Interconnect* | *Location* |
|---|---|---|---|---|
| PU/Xeon64 | Linux Cluster | Xeon x86-64 3.2GHz | Gigabit Ethernet | RCAC |
| PU/Xeon32 | Linux Cluster | Xeon 3.06GHz | Gigabit Ethernet | RCAC |
| PU/Woodcrest | Linux Cluster | Xeon x86-64 Dual Core 2.33GHz | Gigabit Ethernet | RCAC |
| PSC/XT3 | Cray XT3 | Opteron x86-64 2.6GHz | Native | PSC |
| NCSA/Altix | SGI Altix | Itanium2 IA-64 1.6GHz | SGI NUMAlink | NCSA |
| RPI/BGL | BlueGene/L | PowerPC 440 0.7 GHz | Native | RPI |

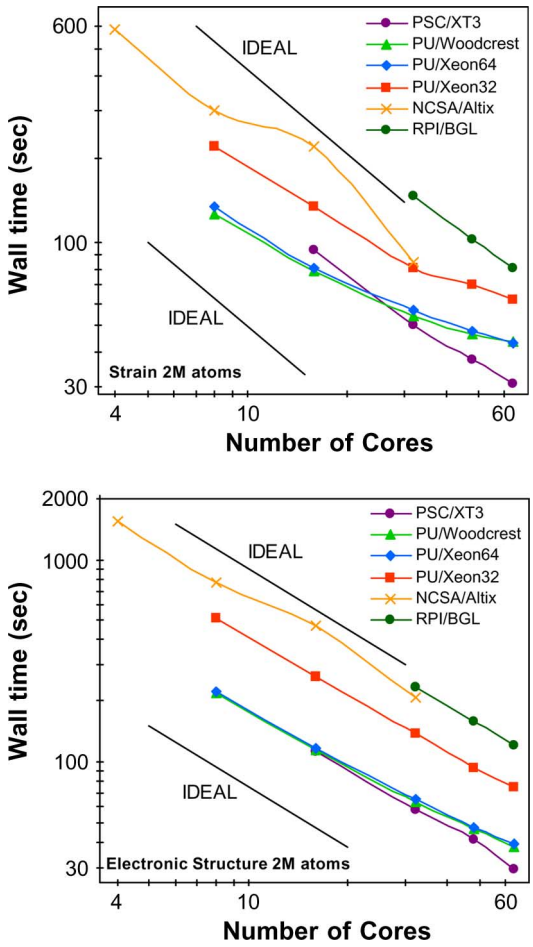



Fig. 4. Parallel performance of NEMO 3-D on some HPC platforms.

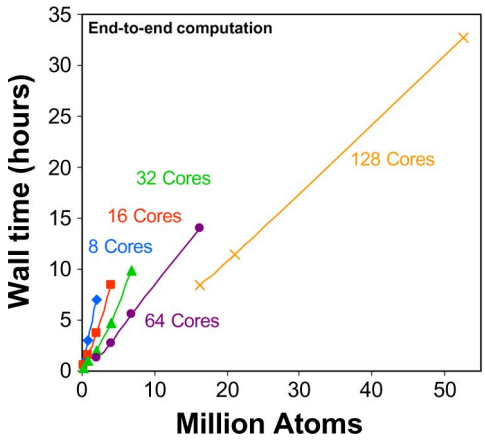



Fig. 5. Wall clock time versus number of atoms for end-to-end computations of the electronic structure of a QD for various numbers of cores on the PU/Woodcrest cluster.

clusters at PU. In the following, the terms processors and cores are used interchangeably.

Fig. 4 shows the performance of NEMO 3-D for each of the architectures. The wall clock times for 100 iterations for the energy minimization in the strain phase and 100 iterations of the Lanczos method for the electronic structure phase are shown as a function of the number of cores. The benchmark problem includes two million atoms. Fig. 4 shows that the PU/Woodcrest cluster and the PU/Xeon64 cluster are very close in performance for the same number of cores. These are both close to the performance of the Cray XT3 for lower core counts, whereas the XT3 performs better for higher core counts due to its faster interconnect. The older cluster, i.e., PU/Xeon32, is slower by a factor of about 2–2.3 as compared to the Woodcrest cluster. The BlueGene's slower performance is consistent with its lower clock speed, whereas the scalability reflects its efficient interconnect. The performance of the Altix is lower than expected.

