Encyclopedia of Complexity and Systems Science
DOI 10.1007/978-3-642-27737-5_343-2

Multimillion Atom Simulation of Electronic and Optical Properties of Nanoscale Devices Using NEMO 3-D

Shaikh Ahmed[b,a]*, Neerav Kharche[a], Rajib Rahman[a], Muhammad Usman[h], Sunhee Lee[a], Hoon Ryu[a], Hansang Bae[a], Steve Clark[c], Benjamin Haley[a], Maxim Naumov[d], Faisal Saied[c], Marek Korkusinski[e], Rick Kennel[c], Michael McLennan[c], Timothy B. Boykin[f] and Gerhard Klimeck[a,g]
[a]School of Electrical and Computer Engineering and Network for Computational Nanotechnology, Purdue University, West Lafayette, IN, USA
[b]Electrical and Computer Engineering Department, Southern Illinois University, Carbondale, IL, USA
[c]Rosen Center for Advanced Computing, Purdue University, West Lafayette, IN, USA
[d]Department of Computer Science, Purdue University, West Lafayette, IN, USA
[e]Institute for Microstructural Sciences, National Research Council of Canada, Ottawa, ON, Canada
[f]Electrical and Computer Engineering Department, The University of Alabama in Huntsville, Huntsville, AL, USA
[g]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[h]School of Physics, University of Melbourne, Parkville, VIC, Australia

Glossary and Notation

Atomistic simulation For device sizes in the range of tens of nanometers, 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 an atomistic simulation, one takes into account both the atomistic/granular and quantum properties of the underlying nanostructure.

Bandstructure Bandstructure of a solid originates from the wave nature of particles and depicts the allowed and forbidden energy states of electrons in the material. The knowledge of the bandstructure is the first and essential step toward the understanding of the device operation and reliable device design for semiconductor devices. Bandstructure is based on the assumption of an infinitely extended (bulk) material without spatial fluctuations (outside a simple repeated unit cell). For nanometer-scale devices with spatial variations on the atomic scale, the traditional concept of bandstructure is called into question.

nanoHUB The nanoHUB is a rich, Web-based resource for research, education, and collaboration in nanotechnology (http://www.nanoHUB.org). It was created by the NSF-funded Network for Computational Nanotechnology (NCN) with a vision to pioneer the development of nanotechnology from science to manufacturing through innovative theory, exploratory simulation, and novel cyberinfrastructure. The nanoHUB offers online nanotechnology simulation tools which one can freely access from his/her Web browser.

Nanostructures Nanostructures have at least two physical dimensions of size less than 100 nm. Their size lies between atomic/molecular and microscopic

*Email: shaikh.ahmed@gmail.com
*Email: ahmed@siu.edu

	structures/particles. Realistically sized nanostructures are usually composed of millions of atoms. 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.
NEMO 3D	NEMO 3D stands for Nanoelectronic Modeling in three dimensions. This versatile, open-source software package currently allows calculating single-particle electronic states and optical response of various semiconductor structures including bulk materials, quantum dots, impurities, quantum wires, quantum wells, and nanocrystals.
Piezoelectricity	A variety of advanced materials of interest, such as GaAs, InAs, and GaN, are piezoelectric. Piezoelectricity arises due to charge imbalances on the bonds between atoms. Modifications of the bond angles or distances result in alterations in charge imbalance. Any spatial nonsymmetric distortion/strain in nanostructures made of these materials will create piezoelectric fields, which may significantly modify the electrostatic potential landscape.
Quantum dots	Quantum dots (QDs) are solid-state nanostructures that provide confinement of charge carriers (electrons, holes, excitons) in all three spatial dimensions typically on the nanometer scale. This work focuses on semiconductor-based quantum dots.
Rappture	Rappture (http://www.rappture.org) is a software toolkit that supports and enables the rapid development of graphical user interfaces (GUIs) for different applications. It is developed by the Network for Computational Nanotechnology at Purdue University, West Lafayette.
Spontaneous (pyroelectric) polarization	The intrinsic asymmetry of the bonding in the equilibrium (unstrained) crystal structure leads to a spontaneous polarization in wurtzite III-N structures even in the absence of external electrical fields. Spontaneous polarization is strain independent and results in a built-in potential that, in many nitride nanostructures, becomes comparable to the piezoelectric counterpart.
Strain	Strain is the deformation caused by the action of stress on a physical body. In nanoelectronic devices, strain typically originates from the assembly of lattice-mismatched semiconductors. Strain can be atomistically inhomogeneous, and a small mechanical distortion of 2–5 % can strongly modify the energy spectrum, in particular, the optical bandgap, of the system by 30–100 %.
Tight binding	Tight binding is an empirical model that enables calculation of single-particle energies and wave functions in a solid. The essential idea is the representation of the electronic states of the valence electrons with a local basis that contains the critical physical elements needed. The basis may contain orthogonal *s*, *p*, and *d* orbitals on one atom that connect/talk to orbitals of a neighboring atom. The connection between atoms and the resulting overlapping wave functions form the bandstructure of a solid.

Encyclopedia of Complexity and Systems Science
DOI 10.1007/978-3-642-27737-5_343-2

Definition of the Subject and Its Importance

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, computation, information processing, biotechnology, and medical industries. At the atomic scale of novel nanostructured semiconductors, the distinction between new device and new material is blurred, and device physics and material science meet. The *quantum mechanical effects* in the electronic states of the device and the *granular*, *atomistic* representation of the underlying material become important. Modeling and simulation approaches based on a *continuum* representation of the underlying material typically used by device engineers and physicists become invalid. Typical ab initio methods used by material scientists do not represent the bandgaps and masses precisely enough for device design or they do not scale to *realistically sized devices which may contain millions of atoms*. The variety of geometries, materials, and doping configurations in semiconductor devices at the nanoscale suggests that a *general* nanoelectronic modeling tool is needed. The nanoelectronic modeling tool (NEMO 3D) has been developed to address these needs. Based on the atomistic valence force field (VFF) method and a variety of nearest-neighbor tight-binding models (s, sp^3s^*, $sp^3d^5s^*$), NEMO 3D enables the computation of strain for over 64 million atoms and of electronic structure for over 52 million atoms, corresponding to volumes of $(110\ \text{nm})^3$ and $(101\ \text{nm})^3$, respectively. Such extreme problem sizes involve very large-scale computations, and NEMO 3D has been designed and optimized to be scalable from single CPUs to large numbers of processors on commodity clusters and the most advanced supercomputers. Excellent scaling to 8192 cores/CPUs has been demonstrated. NEMO 3D is continually developed by the Network for Computational Nanotechnology (NCN) under an open-source license. A Web-based online interactive version for educational purposes is freely available on the NCN portal http://www.nanoHUB.org. This article discusses the theoretical models, essential algorithmic and computational components, and optimization methods that have been used in the development and the deployment of NEMO 3D. Also, successful applications of NEMO 3D are demonstrated in the atomistic calculation of single-particle electronic states of the following realistically sized nanostructures each consisting of multimillion atoms: (1) self-assembled quantum dots including long-range strain and piezoelectricity, (2) stacked quantum dots as used in quantum cascade lasers, (3) phosphorus (P) impurities in silicon used in quantum computation, (4) Si on SiGe quantum wells (QWs), and (5) SiGe nanowires. These examples demonstrate the broad NEMO 3D capabilities and indicate the necessity of multimillion atomistic electronic structure modeling.

Introduction

Emergence of Novel Nanoscale Semiconductor Devices

The new industrial age and the new economy are driven in large measure by unprecedented advances in information technology. The electronics industry is the largest industry in the world with global sales of over one trillion dollars since 1998. If current trends continue, the sales volume of the electronics industry is predicted to reach three trillion dollars and account for about 10 % of gross world product (GWP) by 2010 (Sze and May 2003). Basic to the electronic industry and the new information age are the *semiconductor devices* that implement all needed information processing operations. The revolution in the semiconductor industry was initiated in 1947 with the invention and fabrication of point-contact bipolar devices on *slabs* of polycrystalline germanium (Ge) used as the underlying semiconductor element (Agnello 2002). Later, the development of the planar process and the reliable and high-quality silicon dioxide (SiO_2) growth on silicon wafers, acting as an excellent barrier for the selective diffusion

DOI 10.1007/978-3-642-27737-5_343-2

steps, led to the invention of the silicon-based bipolar integrated circuits in 1959. A metal–oxide–semiconductor field-effect transistor (MOSFET), the most critical device for today's advanced integrated circuits, was reported by Kahng and Atalla in 1960 (Sze and May 2003). By 1968, both complementary metal–oxide–semiconductor devices (CMOS) and polysilicon gate technology allowing self-alignment of the gate to the source/drain of the device had been developed. The industry's transition from bipolar to CMOS technology in the 1980s was mainly driven by the increased power demand for high-performance integrated circuits.

The most important factor driving the continuous device improvement has been the semiconductor industry's relentless effort to reduce the cost *per function* on a chip (Vasileska et al. 2005). This is done by putting more devices on an integrated circuit chip while either reducing manufacturing costs or holding them constant. *Device scaling*, which involves reducing the transistor size while keeping the electric field constant from one generation to the next, has paved the way for a continuous and systematic increase in transistor density and improvements in system performance (described by Moore's Law Moore 1975) for the past 40 years. For example, regarding conventional/classical silicon MOSFETs, the device size is scaled in all dimensions, resulting in smaller oxide thickness, junction depth, channel length, channel width, and isolation spacing. Currently, 65 nm (with a physical gate length of 35 nm) is the state-of-the-art process technology, but even smaller dimensions are expected in the very near future.