In addition to the performance for the benchmark cases, with 100 iterations in the strain and electronic structure cases, end-to-end runs on the PU/Woodcrest cluster are carried out next (Fig. 5). This involves iterating to convergence and computing the eigenstates in the desired range (conduction band and valence band). For each problem size, which is measured in millions of atoms, the end-to-end cases were run to completion for one choice of number of cores. The iteration counts for the Lanczos computation are given in Table IV.

The numerical experiment is designed to demonstrate NEMO 3-D's ability to extract targeted interior eigenvalues and vectors out of virtually identical systems of increasing size. A single dome-shaped InAs QD embedded in GaAs is considered. The GaAs buffer is increased in size to increase the dimension of the system while not affecting confined states in the QD. It is verified [43] that the eigenvectors retain the expected symmetry of the nanostructure.

### D. Visualization

The QD simulation data of NEMO 3-D contain multivariate wave functions and strain profiles of the device structure. For effective 3-D visualizations of these results, a hardware-accelerated direct volume rendering system [44] has been developed, which is combined with a graphical user interface (GUI) based on *Rappture*.[1] This visualization system uses data

[1]*Rappture* is a toolkit supporting rapid application infrastructure, which is developed by Network for Computational Nanotechnology, Purdue University.

TABLE IV
ITERATION COUNTS FOR THE LANCZOS COMPUTATION AS A FUNCTION OF SYSTEM SIZE

| NUMBER OF ATOMS (MILLIONS) | 0.89 | 1.99 | 3.92 | 6.80 | 16.18 | 21.07 | 52.57 |
|---|---|---|---|---|---|---|---|
| Number of Lanczos iterations | 5,061 | 5,121 | 6,141 | 7,921 | 9,621 | 10,401 | 14,691 |

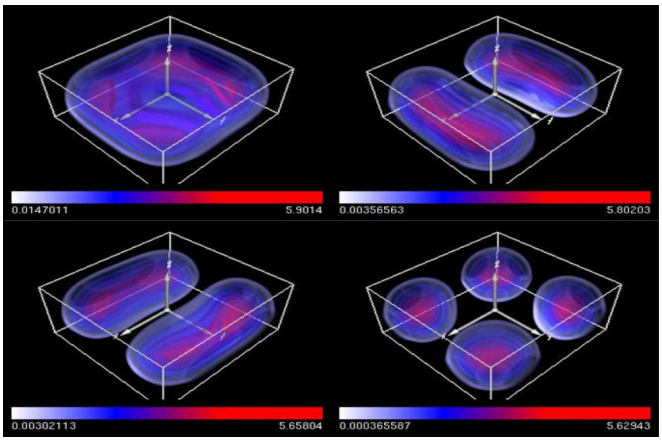



Fig. 6. Wave functions of electron on the first four states in conduction band.

set with *OPEN-DX*[2] format that are directly generated from NEMO 3-D. Fig. 6 shows the wave functions of electron on the first four eigenstates in conduction band of QD, which has 268 800 atoms in the electronic domain.

### *E. Release and Deployment of NEMO 3-D Package*

NEMO 3-D was developed on Linux clusters at the Jet Propulsion Laboratory (JPL) and was released with an open-source license in 2003. The originally released source is hosted at http://www.openchannelfoundation.org. As NEMO 3-D is undergoing further developments by the NCN, we are planning future releases of the NEMO 3-D source through http://www.nanoHUB.org. NEMO 3-D has been ported to different HPC platforms such as the NSF's TeraGrid (the Itanium2 Linux cluster at NCSA), Pittsburgh's Alpha cluster, SGI Altix, IBM p690, and various Linux clusters at PU and JPL.

The NEMO 3-D project is now part of a wider initiative, the NSF NCN. The main goal of this initiative is to support the National Nanotechnology Initiative through research, simulation tools, and education and outreach. Deployment of these services to the science and engineering community is carried out via web-based services, accessible through the nanoHUB portal http://www.nanoHUB.org. The educational outreach of NCN is realized by enabling access to multimedia tutorials, which demonstrate state-of-the-art nanodevice modeling techniques, and by providing space for relevant debates and scientific events (cyberinfrastructure). The second purpose of NCN is to provide a comprehensive suite of nanosimulation tools, which include electronic structure and transport simulators of molecular, biological, nanomechanical, and nanoelectronic systems. Access to these tools is granted to users via the web browsers, without the necessity of any local installation by the remote users. The definition of specific sample layout and parameters is done using a dedicated GUI in the remote desktop (Virtual Network Computing) technology. The necessary computational resources are further assigned to the simulation dynamically by the web-enabled middleware, which automatically allocates the necessary amount of CPU time and memory. The end user, therefore, has access not only to the code, a user interface, and the computational resources necessary to run it but also to the scientific and engineering community responsible for its maintenance.