However, recent studies by many researchers around the globe reveal the fact that the exponential growth in integrated circuit complexity as achieved through conventional scaling is finally facing its limits and will slow down in the very near future. Critical dimensions, such as transistor gate length and oxide thickness, are reaching physical limitations (Vasileska et al. 2005). Maintaining dimensional integrity at the limits of scaling is a challenge. Considering the manufacturing issues, photolithography becomes difficult as the feature sizes approach the wavelength of ultraviolet light. In addition, it is difficult to control the oxide thickness when the oxide is made up of just a few monolayers. Processes will be required approaching atomic-layer precision. In addition to the processing issues, there are also some fundamental device issues (Wong 2002). As the silicon industry moves into the 45-nm node regime and beyond, two of the most important challenges facing us are the growing dissipation of *standby power* and the increasing variability and mismatch in device characteristics.

The Semiconductor Industry Association (SIA) forecasts (Semiconductor Industry 2001) that the current rate of transistor performance improvement can be sustained for another 10–15 years, but only through the development and introduction of *new materials and transistor structures*. In addition, a major improvement in lithography will be required to continue size reduction. It is expected that these new technologies may extend MOSFETs to the 22-nm node (9-nm physical gate length) by 2016. Intrinsic device speed may exceed 1 THz and integration densities will be more than one billion transistors/cm^2. In many cases, the introduction of a new material requires the use of a new device structure, or vice versa. To fabricate devices beyond current scaling limits, IC companies are simultaneously pushing the planar, bulk silicon CMOS design while exploring alternative gate stack materials (high-k dielectric (Zhu et al. 2004) and metal gates), band engineering methods (using strained Si (Welser et al. 1992) or SiGe (Oberhuber et al. 1998)), and alternative transistor structures. The concept of a band-engineered transistor is to enhance the mobility of electrons and/or holes in the channel by modifying the bandstructure of silicon in the channel in a way such that the physical structure of the transistor remains substantially unchanged. This enhanced mobility increases the transistor transconductance (g_m) and on-drive current (I_{on}). A SiGe layer or a strained silicon on relaxed SiGe layer is used as the enhanced-mobility channel layer. Today, there is also an extensive research in double-gate (DG) structures and FinFET transistors (Colinge 2004), which have better electrostatic integrity and theoretically have better transport properties than single-gated FETs. Some novel and revolutionary technology such as carbon nanotubes, silicon nanowires, or

DOI 10.1007/978-3-642-27737-5_343-2

Table 1 Major semiconductor devices with the approximate date of their introduction

1874: Metal–semiconductor contact
1947: Bipolar junction transistors (BJT)
1954: Solar cell
1957: Heterojunction bipolar transistor (HBT)
1958: Tunnel diode
1959: Integrated circuits
1960: Field-effect transistors (FETs)
1962: Semiconductor lasers
1966: Metal–semiconductor FET
1967: Nonvolatile semiconductor memory
1974: Resonant tunneling diode (RTD)
1990: Magnetoresistive random-access memory (MRAM)
1991: Carbon nanotubes
1994: Room-temperature single-electron memory cell (SEMC)
1994: Quantum cascade laser
1998: Carbon nanotube FET
1998: Proposal for Kane quantum computer
2001: 15 nm MOSFET
2003: High-performance silicon nanowire FET

molecular transistors might be seen on the horizon, but it is not obvious, in view of the predicted future capabilities of CMOS, how competitive they will be.

A recent analysis based on fundamental quantum mechanical principles, restated by George Bourianoff of the Intel Corporation, reveals that heat/power dissipation will ultimately limit any logic device using an electronic charge (Zhirnov et al. 2003) and operating at room temperature. This limit is about 100 W per square centimeter for passive cooling techniques with no active or electrothermal elements. These fundamental limits have led to pessimistic predictions of the imminent end of technological progress for the semiconductor industry and simultaneously have increased interest in advanced alternative technologies that rely on something other than electronic charge – such as spin or photon fields – to store computational state. Many advocate a focus on quantum computers that make use of distinctively quantum mechanical phenomena, such as entanglement and superposition, to perform operations on data. Among a number of quantum computing proposals, the Kane scalable quantum computer is based on an array of individual phosphorus (P) donor atoms embedded in a pure silicon lattice (Kalden et al. 2010). Both the nuclear spins of the donors and the spins of the donor electrons participate in the quantum computation. The Loss–DiVincenzo quantum computer (Loss and DiVincenzo 1998), also a scalable semiconductor-based quantum computer, makes use of the intrinsic spin degree of freedom of individual electrons confined to quantum dots as qubits.

Since the invention of the point-contact bipolar transistor in 1947, advanced fabrication technologies, introduction of new materials with unique properties, and broadened understanding of the underlying physical processes have resulted in tremendous growth in the number and variety of semiconductor devices and literally changed the world. To date, there are about 60 major devices, with over 100 device variations related to them. A list of most of the basic semiconductor devices (mainly based on Sze and May 2003) discovered and used over the past century with the date of their introduction is shown in Table 1.

Need for Simulations

Simulation is playing a key role in device development today. Two issues make simulation important (Vasileska et al. 2005). Product cycles are getting shorter with each generation, and the demand for production wafers shadows development efforts in the factory. Consider the product cycle issue first. In order for companies to maintain their competitive edge, products have to be taken from design to production in less than 18 months. As a result, the development phase of the cycle is getting shorter. Contrast this requirement with the fact that it takes 2–3 months to run a wafer lot through a factory, depending on its complexity. The specifications for experiments run through the factory must be near the final solution. While simulations may not be completely predictive, they provide a good initial guess. This can ultimately reduce the number of iterations during the device development phase.

The second issue that reinforces the need for simulation is the production pressures that factories face. In order to meet customer demand, development factories are making way for production space. It is also expensive to run experiments through a production facility. The displaced resources could have otherwise been used to produce sellable product. Again, device simulation can be used to decrease the number of experiments run through a factory. Device simulation can be used as a tool to guide manufacturing down a more efficient path, thereby decreasing the development time and costs.

Besides offering the possibility to test hypothetical devices which have not (or could not have) yet been manufactured, device simulation offers unique insight into device behavior by allowing the observation of internal phenomena that cannot be measured. Thus, 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. But the tool development itself is not enough. The tool needs to be deployed to the user community so it can be made more reliable, flexible, and accurate.

Goal of This Article

The rapid progress in nanofabrication technologies has led to the development of novel devices and structures which could revolutionize many high-technology 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. For device sizes in the range of tens of nanometers, 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.

The goal of this entry is to describe the theoretical models and the essential algorithmic and computational components that have been used in the development and deployment of the nanoelectronic modeling tool NEMO 3D on http://www.nanoHUB.org and to demonstrate successful applications of NEMO 3D in the atomistic calculation of single-particle electronic states of different, realistically sized nanostructures, each consisting of multimillion atoms. We present some of the new capabilities that have been recently added to NEMO 3D to make it one of the premier simulation tools for design and analysis of realistic nanoelectronic devices and thus 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 the NEMO 3D code is demonstrated on the IBM Blue Gene, the Cray XT3, an Intel Woodcrest cluster, and other Linux clusters. The largest electronic structure calculation, with *52 million atoms*, involved a Hamiltonian matrix with over one *billion* complex degrees of freedom. The performance impact of storing the Hamiltonian versus recomputing the matrix, when needed, is explored. We describe the state-of-the-art algorithms that have been incorporated in the code, including very effective Lanczos, block Lanczos, and Tracemin eigenvalue

solvers, and present a comparison of the different solvers. While system sizes of tens of millions of atoms appear at first sight huge and wasteful, we demonstrate that some physical problems require such large-scale analysis. We recently showed (Kharche et al. 2007) that the analysis of valley splitting in strained Si quantum wells grown on strained SiGe required atomistic analysis of ten million atoms to match experimental data. The insight that disorder in the SiGe buffer increases valley splitting in the Si quantum well would probably not be predictable in a continuum effective mass model. Similarly, the simulations of P impurities in silicon required multimillion-atom simulations (Rahman et al. 2007). In the following, we describe NEMO 3D capabilities in the simulation of different classes of nanodevices having carrier confinement in 3, 2, and 1 dimensions in the GaAs/InAs and SiGe material systems.

Single and Stacked Quantum Dots (*Confinement in Three Dimensions*). Quantum dots (QDs) are solid-state semiconducting nanostructures that provide confinement of charge carriers (electrons, holes, excitons) in all three spatial dimensions resulting in strongly localized wave functions, discrete energy eigenvalues, and interesting physical and novel device properties (Michler et al. 2000; Reed et al. 1988; Reed 1993; Bester and Zunger 2005; Petroff 2003; Moreau et al. 2001). Existing nanofabrication techniques tailor QDs in a variety of types, shapes, and sizes. Within bottom-up approaches, QDs can be realized by colloidal synthesis at benchtop conditions. Quantum dots thus created have dimensions ranging from 2 to 10 nanometers, corresponding to 100–100,000 atoms.