Recently, a prototype GUI based on the *Rappture* package (http://www.rappture.org) is incorporated within the NEMO 3-D package, and a web-based online *interactive* version (Quantum Dot Lab) for educational purposes is freely available on http://www.nanohub.org [45]. The currently deployed educational version is restricted to a single $s$-orbital basis (single-band effective mass) model and runs in seconds. Quantum Dot Lab was deployed in November 2005, and its usage during the past year increased to 924 users conducting 6127 simulation runs (Fig. 7). Users can generate and freely rotate 3-D wave functions interactively powered by a remote visualization service.

The complete NEMO 3-D package is available to selected members of the NCN community through the use of a nanoHUB workspace. A nanoHUB workspace presents a complete Linux workstation to the user within the context of a web browser. The workstation persists beyond the browser lifetime, enabling the user to perform long-duration simulations without requiring their constant attention. As shown in this paper, the computational resources that are required to perform device-scale simulations are considerable and beyond the reach of many researchers. With this requirement in mind, NCN has joined forces with Teragrid [46] and the Open Science Grid [47] to seamlessly provide the necessary backend computational capacity to do scientifically significant computing. Computational resources necessary for large-scale parallel computing are linked to nanoHUB through the Teragrid *Science Gateways* program. Access to a Teragrid allocation is provided for members of the NCN community. Development of a more comprehensive NEMO 3-D user interface continues. The more comprehensive interface will provide access to a broader audience and encourage the continued growth of the nanoHUB user base.

## IV. CONCLUSION

NEMO 3-D is introduced to the IEEE Nanoelectronics community as a versatile open-source electronic structure code that can handle device domains relevant for realistic large devices. Realistic devices containing millions of atoms can be computed with reasonably easily available cluster computers. NEMO 3-D

[2]*OPEN-DX* is a package of open-source visualization software based on IBM's Visualization Data Explorer.

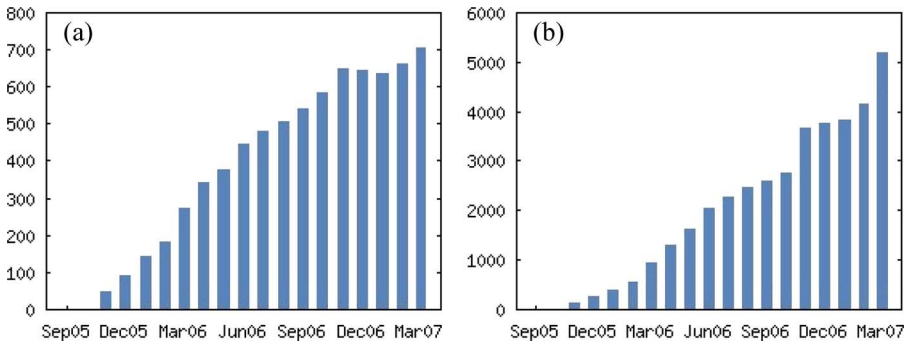



Fig. 7. (a) Number of annual users who have run at least one simulation. (b) Annualized simulation runs executed by nanoHUB users.

employs a VFF Keating model for strain and the 20-band $sp^3d^5s^*$ empirical tight-binding model for the electronic structure computation. It is released under an open-source license and maintained by the NCN, an organization dedicated to develop and deploy advanced nanoelectronic modeling and simulation tools. NEMO 3-D is not limited to research computing alone: A first educational version including visualization capabilities has been released on and has been used by hundreds of users for thousands of simulations. In the next part, the use of NEMO 3-D is demonstrated in the modeling and calculation of single-particle electronic states of a large variety of relevant realistically sized nanoelectronic devices.