Self-assembled quantum dots (SAQDs) grown in the coherent Stranski–Krastanov heteroepitaxial growth mode nucleate spontaneously within a lattice-mismatched material system (e.g., InAs grown on GaAs substrate) under the influence of strain in certain physical conditions during molecular beam epitaxy (MBE) and metalorganic vapor phase epitaxy (MOVPE) (Arakawa and Sasaki 1982). The strain produces coherently strained quantum-sized islands on top of a two-dimensional wetting layer. The islands can be subsequently buried. Semiconducting QDs grown by self-assembly are of particular importance in quantum optics (Fafard et al. 1996; Maximov et al. 1998), since they can be used as detectors of infrared radiation, optical memories, and in laser applications.

The strongly peaked energy dependence of density of states and the strong overlap of spatially confined electron and hole wave functions provide ultralow laser threshold current densities, high temperature stability of the threshold current, and high material and differential quantum gain/yield. Strong oscillator strength and nonlinearity in the optical properties have also been observed (Maximov et al. 1998). Self-assembled quantum dots also have potential for applications in quantum cryptography as single-photon sources and quantum computation (Kalden et al. 2010; Chen et al. 2001). In electronic applications, QDs have been used to operate like a single-electron transistor and demonstrate a pronounced Coulomb blockade effect. Self-assembled QDs, with an average height of 1–5 nm, are typically of size (base length/diameter) 5–50 nm and consist of 5,000–2,000,000 atoms. Arrays of quantum mechanically coupled (stacked) self-assembled quantum dots can be used as optically active regions in high-efficiency, room-temperature lasers. Typical QD stacks consist of 3–7 QDs with typical lateral extension of 10–50 nm and dot height of 1–3 nm. Such dots contain 5–50 million atoms in total, where atomistic details of interfaces are extremely important (Usman et al. 2006).

Impurities (*Confinement in Three Dimensions*). Impurities have always played a vital role in semiconductors since the inception of the transistor. Till the end of last century, scientists and engineers had been interested in the macroscopic properties of an ensemble of dopants in a semiconductor. As technology enters the era of nanoscale electronics, devices which contain a few discrete dopants are becoming increasingly common. In recent years, there have been proposals of novel devices that operate on purely quantum mechanical principles using the quantum states of isolated or coupled donors/impurities (Kalden et al. 2010; Vrijen et al. 2000; Hollenberg et al. 2004). The ongoing extensive research effort on the phosphorus (P) donor-based quantum computer architecture of Kane (Kalden et al. 2010) exemplifies an effort to harness the quantum nature of materials for the development of next-generation

electronics. As researchers strive to establish atomic-scale quantum control over single impurities (Sellier et al. 2006; Bradbury et al. 2006; Stegner et al. 2006), precision modeling techniques are required to explore this new regime of device operations (Debernardi et al. 2006; Martins et al. 2004; Rahman et al. 2007; Friesen et al. 2005).

Although effective mass-based approaches have been predominantly used in literature to study the physics of impurities, realistic device modeling using this technique has proved difficult in practice. Tight-binding methods (Slater and Koster 1954) consider a more extensive Bloch structure for the host material and can treat interfaces, external gates, strain, magnetic fields, and alloy disorder within a single framework. When applied to realistic nanodevices of several million atoms, this technique can prove very effective for device modeling (Klimeck et al. 2002). We present a semiempirical method for modeling impurities in Si that can be used for a variety of applications such as *quantum computer* architecture, discretely doped FinFETs, and impurity scattering problems. Although we focus on P impurities in Si here, the method is sufficiently general to be used on other impurities and hosts.

Quantum Wires (*Confinement in Two Dimensions*). For quite some time, nanowires have been considered a promising candidate for future building block in computers and information processing machines (Klimeck et al. 2006; Wang et al. 2005; Zheng et al. 2005; Luisier et al. 2006; Boykin et al. 2007a). Nanowires are fabricated from different materials (metal, semiconductor, insulator, and molecular) and assume different cross-sectional shapes, dimensions, and diameters. Electrical conductivity of nanowires is greatly influenced by edge effects on the surface of the nanowire and is determined by quantum mechanical conductance. In the nanometer regime, the impact of surface roughness or alloy disorder on electronic bandstructure must be atomistically studied to further gauge the transport properties of nanowires.

Quantum Wells (*Confinement in One Dimension*). QW devices are already a de facto standard technology in MOS devices and QW lasers. They continue to be examined carefully for ultra-scaled devices where interfacial details turn out to be critical. Composite channel materials with GaAs, InAs, InSb, GaSb, and Si are being considered (Rahman et al. 2005; Prada et al. 2007), which effectively constitute QWs. Si QWs buffered/strained by SiGe are considered for quantum computing (QC) devices where valley splitting (VS) is an important issue (Eriksson et al. 2004). Si is desirable for QC due to its long spin decoherence times, scaling potential, and integrability within the present microelectronic manufacturing infrastructure. In strained Si, the sixfold valley degeneracy of Si is broken into lower twofold and raised fourfold valley degeneracies. The presence of twofold valley degeneracy is a potential source of decoherence which leads to leakage of quantum information outside qubit Hilbert space. Therefore, it is of great interest to study the lifting of the remaining twofold valley degeneracy in strained Si due to sharp confinement potentials in recently proposed (Eriksson et al. 2004) SiGe/Si/SiGe quantum well (QW) heterostructure-based quantum computing architectures.

Nanoscale Device Modeling and Simulation Challenges

The theoretical knowledge of the electronic structure of nanoscale semiconductor devices is the first and most essential step toward the interpretation and the 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.

1. *Full Three-Dimensional Atomistic Representation*: The lack of *spatial symmetry* in the overall geometry of the nanodevices usually requires explicit three-dimensional representation. For example, Stranski–Krastanov growth techniques tend to produce self-assembled InGaAs/GaAs quantum dots

(Michler et al. 2000; Petroff 2003; Reed 1993; Reed et al. 1988) with some rotational symmetry, e.g., disks, truncated cones, domes, or pyramids (Bester and Zunger 2005). These structures 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 (Pryor et al. 1998) and ***k•p*** (Grundmann et al. 1995; Stier et al. 1999) typically ignore such crystal symmetry and atomistic resolution.

The required simulation domain sizes of ∼1 M atoms prevent the usage of ab initio methods. Empirical methods which 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 (Calderón et al. 2006) and (2) tight binding (Klimeck et al. 2002). 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 hetero-interfaces. This means that the basis needs to be adjusted (by an expert) for every different device, which limit 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 hetero-interfaces, and it results in very sparse matrices. The requirements of storage and processor communication are therefore minimal compared to pseudopotentials, and actual implementations perform extremely well on inexpensive clusters (Klimeck et al. 2002).

Tight binding has the disadvantage that it is based on empirical fitting and some in the community continue to question 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 (Boykin et al. 2002), electromagnetic fields (Boykin et al. 2001), and Coulomb matrix elements (Lee et al. 2002) and fit them to well-known and accepted bulk parameters (Klimeck et al. 2000a, b, 2002). With tight binding, the NEMO team was able early on to match experimentally verified, high-bias current–voltage curves of resonant tunneling (Bowen et al. 1997; Klimeck et al. 1997) 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 we are in the process of benchmarking it against more fundamental ab initio approaches and pseudopotential approaches. Our current Si/Ge parameterization is described in references (Boykin et al. 2004a, 2007b). Figure 1 depicts a range of phenomena that represent new challenges presented by new trends in nanoelectronics and lays out the NEMO 3D modeling agenda.

2. *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 quantum dots, this mismatch is around 7 % and leads to a strong *long-range* strain field within the extended neighborhood (typically ∼25 nm) of each quantum dot (Ahmed et al. 2007). Si/Ge core/shell structured nanowires are another example of strain-dominated atom arrangements (Liang et al. 2006). Si quantum wells and SiGe quantum computing architectures rely on strain for state separation (Eriksson et al. 2004). The strain can be atomistically inhomogeneous, involving not only biaxial components but also non-negligible shear components. Strain strongly influences the core and barrier material bandstructures, 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

DOI 10.1007/978-3-642-27737-5_343-2

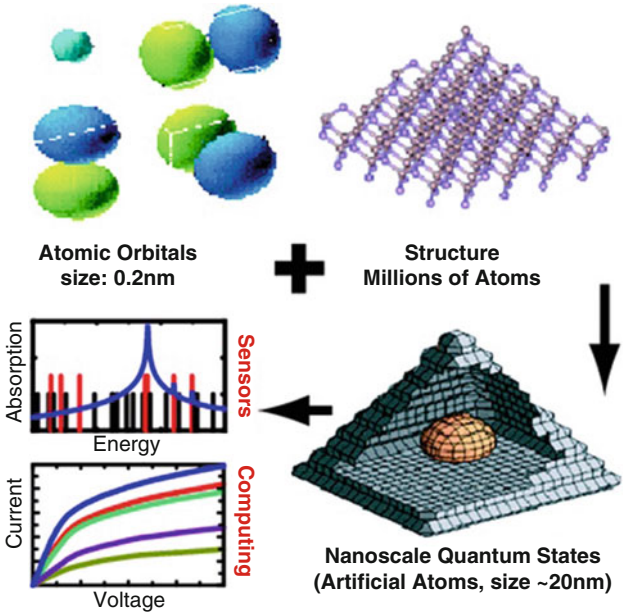


Fig. 1 NEMO 3D modeling agenda: map electronic properties of individual atoms into realistic structures containing millions of atoms, computation of nanoscale quantum dots that map into real applications

carriers and the long-distance atomistic strain effects with proper boundary conditions on equal footing (Williamson et al. 2000; Lazarenkova et al. 2004).

3. *Piezoelectric Field*: A variety of III–IV materials such as GaAs, InAs, and GaN, are piezoelectric. Any spatial nonsymmetric distortion in nanostructures made of these materials will create piezoelectric fields, which will modify the electrostatic potential landscape. Recent spectroscopic analyses of self-assembled QDs demonstrate polarized transitions between confined hole and electron levels (Bester and Zunger 2005). While the continuum models (effective mass or $\boldsymbol{k \cdot p}$) can reliably predict aspects of the single-particle energy states, they fail to capture the observed non-degeneracy and optical polarization anisotropy of the excited energy states in the (001) plane. These methods fail because they use a confinement potential which is assumed to have only the *shape symmetry* of the nanostructure, and they ignore the underlying crystal symmetry. The experimentally measured symmetry is significantly lower than the assumed continuum symmetry because of (a) underlying crystalline symmetry, (b) atomistic strain relaxation, and (c) piezoelectric field. For example, in the case of pyramid-shaped quantum dots with square bases, continuum models treat the underlying material in C_{4v} symmetry, while the atomistic representation lowers the crystal symmetry to C_{2v}. The 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.

NEMO 3D Simulation Package

Basic Features: Simulation Domains

NEMO 3D (Klimeck et al. 2002; Korkusinski and Klimeck 2006; Oyafuso et al. 2003a, b; Korkusinski et al. 2005b) bridges the gap between the large size, classical semiconductor device models and the molecular level modeling. This package currently allows calculating single-particle electronic states and optical response of various semiconductor structures including bulk materials, quantum dots, quantum

DOI 10.1007/978-3-642-27737-5_343-2

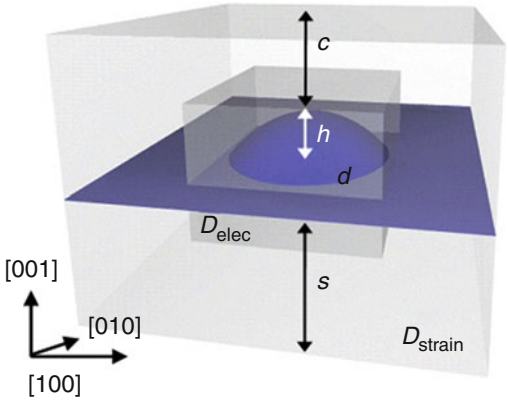


Fig. 2 Simulated dome-shaped InAs quantum dot buried in GaAs. Two simulation domains are shown: D_{elec}, central smaller domain for electronic structure calculation, and D_{strain}, outer larger domain for strain calculation. In the figure, s is the substrate height, c is the cap layer thickness, h is the dot height, d is the dot diameter

wires, quantum wells, and nanocrystals. NEMO 3D 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 (Klimeck et al. 2002; Graf and Vogl 1995; Boykin et al. 2001). In a schematic view of InAs quantum dot embedded in a GaAs barrier material, the sample is presented in Fig. 2. The quantum dot 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_{strain}, while the electronic structure computation is restricted to the smaller domain D_{elec}. 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 3D enables the computation of strain and electronic structure in an atomistic basis for over 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 any closed 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 3D runs on serial and parallel platforms, local cluster computers, as well as the NSF TeraGrid.

Components and Models

The NEMO 3D 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; stored are also its type (atomic number in short integer), 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 CPUs, i.e., 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 do 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.
2. *Strain.* The materials making up the QD nanostructure may differ in their lattice constants; for the InAs/GaAs system, this difference is of 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 3D computes strain field using an atomistic valence force field (VFF) method (Keating 1966) 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

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

where the sum is carried out over the n nearest neighbors j of atom i; $\vec{d}_{ij}$ and $\vec{R}_{ij}$ are the bulk and actual (distorted) distances between neighbor atoms, respectively; and α_{ij} and β_{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 (Williamson et al. 2000; Lazarenkova et al. 2004) are also implemented in NEMO 3D. 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:

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

where $c_{i,v}^{+}$ ($c_{i,v}$) is the creation (annihilation) operator of an electron on the orbital v localized on atom i. In the above equation, the first term describes the on-site orbital terms, found on the diagonal of the Hamiltonian matrix. The second term describes coupling between different orbitals 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 the atoms i and j be nearest neighbors.

The characteristic parameters ε 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 σ and π bonds between the atomic orbitals. For example, for a simple cubic lattice, the interaction between the s orbital localized on the atom i at origin and the orbital p_x localized on the atom j with coordinate $\vec{d}_{ij} = a\hat{x}$ with respect to the atom i would simply be expressed as $t_{ij}^{(s,p_x)} = V_{sp\sigma}$. 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 3D vector $\vec{d}_{ij} = l\hat{x} + m\hat{y} + n\hat{z}$, with l, m, and n being the directional cosines. These cosines rescale the interaction constants, so that the element describing the interaction of the orbitals s and p_x is $t_{ij}^{(s,p_x)} = lV_{sp\sigma}$. The parameterization of all bonds using analytical forms of directional cosines for various tight-binding models is given in (Slater and Koster 1954). NEMO 3D provides the user with choices of the $sp^3d^5s^*$, sp^3s^*, and single s orbital models with and without spin, in zinc blende, wurtzite, and simple cubic lattices.

DOI 10.1007/978-3-642-27737-5_343-2

Additional complications arise in strained structures, where the atomic positions deviate from the ideal (bulk) crystal lattice (Jancu et al. 1998). The presence of strain leads to distortions not only of bond directions, but also bond lengths. In this case, the discussed interaction constant $t_{ij}^{(s,p_x)} = l' V_{sp\sigma} \left(\frac{d}{d_0}\right)^{\eta(sp\sigma)}$, 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 available data. The energy constants parameterizing the on-site interaction change as well due to bond renormalization (Klimeck et al. 2002; Boykin et al. 2002).

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 unit cell only. The model makes it possible to compute the bandstructure 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 and along the Δ line from Γ to X 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, described in detail in Klimeck et al. (2002) and Graf and Vogl (1995). 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. *Post Processing of Eigenstates.* From the single-particle eigenstates, various physical properties can be calculated in NEMO 3D such as optical matrix elements (Boykin and Vogl 2001), Coulomb and exchange matrix elements (Lee et al. 2002), and approximate single-cell bandstructures from supercell bandstructure (Boykin and Klimeck 2005; Boykin et al. 2007a, c).

Algorithmic and Numerical Aspects

1. *Parallel Implementation*. The complexity and generality of physical models in NEMO 3D can place high demands on computational resources. For example, in the 20-band electronic calculation, the discrete Hamiltonian matrix is of order 20 times the number of atoms. Thus, in a computation with 20 million atoms, the matrix is of order 400 million. Computations of that size can be handled because of the parallelized design of the package. NEMO 3D 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 slabs along one dimension. 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

DOI 10.1007/978-3-642-27737-5_343-2

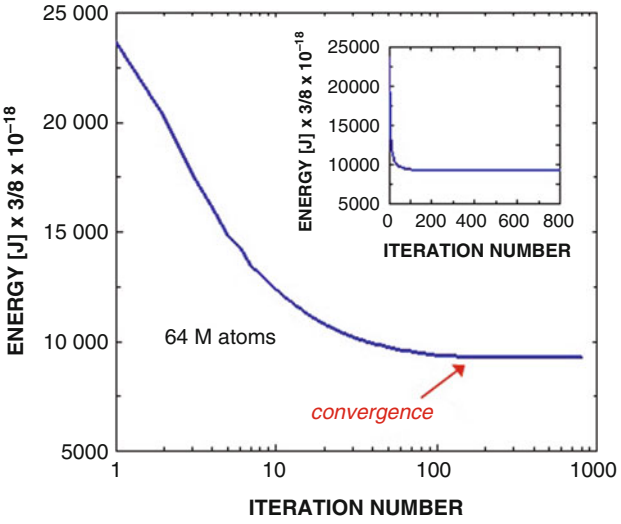


Fig. 3 Elastic energy convergence profile in a typical simulation of an InAs/GaAs quantum dot with a total 64 million of atoms (*inset* – linear scale)

implemented such that all inter-processor communications can be performed in two steps (Klimeck et al. 2002).

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. Figure 3 shows the energy convergence behavior in a typical simulation of an InAs/GaAs quantum dot with a total of around 64 million atoms. 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 3D requires six vectors, each of the total size of 3× number of atoms (to store atomic coordinates, gradients, and intermediate data); however, all those vectors are divided into slabs and distributed among CPUs as discussed above. The final atom position vectors are by default stored on all the CPUs for some technical output details.

   The electronic computation involves a very large eigenvector computation (matrices of order of hundreds of millions or even billion). The algorithms/solvers available in NEMO 3D include the PARPACK library (Maschhoff and Sorensen 1996), a custom implementation of the Lanczos method, block Lanczos method, the spectrum folding method (Wang and Zunger 1994), and the Tracemin method (Sameh and Tong 2000). The research group is also exploring implementations of Lanczos with deflation method.