## ACKNOWLEDGMENT

This work was carried out in part at the JPL, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The development of the NEMO 3-D tool involved a large number of individuals at JPL and PU, whose work has been cited. Dr. R. C. Bowen, Dr. F. Oyafuso, and Dr. S. Lee were key contributors in this large effort at JPL. Access to the Bluegene was made available through the auspices of the Computational Center for Nanotechnology Innovations, Rensselaer Polytechnic Institute. The authors would like to thank the NSF for Teragrid Award DMR070032 and RCAC, PU, for their support. nanoHUB computational resources were used for part of this work.

## REFERENCES

[1] N. Kharche, M. Prada, T. B. Boykin, and G. Klimeck, "Valley-splitting in strained silicon quantum wells modeled with 2 degree miscuts, step disorder, and alloy disorder," *Appl. Phys. Lett.*, vol. 90, no. 9, p. 092 109, Feb. 2007.

[2] P. M. Petroff, "Single quantum dots: Fundamentals, applications, and new concepts," in P. Michler, Ed. Berlin, Germany: Springer-Verlag, 2003.

[3] P. Michler, A. Kiraz, C. Becher, W. V. Schoenfeld, P. M. Petroff, L. Zhang, E. Hu, and A. Imamolu1, "A quantum dot single-photon turnstile device," *Science*, vol. 290, no. 5, pp. 2282–2285, Dec. 2000.

[4] M. A. Reed, J. N. Randall, R. J. Aggarwal, R. J. Matyi, T. M. Moore, and A. E. Wetsel, "Observation of discrete electronic states in a zero-dimensional semiconductor nanostructure," *Phys. Rev. Lett.*, vol. 60, no. 6, pp. 535–537, Feb. 1988.

[5] M. A. Reed, "Quantum Dots," *Sci. Amer.*, vol. 268, no. 1, p. 118, 1993.

[6] G. Bester and A. Zunger, "Cylindrically shaped zinc-blende semiconductor quantum dots do not have cylindrical symmetry: Atomistic symmetry, atomic relaxation, and piezoelectric effects," *Phys. Rev. B, Condens. Matter*, vol. 71, p. 045 318, 2005. please see references therein.

[7] C. Pryor, J. Kim, L. W. Wang, A. J. Williamson, and A. Zunger, "Comparison of two methods for describing the strain profiles in quantum dots," *J. Appl. Phys.*, vol. 83, no. 8, pp. 2548–2554, Mar. 1998.

[8] M. Grundmann, O. Stier, and D. Bimberg, "InAs/GaAs pyramidal quantum dots: Strain distribution, optical phonons, and electronic structure," *Phys. Rev. B, Condens. Matter*, vol. 52, no. 16, p. 11 969, Oct. 1995.

[9] O. Stier, M. Grundmann, and D. Bimberg, "Electronic and optical properties of strained quantum dots modeled by 8-band $k \cdot p$ theory," *Phys. Rev. B, Condens. Matter*, vol. 59, no. 8, p. 5688, Feb. 1999.

[10] A. Canning, L. W. Wang, A. Williamson, and A. Zunger, "Parallel empirical pseudopotential electronic structure calculations for million atom systems," *J. Comput. Phys.*, vol. 160, no. 1, pp. 29–41, May 2000.

[11] G. Klimeck, F. Oyafuso, T. Boykin, R. Bowen, and P. von Allmen, "Development of a nanoelectronic 3-D (NEMO 3-D) simulator for multimillion atom simulations and its application to alloyed quantum dots," *Comput. Model. Eng. Sci.*, vol. 3, p. 601, 2002.

[12] T. B. Boykin, G. Klimeck, R. C. Bowen, and F. Oyafuso, "Diagonal parameter shifts due to nearest-neighbor displacements in empirical tight-binding theory," *Phys. Rev. B, Condens. Matter*, vol. 66, no. 12, p. 125 207, Sep. 2002.

[13] T. B. Boykin, R. C. Bowen, and G. Klimeck, "Electromagnetic coupling and gauge invariance in the empirical tight-binding method," *Phys. Rev. B, Condens. Matter*, vol. 63, no. 24, p. 245 314, Jun. 2001.

[14] S. Lee, J. Kim, L. Jönsson, J. W. Wilkins, G. Bryant, and G. Klimeck, "Many-body levels of multiply charged and laser-excited InAs nanocrystals modeled by empirical tight binding," *Phys. Rev. B, Condens. Matter*, vol. 66, p. 235 307, 2002.