   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 quantum 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 Lanczos algorithm is that it is fast, while 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 by a factor of 40 for the NEMO 3D matrix than PARPACK. Block Lanczos with block

DOI 10.1007/978-3-642-27737-5_343-2

Table 2 Performance comparison of different eigenvalue solvers on 32 processors of Purdue University Linux cluster (Xeon x86-64 Dual Core 2.33GHz). Simulation was performed on an InAs QD structure with 268800 atoms. Time (in hours), relative time, number of matrix–vector products (#MVP), relative matrix–vector products, memory (in GB), and number of correct eigenvalues and their multiplicity (#Eig(mul)) for Lanczos, block Lanczos with block size 2 (BLanczos2), PARPACK, Tracemin with quadratic mapping (QTracemin), and Tracemin with Chebyshev polynomial mapping (CTracemin)

Algorithm	Time (h)	Relative time	#MVP (×1,000)	Relative MVP	Memory (GB)	#Eig(mul)
Lanczos	0.428	1.0	10.9	1.0	2.64	20(1)
BLanczos2	1.385	3.2	11.8	1.1	2.77	8(2)
PARPACK	18.04	42.2	59.3	5.4	2.64	8(2), 4(1)
QTracemin	15.71	36.7	317.0	29.1	2.77	10(2)
CTracemin	13.70	32.1	528.8	48.5	2.64	10(2)

Table 3 List of spectrum between 1.0 ∼ 1.3 eV and the number of multiplicities obtained from different solvers. *Number* of searched eigenvalues was kept constant for these methods

Eigenvalues	Lanczos	BLanczos2	PARPACK	QTracemin	CTracemin
1.0361	1	-	-	2	2
1.0969	1	2	-	2	2
1.0976	1	2	1	2	2
1.1624	1	2	2	2	2
1.1645	1	2	2	2	2
1.1748	1	2	2	2	2
1.2304	1	2	2	2	2
1.2312	1	2	2	2	2
1.2445	1	2	2	2	2
1.2448	1	-	2	2	2
1.2975	1	-	2	-	-

size p finds p degenerate eigenvalues relatively fast compared to PARPACK and Tracemin; however, a potential instability exists as well. The 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 Tracemin and missing some of the eigenvalues and their multiplicity. Tables 2 and 3 give a comparison of Lanczos, block Lanczos, PARPACK, and Tracemin with the number of eigenvalues searched 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 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 to not 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 on memory-poor platforms like an IBM Blue Gene.

Figure 4 shows the memory requirements for the dominant phase of the code (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. 4, on the vertical axis on the right side of each plot. This figure shows that the number of atoms that can be

DOI 10.1007/978-3-642-27737-5_343-2

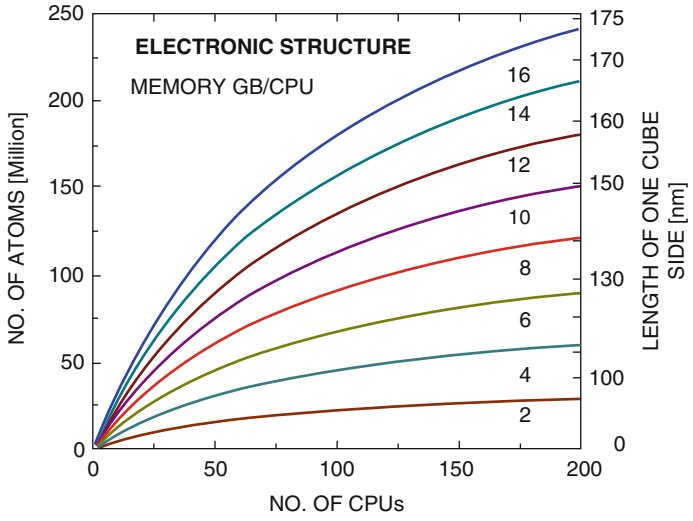


Fig. 4 Number of atoms that can be treated, as a function of the number of CPUs for different amounts of memory per CPU for the electronic structure calculation. The vertical axis on the right side of each plot gives the equivalent length in nm of one side of the cube that would contain the given number of atoms

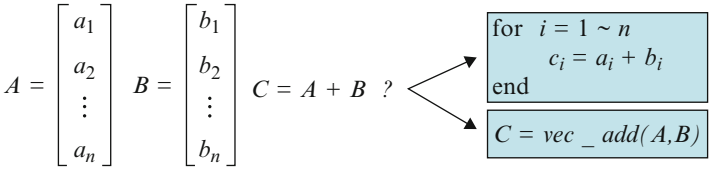


Fig. 5 The conceptual diagram of vectorization. In vectorized CPU, it is capable of n simultaneous operations in a single CPU cycle

treated in NEMO 3D continues to grow for larger CPU counts. The strain calculations have so far never been memory limited. NEMO 3D is typically size limited in the electronic structure calculation.

3. *Optimization in NEMO 3D*. In running a scientific application that requires massive computation power, we have to consider various issues that may occur, mainly due to limited resource in a computer: Too small memory per core can limit the size of the problem, and unnecessary loops in the code consume additional time for calculation. It is crucial to design an application in a way to maximize floating operations per second and avoid inefficient loops. In NEMO 3D, several optimization ideas are implemented, and those are introduced in the following sections.
   (a) *Vectorization*. Vectorization is a hardware-dependent optimization scheme that converts multiple single scalar operations to single vector operation. The concept is shown in Fig. 5. It is commonly used in graphic processors and supercomputers (e.g., Cray X1E machines) where massive computation load and fast processing are needed. Even recent processors in desktop computers support similar parallel data processing scheme. The most common technique to support parallelism is single instruction, multiple data (SIMD) algorithm. It was Intel who first developed instruction sets known as Streaming SIMD Extensions, or SSE, to support in their Pentium III processors in 1999 http://www.intel.com/cd/software/products/asmo-na/eng/307757.htm. Nowadays, AMD, Transmeta, and Via also support SSE features, and new enhancements are developed continuously (as of Oct. 2007, SSE5 is the latest version). A couple of single- and double-precision arithmetic can be carried out simultaneously resulting in fast computation. Therefore, it is possible to make use of SSE scheme in scientific applications with heavy complex number calculations. In NEMO

DOI 10.1007/978-3-642-27737-5_343-2

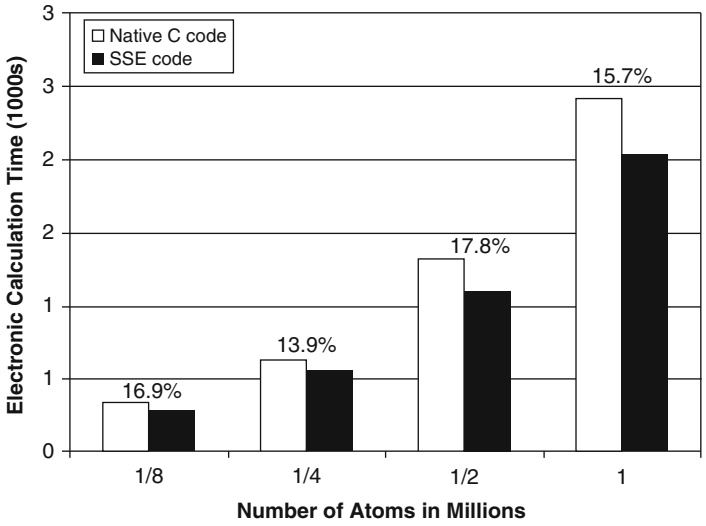


Fig. 6 Comparison of electronic calculation time between SSE-optimized code and native C code. Simulated on a single node of Xeon x86-64 Dual Core 2.33GHz CPU computers

3D, complex multiplication and addition occurs frequently in matrix–matrix multiply routine. To this certain application, major improvement was achieved in real-complex multiplies. Figure 6 shows the speed improvement observed in NEMO 3D by replacing SSE instructions to real-complex multiplication.

(b) *Matrix–Matrix Multiplier and BLAS*. The Basic Linear Algebra Subprograms, or BLAS, are standardized interface for performing basic matrix–vector and matrix–matrix multiplication. The BLAS package is widely used in high-performance computing, and it has been optimized to maximize the number of floating-point operations for specific CPUs. For example, Intel develops its own BLAS package in the Math Kernel Library (MKL BLAS) highly optimized to their processors. Compared to native C code with double-nested loops, benefits can be made from BLAS, especially with matrix–matrix multiplication. From the experiment shown in Fig. 7, highly optimized BLAS matrix–matrix multiply instruction, or ZGEMM, is capable of utilizing the CPU to perform more floating-point operations per second, reducing the total calculation time. Even for the block sizes $N = 10$, $N = 20$ corresponding to $sp^3d^5s^*$ bands, significant improvement can be seen by performing block-wise operations. The data in Fig. 7 indicates an excellent incentive for the block Lanczos and the Tracemin algorithms that perform multiple matrix–vector multiplies for the same matrix to be blocked. For example, at $N = 10$, a single vector multiply can be performed at about 1.5 GFlops, while eight multiplies can be performed at a rate of 3.6 GFlops. With the increase in relative performance for increased block size, the required total CPU time increased sublinearly. Subsequent NEMO 3D development for general 3D spatial structures will utilize the ZGEMM multiply by arranging the data structures such that no copy is needed.