[15] G. Klimeck, R. C. Bowen, T. B. Boykin, C. Salazar-Lazaro, T. A. Cwik, and A. Stoica, "Si tight-binding parameters from genetic algorithm fitting," *Superlattices Microstruct.*, vol. 27, no. 2/3, pp. 77–88, Mar. 2000.

[16] G. Klimeck, R. C. Bowen, T. B. Boykin, and T. A. Cwik, "$sp^3s^*$ tight-binding parameters for transport simulations in compound semiconductors," *Superlattices Microstruct.*, vol. 27, no. 5, pp. 519–524, May 2000.

[17] R. C. Bowen, G. Klimeck, R. Lake, W. R. Frensley, and T. Moise, "Quantitative resonant tunneling diode simulation," *J. Appl. Phys.*, vol. 81, p. 3207, 1997.

[18] G. Klimeck, T. B. Boykin, R. C. Bowen, R. Lake, D. Blanks, T. Moise, Y. C. Kao, and W. R. Frensley, "Quantitative simulation of strained InP-based resonant tunneling diodes," in *Proc. 55th IEEE Device Res. Conf. Dig.*, 1997, p. 92.

[19] T. B. Boykin, G. Klimeck, and F. Oyafuso, "Valence band effective mass expressions in the $sp^3d^5s^*$ empirical tight-binding model applied to a new Si and Ge parameterization," *Phys. Rev. B, Condens. Matter*, vol. 69, no. 11, p. 115 201, Mar. 2004.

[20] T. B. Boykin, N. Kharche, and G. Klimeck, "Brillouin zone unfolding of perfect supercells composed of non-equivalent primitive cells," *Phys. Rev. B, Condens. Matter*. submitted for publication.

[21] S. Ahmed, M. Usman, C. Heitzinger, R. Rahman, A. Schliwa, and G. Klimeck, "Atomistic simulation of non-degeneracy and optical polarization anisotropy in zincblende quantum dots," in *Proc. 2nd Annu. IEEE Int. Conf. Nano/Micro Eng. and Molecular Syst. (IEEE-NEMS)*, Bangkok, Thailand, Jan. 16–19, 2007.

[22] G. Liang, J. Xiang, N. Kharche, G. Klimeck, C. M. Lieber, and M. Lundstrom, *Performance Analysis of a Ge/Si Core/Shell Nanowire Field Effect Transistor*. cond-mat/0611226.

[23] M. A. Eriksson, M. Friesen, S. N. Coppersmith, R. Joynt, L. J. Klein, K. Slinker, C. Tahan, P. M. Mooney, J. O. Chu, and S. J. Koester,