(c) *Explicit Construction of Hamiltonian in Recompute Mode*. The recompute mode enables NEMO 3D to run on limited memory computers by eliminating storage of the Hamiltonian altogether and recomputing the matrix elements as they are needed. However, since the construction of the Hamiltonian consumes significant time, reducing the number of calculations in the construction of a matrix element enhances the performance. In cases where no external magnetic field is present, duplicate calculations due to the spin degeneracy can be avoided. Also, since the orbital interactions are known, unnecessary loops can be avoided, and nonzero elements may be explicitly evaluated. The doubly nested switch statements at the core of the orbital–orbital interaction loops

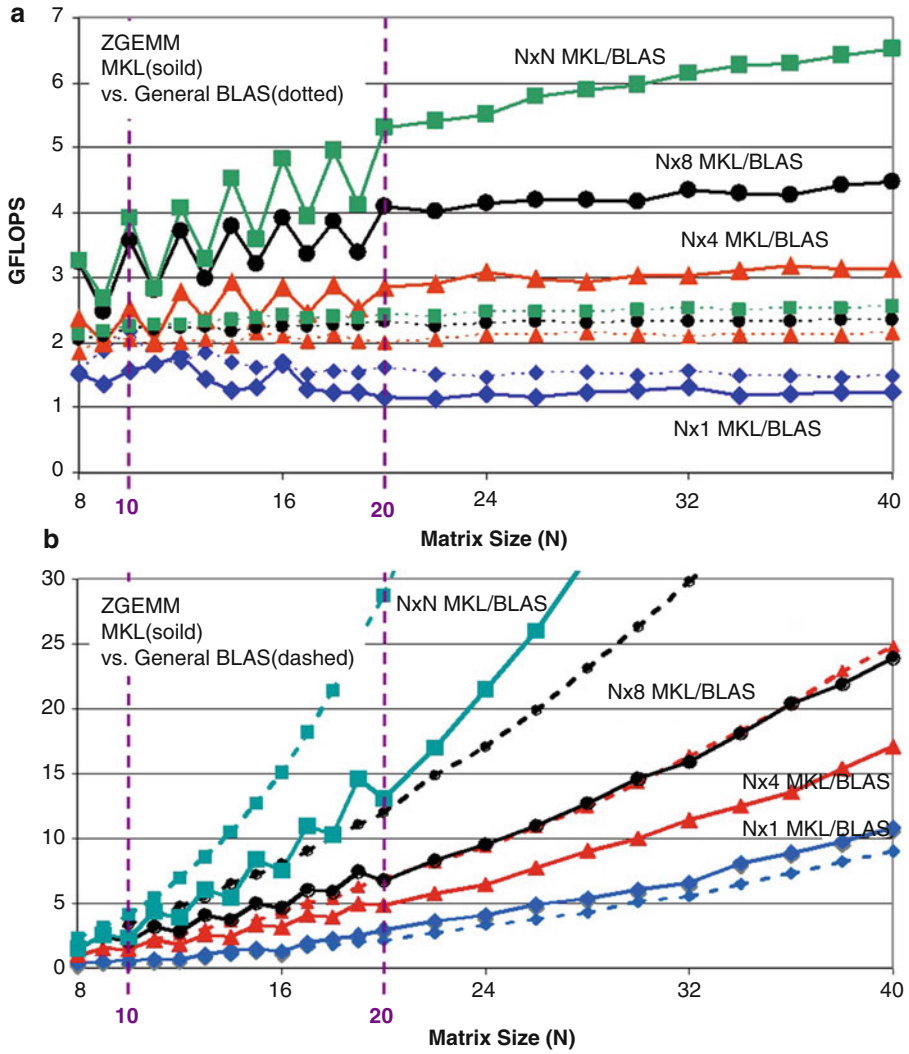


Fig. 7 (**a**) Performance plots of ZGEMM (Y = AX) included in different BLAS libraries. GFLOPS (10^9 floating operations/second) measures of ZGEMM from MKL/BLAS (*solid line*) and general BLAS/LAPACK library (open markers) are plotted with varying size of $A(N \times N)$ and column size of $X(N \times M)$. Simulated on a single node of Xeon x86-64 Dual Core 2.33 GHz CPU computer. (**b**) Total compute time of data in (**a**)

have been replaced by customized expressions for the matrix elements for specific tight-binding orbital arrangements such as sp^3s^* and $sp^3d^5s^*$. Simulation result indicates that the electronic calculation time is reduced up to 40 % (Fig. 8). This customization increases computational performance but reduces the algorithmic generality.

4. *Scaling*. Out of the two phases of NEMO 3D, the strain calculation is algorithmically and computationally less challenging than the Lanczos diagonalization of the Hamiltonian matrix.

To investigate the performance of NEMO 3D package, computation was performed in a single dome-shaped InAs quantum dot nanostructure embedded in a GaAs barrier material as shown in Fig. 2. The HPC platform used in the performance studies are shown in Table 4. These include a Linux cluster at the Rosen Center for Advanced Computing (RCAC) at Purdue with Intel processors (dual core Woodcrest). The other five platforms are a Blue Gene at the Rensselaer Polytechnic Institute (RPI), the Cray XT3 at the

DOI 10.1007/978-3-642-27737-5_343-2

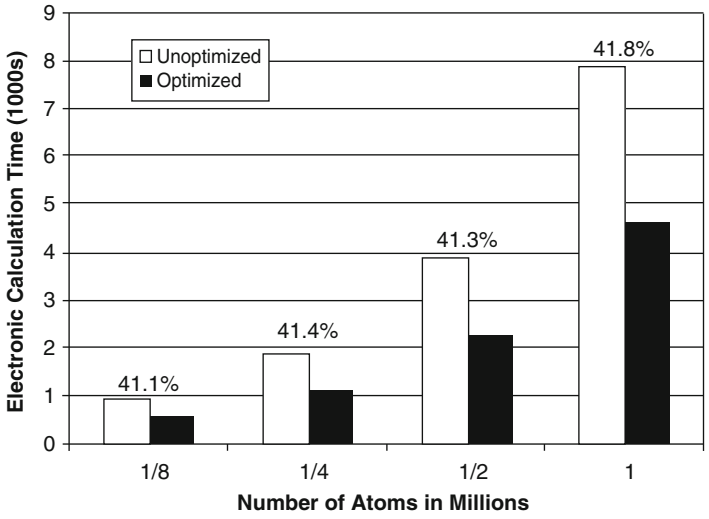


Fig. 8 The electronic calculation time comparison of optimized/unoptimized Hamiltonian construction in recompute mode. Simulated on four nodes of Xeon x86-64 Dual Core 2.33GHz CPU computers

Table 4 Specifications for the HPC platforms used in the performance comparisons

Platform	Type	CPU	No. of Cores	Memory/ core	Interconnect	Top 500 June 2007	Location
ORNL/ Jaguar	Cray XT3/4	Opteron x86-64 2.6GHz	23,016	2GB	Native	#2	ORNL
RPI/BGL	Blue Gene/L	PowerPC 440 0.7 GHz	32,768	256 MB	Native	#7	RPI
IUPU/Big Red	IBM JS21	PowerPC 970 2.5 GHz	3,072	2GB	Myrinet	#8	IUPU
PSC/XT3	Cray XT3	Opteron x86-64 2.6GHz	4,136	1GB	Native	#30	PSC
PU/Xeon D	Linux cluster	Xeon x86-64 dual core 2.33 GHz	672	2GB/ 4GB	Gigabit ethernet	#46	RCAC Purdue

Pittsburgh Supercomputing Center (PSC), the Cray XT3/4 at ORNL, JS21 at Indiana University, and a Woodcrest machine at NCSA. Table 4 provides the relevant machine details. These platforms have proprietary interconnects that are higher performance than Gigabit Ethernet (GigE) for the three Linux clusters at Purdue. In the following, the terms processors and cores are used interchangeably.

Figure 9 shows the performance of NEMO 3D for each of the architectures. The wall clock times for 500 iterations of the Lanczos method for the electronic structure phase are shown as a function of the number of cores. The benchmark problem includes eight million atoms. Figure 9 shows that the PU/Woodcrest cluster is close to the performance of the Cray XT3 for lower core counts, while the XT3 performs better for higher core counts, due to its faster interconnect. The Blue Gene's slower performance is consistent with its lower clock speed, while the scalability reflects its efficient interconnect.

Recomputing the Hamiltonian causes a performance reduction of about a factor of 4–6. Since the IBM Blue Gene L is memory poor, we can operate NEMO 3D only in the Hamiltonian recomputed mode. Since the IBM Blue Gene runs about a factor of 4× slower than the other HPC platforms, one can see about a factor of 16× better performance in Cray XT3/4 since it runs fast and has enough memory.