"Spin-based quantum dot quantum computing in silicon," *Quantum Inf. Process.*, vol. 3, no. 1–5, pp. 133–146, Oct. 2004.
[24] A. J. Williamson, L. W. Wang, and A. Zunger, "Theoretical interpretation of the experimental electronic structure of lens-shaped self-assembled InAs/GaAs quantum dots," *Phys. Rev. B, Condens. Matter*, vol. 62, no. 19, pp. 12 963–12 977, Nov. 2000.
[25] O. L. Lazarenkova, P. von Allmen, F. Oyafuso, S. Lee, and G. Klimeck, "Effect of anharmonicity of the strain energy on band offsets in semiconductor nanostructures," *Appl. Phys. Lett.*, vol. 85, no. 18, p. 4193, Nov. 2004.
[26] M. Korkusinski and G. Klimeck, "Atomistic simulations of long-range strain and spatial asymmetry molecular states of seven quantum dots," *J. Phys.: Conf. Ser.*, vol. 38, pp. 75–78, 2006.
[27] F. Oyafuso, G. Klimeck, P. von Allmen, T. B. Boykin, and R. C. Bowen, "Strain effects in large-scale atomistic quantum dot simulations," *Phys. Stat. Sol. B*, vol. 239, no. 1, pp. 71–79, Sep. 2003.
[28] F. Oyafuso, G. Klimeck, R. C. Bowen, T. B. Boykin, and P. von Allmen, "Disorder induced broadening in multimillion atom alloyed quantum dot systems," *Phys. Stat. Sol. C*, vol. 4, p. 1149, 2003.
[29] M. Korkusinski, F. Saied, H. Xu, S. Lee, M. Sayeed, S. Goasguen, and G. Klimeck, "Large scale simulations in nanostructures with NEMO 3-D on Linux clusters," in *Proc. Linux Cluster Inst. Conf.*, Raleigh, NC, Apr. 2005.
[30] M. Graf and P. Vogl, "Electromagnetic fields and dielectric response in empirical tight-binding theory," *Phys. Rev. B, Condens. Matter*, vol. 51, no. 8, pp. 4940–4949, Feb. 1995.
[31] P. Keating, "Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure," *Phys. Rev.*, vol. 145, no. 2, pp. 637–644, May 1966.
[32] J. C. Slater and G. F. Koster, "Simplified LCAO method for the periodic potential problem," *Phys. Rev.*, vol. 94, no. 6, pp. 1498–1524, 1954.
[33] J. M. Jancu, R. Scholz, F. Beltram, and F. Bassani, "Empirical $spds^*$ tight-binding calculation for cubic semiconductors: General method and material parameters," *Phys. Rev. B, Condens. Matter*, vol. 57, no. 11, pp. 6493–6507, Mar. 1998.
[34] T. B. Boykin and P. Vogl, "Dielectric response of molecules in empirical tight-binding theory," *Phys. Rev. B, Condens. Matter*, vol. 65, no. 3, p. 035 202, Jan. 2001.
[35] T. B. Boykin and G. Klimeck, "Practical application of zone-folding concepts in tight-binding," *Phys. Rev. B, Condens. Matter*, vol. 71, no. 11, p. 115 215, Mar. 2005.
[36] T. B. Boykin, N. Kharche, G. Klimeck, and M. Korkusinski, "Approximate bandstructures of semiconductor alloys from tight-binding supercell calculations," *J. Phys: Condensed Matter*, vol. 19, no. 3, p. 036 203, Jan. 2007.
[37] T. B. Boykin, M. Luisier, A. Schenk, N. Kharche, and G. Klimeck, "The electronic structure and transmission characteristics of disordered AlGaAs nanowires," *IEEE Trans. Nanotechnol.*, vol. 6, no. 1, pp. 43–47, Jan. 2007.
[38] K. Maschhoff and D. Sorensen, "A portable implementation of ARPACK for distributed memory parallel architectures," in *Proc. Copper Mountain Conf. Iterative Methods*, 1996.
[39] C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *J. Res. Natl. Bur. Stand.*, vol. 45, no. 4, pp. 255–282, 1950.
[40] L. W. Wang and A. Zunger, "Solving Schrödinger's equation around a desired energy: Application to silicon quantum dots," *J. Chem. Phys.*, vol. 100, no. 3, pp. 2394–2397, Feb. 1994.
[41] A. Sameh and Z. Tong, "The trace minimization method for the symmetric generalized eigenvalue problem," *J. Comput. Appl. Math*, vol. 123, no. 1, pp. 155–175, Nov. 2000.
[42] G. L. G. Sleijppen and H. A. Van der Vorst, "A Jacobi–Davidson iteration method for linear eigenvalue problems," *SIAM J. Matrix Anal. Appl.*, vol. 17, no. 2, pp. 401–425, 1996.
[43] H. Bae, S. Clark, B. Haley, G. Klimeck, M. Korkusinski, S. Lee, M. Naumov, H. Ryu, and F. Saied, "Electronic structure computations of quantum dots with a billion degrees of freedom," in *Proc. Supercomputing*, Reno, NV, Nov. 2007.
[44] W. Qiao, M. Mclennan, R. Kennell, D. Ebert, and G. Klimeck, "Hub-based simulation and graphics hardware accelerated visualization for nanotechnology applications," *IEEE Trans. Vis. Comput. Graphics*, vol. 12, no. 5, pp. 1061–1068, Sep./Oct. 2006.
[45] [Online]. Available: https://www.nanohub.org/simulation_tools/qdot_tool_information
[46] *TeraGrid*. [Online]. Available: http://www.teragrid.org
[47] *Open Science Grid*. [Online]. Available: http://www.opensciencegrid.org

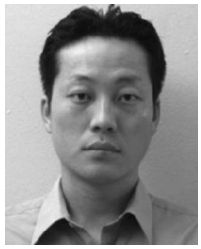
**Gerhard Klimeck** (S'93–M'94–SM'04) received the German electrical engineering degree from Ruhr-University Bochum, Bochum, Germany, in 1990 and the Ph.D. degree from Purdue University, West Lafayette, IN, in 1994.

He was the Technical Group Supervisor with the Applied Cluster Computing Technologies Group, National Aeronautics and Space Administration Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA. He was a member of Technical Staff with the Central Research Laboratory, Texas Instruments Incorporated. He is currently the Technical Director of the Network for Computational Nanotechnology and a Professor of electrical and computer engineering with Purdue University. He is also currently with the National Aeronautics and Space Administration Jet Propulsion Laboratory as a Principal Member on a faculty part-time basis. He leads the development and deployment of Web-based simulation tools that are hosted on the http://nanohub.org a community website that is utilized by over 22 000 users annually. His work is documented in over 170 peer-reviewed publications and over 270 conference proceedings. His research interests include the modeling of nanoelectronic devices, parallel cluster computing, and genetic algorithms. He led the development of NEMO 3-D, a tool that enables the simulation of tens-of-million atom quantum dot systems, and NEMO 1-D, which is the first nanoelectronic CAD tool.

Dr. Klimeck is a member of the American Physical Society, Eta Kappa Nu, and Tau Beta Pi.

**Shaikh Shahid Ahmed** received the B.S. degree in electrical and electronic engineering from Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 1998 and the M.S. and Ph.D. degrees in electrical engineering from Arizona State University, Tempe, in 2003 and 2005, respectively.

He is currently a Postdoctoral Research Associate with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN. He has published more than 40 journal articles and conference proceedings and two book chapters. His main research interests are atomistic electronic structure and quantum transport models and methodologies for nanostructures (including quantum dots and qubits, carbon nanotube and ribbons, nanowires and sensors, and solid-state light), classical and nonclassical semiconductor device and process simulations, novel numerical algorithms, particularly for molecular dynamics and Green's function computation, large-scale high-performance parallel cluster and distributed computing and the development of community nanotechnology application packages, and analog and digital design.



**Hansang Bae** received the B.S. degree in electronics engineering from Korea University, Seoul, South Korea, in 1999 and the M.S. degree in computer engineering from Purdue University, West Lafayette, IN, in 2003. He is currently working toward the Ph.D. degree in computer engineering in the School of Electrical and Computer Engineering, Purdue University.

He was a member of the Rosen Center for Advanced Computing, Purdue University, working on the performance characterization of HPC applications.

**Neerav Kharche** received the B.Tech. and M.Tech. degrees in metallurgical engineering and materials science from the Indian Institute of Technology Bombay, Mumbai, India, in 2003. He is currently working toward the Ph.D. degree in electrical and computer engineering at Purdue University, West Lafayette, IN.

His research interests are electronic structure and quantum transport in nanoscale devices. He is currently working on modeling the effects of disorder on the electronic structure of nanowires and quantum wells.

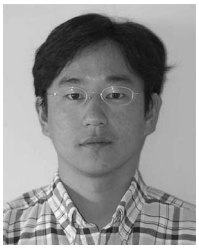
**Rajib Rahman**, photograph and biography not available at the time of publication.

**Steve Clark** received the Ph.D. degree in chemical engineering from Purdue University, West Lafayette, IN, in 1985.

He is currently an Application Engineer for Scientific Computing with the Rosen Center for Advanced Computing, Purdue University, maintaining and advancing the Web-based application pool hosted on http://nanoHUB.org. Before starting his current post, he was involved with projects on noncontinuous process simulation modeling.

**Benjamin Haley** received the B.S. degree in physics from Purdue University, West Lafayette, IN, in 1998 and the M.S. and Ph.D. degrees in applied science from the University of California, Davis, in 2001 and 2005, respectively.

He was with Cummins Engine Company and Intel Corporation. He is currently a Postdoctoral Research Associate with the School of Electrical Engineering, Purdue University, working on quantum transport and numerical algorithms.



**Sunhee Lee** received the B.S. and M.S. degrees in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 1999 and 2001, respectively. He is currently working toward the Ph.D. degree at Purdue University, West Lafayette, IN.

From 2001 to 2006, he was with the Telecommunication Division, Samsung Electronics, as a Research Engineer. His research interests include the implementation and optimization of nanoelectronic simulators.