DOI 10.1007/978-3-642-27737-5_343-2

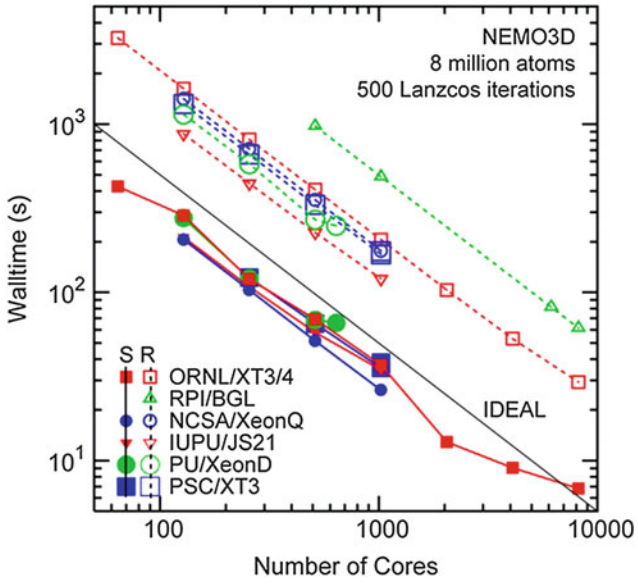


Fig. 9 Strong scaling of a constant problem size (eight million atoms) on six different HPC platforms. *Solid/dashed lines* correspond to a stored/recomputed Hamiltonian matrix. The largest number of cores available was 8,192 on Cray XT3/4 and IBM Blue Gene

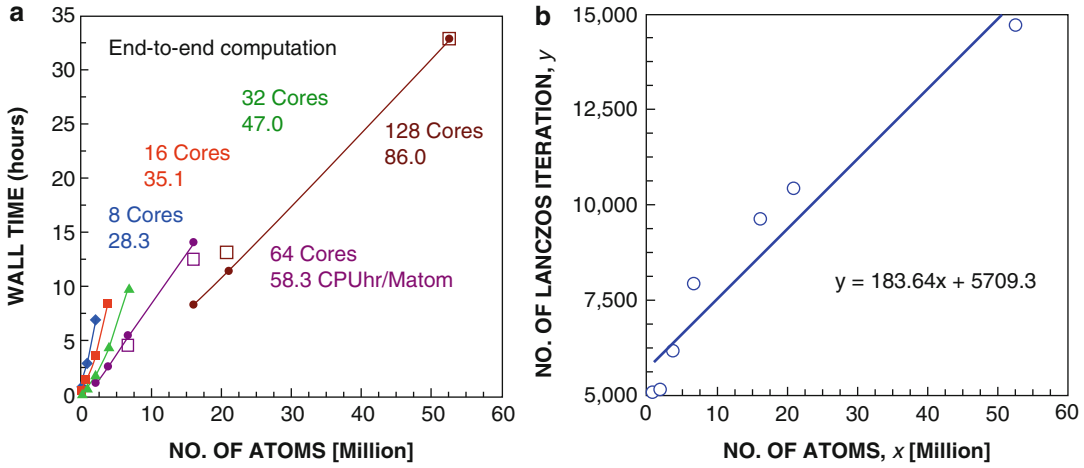


Fig. 10 (**a**) Wall clock time versus number of atoms for end-to-end computations of the electronic structure of a quantum dot, for various numbers of cores on the PU/Woodcrest cluster. Listed next to the number of cores are the CPU hours/million of atoms needed in the simulation. (**b**) Number of Lanczos iteration versus number of atoms for one choice of number of cores

In addition to the performance for the benchmark cases, end-to-end runs on the PU/Woodcrest cluster are carried out next (Fig. 10). This involves iterating to convergence and computing the eigenstates in the desired range (four conduction band and four valence band states). For each problem size, measured in millions of atoms, the end-to-end cases were run to completion, for one choice of number of cores.

The numerical experiment is designed to demonstrate NEMO 3D's ability to extract targeted interior eigenvalues and vectors out of virtually identical systems of increasing size. A single dome-shaped InAs quantum dot 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 (Bae et al. 2007) that the eigenvectors retain the expected symmetry of the nanostructure.

DOI 10.1007/978-3-642-27737-5_343-2

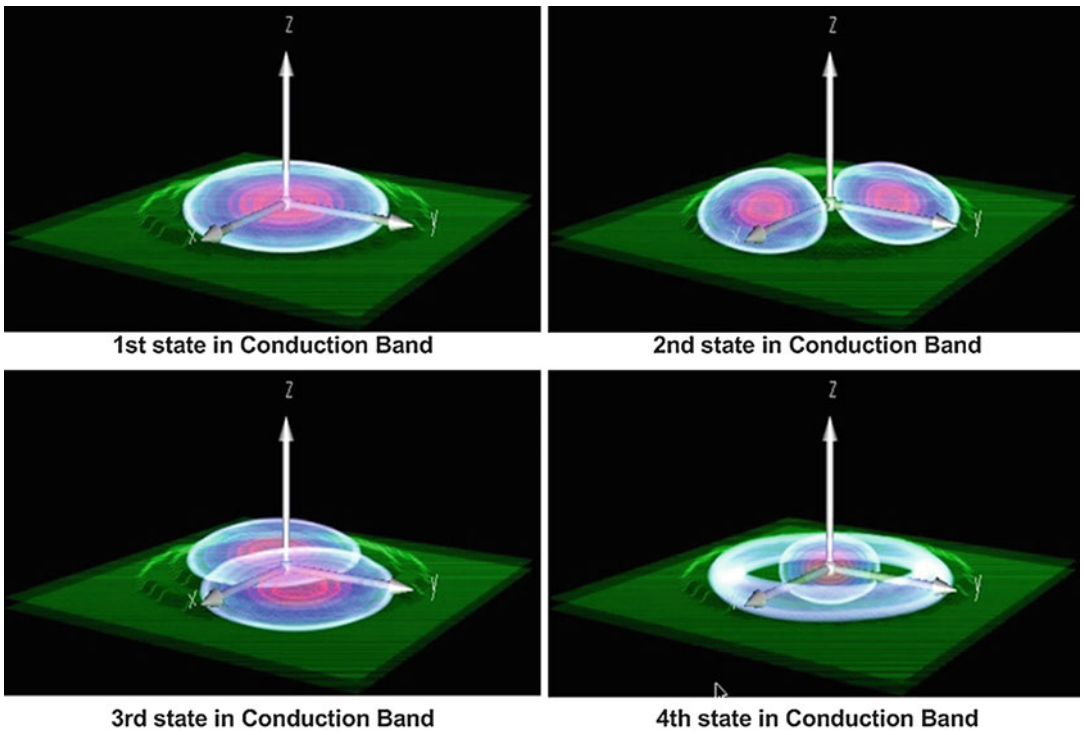


Fig. 11 Wave function profiles of first four-electron eigenstates in the conduction band. *Green* color shows active InAs region where confinement takes place

Visualization

The simulation data of NEMO 3D contains multivariate wave functions and strain profiles of the device structure. For effective 3D visualizations of these results, a hardware-accelerated direct volume rendering system (Qiao et al. 2006) has been developed, which is combined with a graphical user interface based on *Rappture*. Rappture is a toolkit that supports and enables the rapid development of *graphical user interfaces* (GUIs) for applications, which is developed by Network for Computational Nanotechnology at Purdue University. Two approaches can be followed: (1) The legacy application is not modified at all and a *wrapper script* translates Rappture I/O to the legacy code. (2) Rappture is integrated into the source code to handle all I/O. The first step is to declare the parameters associated with one's tool by describing Rappture objects in the Extensible Markup Language (XML). Rappture reads the XML description for a tool and generates the GUI automatically. The second step is that the user interacts with the GUI, entering values, and eventually presses the Simulate button. At that point, Rappture substitutes the current value for each input parameter into the XML description and launches the simulator with this XML description as the driver file. The third step shows that, using parser calls within the source code, the simulator gets access to these input values. Rappture has parser bindings for a variety of programming languages, including C/C++, Fortran, Python, and MATLAB. And finally, the simulator reads the inputs, computes the outputs, and sends the results through run file back to the GUI for the user to explore. The visualization system uses data set with *OPEN-DX* format that are directly generated from NEMO 3D. *OPEN-DX* is a package of open-source visualization software based on IBM's Visualization Data Explorer. Figure 11 shows the wave functions of electron on the first four eigenstates in conduction band of quantum dot which has 268,800 atoms in the electronic domain.

DOI 10.1007/978-3-642-27737-5_343-2

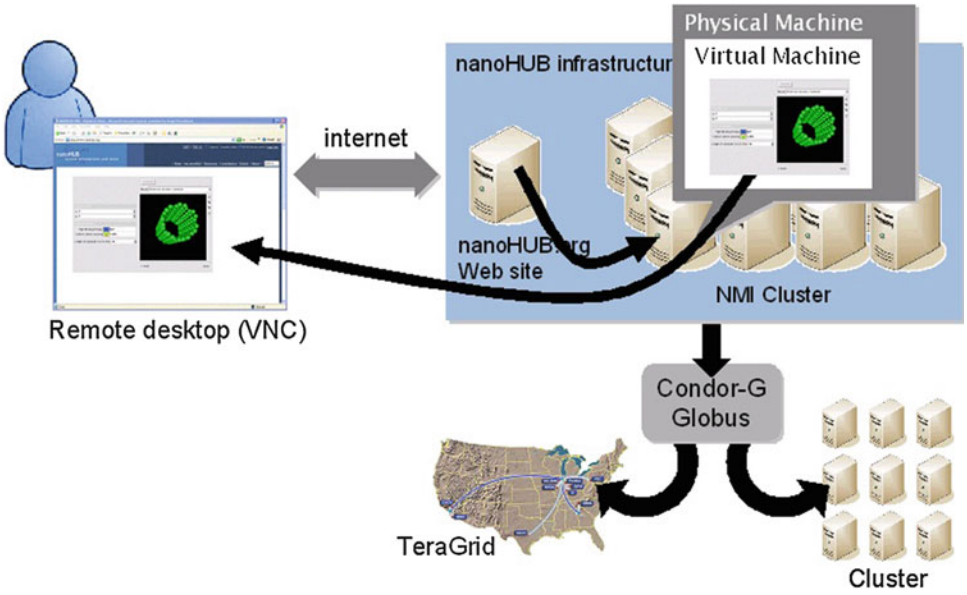


Fig. 12 Deployment of the NCN nanotechnology tools on http://www.nanoHUB.org: remote access to simulators and compute power

Release and Deployment of NEMO 3D Package

NEMO 3D was developed on Linux clusters at the Jet Propulsion Lab (JPL) and was released with an open-source license in 2003. The originally released source is hosted at http://www.openchannelfoundation.org Web site. As NEMO 3D is undergoing further developments by the NCN, we are planning future releases of the NEMO 3D source through http://www.nanoHUB.org. NEMO 3D has been ported to different high-performance computing (HPC) platforms such as the NSF's TeraGrid (the Itanium2 Linux cluster at NCSA), Pittsburgh's Alpha cluster, Cary XT3, SGI Altix, IBM p690, and various Linux clusters at Purdue University and JPL.