**Maxim Naumov** received the B.Sc. degree in computer science and mathematics from Purdue University, West Lafayette, IN, in 2003. He is currently working toward the Ph.D. degree in computational science and engineering at Purdue University.

His research interests are numerical linear algebra, numerical analysis, and high-performance computing. He has participated in projects related to the numerical solution of linear systems and eigenvalue problems (using sequential and parallel algorithms), with applications in nanoelectronics and computational fluid dynamics.

**Hoon Ryu** received the B.S.E.E. degree from Seoul National University, Seoul, South Korea, in 2003 and the M.S.E.E. degree from Stanford University, Stanford, CA, in 2005. He is currently working toward the Ph.D. degree at Purdue University, West Lafayette, IN.

He was with the Semiconductor Division, Samsung Electronics, as a Research Engineer in 2006. His research interests include the implementation and optimization of nanoelectronic computation.

**Faisal Saied** received the Ph.D. degree in computer science from Yale University, New Haven, CT, in 1990.

He was the Manager of the Performance Engineering Group, National Center for Supercomputing Applications (NCSA), University of Illinois, Urbana-Champaign, directing the group to help users optimize their codes for the HPC platforms at NCSA. He is currently a Senior Research Scientist with the Computing Research Institute and the Rosen Center for Advanced Computing, Purdue University, West Lafayette, IN. His research interests include scientific computing, high-performance computing, parallel numerical algorithms, and the performance analysis of parallel programs. He has collaborated with computational scientists from a variety of application areas. These include developing parallel eigenvalue solvers for nanoelectronics, parallel multigrid methods for groundwater modeling, multigrid methods for biophysics, and preconditioned iterative solvers for structural dynamics.

**Marta Prada** received the degree in physics from the Universidad Autónoma de Madrid, Madrid, Spain, in 1998, and the Ph.D. degree from the University of Leeds, Leeds, U.K. in December 2005.

She was with the European Council for Nuclear Research (CERN), Geneva, Switzerland, for a couple years. She held a postdoctoral position at Purdue University, West Lafayette, IN, in 2006. She was with the School of Electrical and Computer Engineering and also with the Network for Computational Nanotechnology, Purdue University.

**Marek Korkusinski** received the Bachelor's degree from the Wroclaw University of Technology, Wroclaw, Poland, in 1998 and the Ph.D. degree from the University of Ottawa, Ottawa, ON, Canada, in 2004.

He is a Research Associate with the Quantum Theory Group, Institute for Microstructural Sciences, National Research Council of Canada, Ottawa, ON, Canada. Prior to his current appointment, he collaborated with Prof. G. Klimeck of Purdue University on multimillion-atom computational projects. His main research interests include theoretical condensed matter physics, particularly optical and electronic properties of semiconductor nanostructures, correlations in electronic and electron-hole systems, and nanospintronics. His current projects involve using the tb-DFT and *ab initio* methods in the calculation of the properties of semiconductor and carbon-based nanostructures.

**Timothy B. Boykin** (S'86–M'91–SM'05) received the B.S. degree in electrical engineering (summa cum laude) from Rice University, Houston, TX, in 1987 and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1988 and 1992, respectively.

He joined the Department of Electrical and Computer Engineering, University of Alabama in Huntsville (UAH), Huntsville, in September 1992, where he was an Associate Professor in August 1997, was granted the tenure in August 1998, and has been a Full Professor since August 2007. He is the author or coauthor of over 55 refereed journal articles, and his first authored papers alone have been cited by other (noncoauthor) researchers over 330 times. His main research interests are detailed semiconductor heterostructure and device modeling, particularly the first numerically stable localized-orbital basis heterostructure model, the impact of nonzero in-plane wave vector on tunneling, incompleteness in tight-binding models, improved methods for calculating boundary conditions in tight-binding models, optimizations and capabilities of tight-binding models, consequences of incompleteness for the quantum-mechanical continuity equation, the representation of electromagnetic effects in tight-binding models, valley splitting in silicon quantum wells, and modeling alloy nanostructures.

Dr. Boykin is a member of Phi Beta Kappa, Tau Beta Pi, Eta Kappa Nu, the American Physical Society, and Sigma Xi. He received the UAH Foundation Award for Research and Creative Achievement in Applied Science/Technology in 2001.