The NEMO 3D project is now part of a wider initiative, the NSF Network for Computational Nanotechnology (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. The second purpose of NCN is to provide a comprehensive suite of nano simulation 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 graphical user interface (GUI) in the remote desktop (VNC) 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. The nanoHUB is currently considered one of the leaders in science gateways and cyberinfrastructure.

The process of Web-based deployment of these tools is depicted in Fig. 12. A user visits the www.nanohub.org site and finds a link to a tool. Clicking on that link will cause our middleware to create a

DOI 10.1007/978-3-642-27737-5_343-2

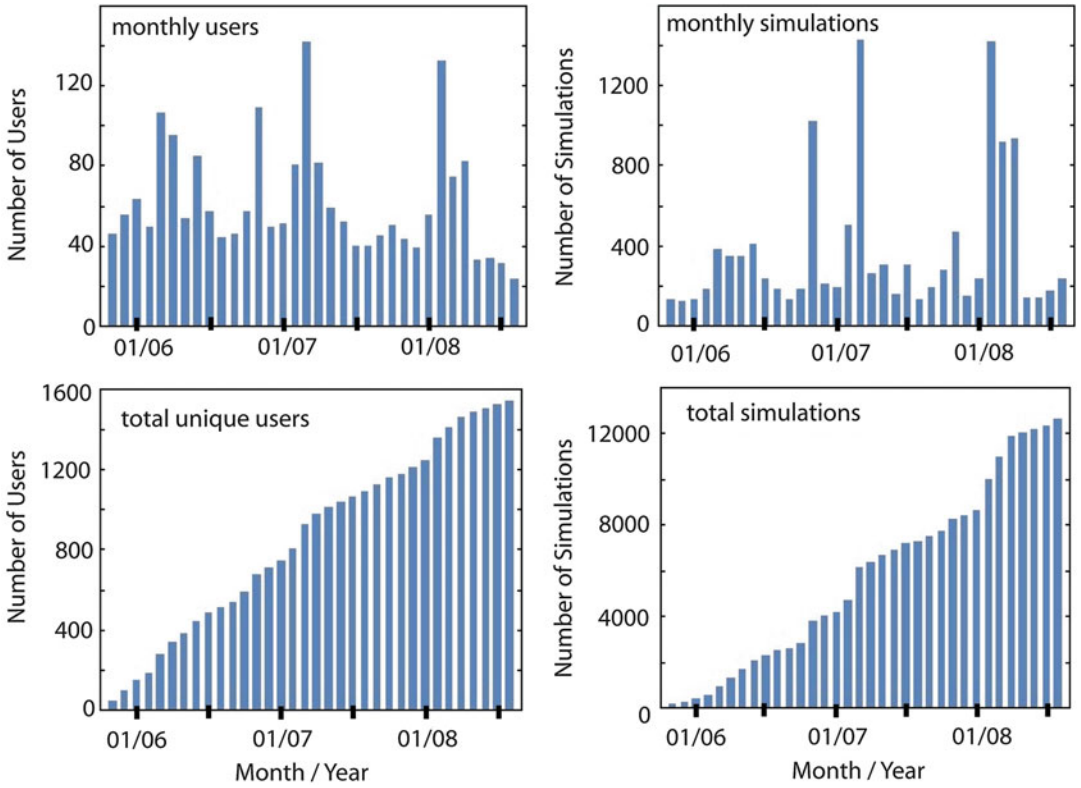


Fig. 13 (*First row*) Number of *monthly* users who have run at least one simulation and number of *monthly* simulation runs executed by nanoHUB users. (*Second row*) Number of *total* users who have run at least one simulation and *total* simulation runs executed by nanoHUB users

virtual machine running on some available CPU. This virtual machine gives the user his/her own private file system. The middleware starts an application and exports its image over the Web to the user's browser. The application looks like an applet running in the browser. The user can click and interact with the application in real time taking advantage of high-performance distributed computing power available on local clusters at Purdue University and on the NSF TeraGrid or the Open Science Grid.

Recently, a prototype graphical user interface (GUI) based on the *Rappture* package (http://www.rappture.org) is incorporated within the NEMO 3D package and a Web-based online *interactive* version (Quantum Dot Lab) for educational purposes is freely available on www.nanohub.org (Klimeck et al. 2008). The currently deployed NEMO 3D educational version is restricted to a single *s* orbital basis (single-band effective mass) model and runs in seconds. Users can generate and freely rotate 3D wave functions interactively powered by a remote visualization service. Quantum Dot Lab was deployed in November 2005 and has been a popular tool used by 4,276 users who ran 47,402 simulations up to May 2013. For example, monthly and annualized users and simulation numbers from 2006 to 2008 are shown in Fig. 13.

The complete NEMO 3D 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 to user to perform long-duration simulations without requiring their constant attention. As shown in this entry, the computational resources 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 (http://www.teragrid.org) and the Open Science Grid (http://www.opensciencegrid.org) to

DOI 10.1007/978-3-642-27737-5_343-2

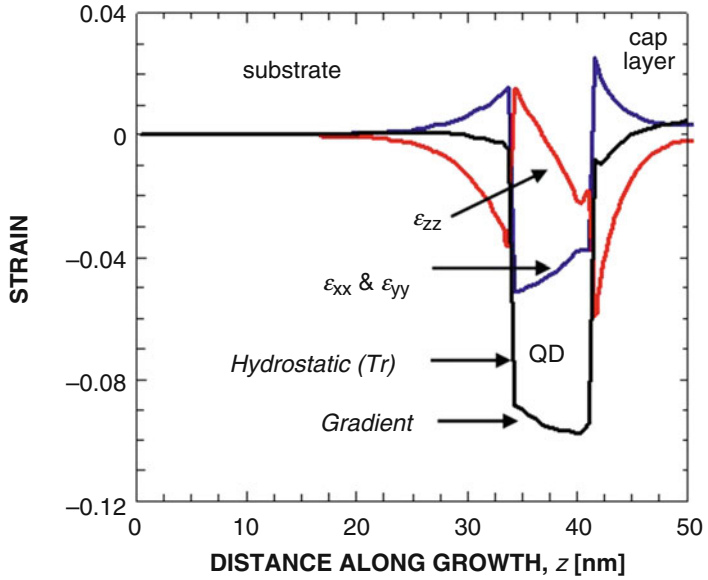


Fig. 14 Atomistic *diagonal* strain profile along the [001], z direction. Dome-shaped dot with diameter $d = 11.3$ nm and height $h = 5.65$ nm. Strain is seen to penetrate deep inside the substrate and the cap layer. Also, noticeable is the gradient in the trace of the hydrostatic strain curve (*Tr*) inside the dot region that results in optical polarization anisotropy and non-degeneracy in the electronic conduction band *P*. Atomistic strain thus lowers the symmetry of the dot

seamlessly provide the necessary backend computational capacity to do computationally intensive 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 3D user interface continues. The more comprehensive interface will provide access to a broader audience and encourage the continued growth of the nanoHUB user base.

Simulation Results

Strain and Piezoelectricity in InAs/GaAs Single QDs

The dome-shaped InAs QDs that are studied first in this work are embedded in a GaAs barrier material (schematic shown in Fig. 2) and have diameter and height of 11.3 nm and 5.65 nm, respectively, and are positioned on a 0.6-nm-thick wetting layer (Bester and Zunger 2005; Lee et al. 2004a). The simulation of strain is carried out in the larger computational box (width D_{strain} and height H), while the electronic structure computation is usually restricted to the smaller domain (width D_{elec} and height H_{elec}). All the strain simulations in this category fix the atom positions on the bottom plane to the GaAs lattice constant, assume periodic boundary conditions in the lateral dimensions, and open boundary conditions on the top surface. The inner electronic box assumes closed boundary conditions with passivated dangling bonds (Lee et al. 2004b). The strain domain contains ∼3 M atoms while the electronic structure domain contains ∼0.3 M atoms.

Impact of Strain. Strain modifies the effective confinement volume in the device, distorts the atom bonds in length and angles, and hence modulates the local bandstructure and the confined states. Figure 14 shows the diagonal (biaxial) components of strain distribution along the [001] direction in the quantum dot (cut through the center of the dot). There are two salient features in this plot: (a) The atomistic strain is long-ranged and penetrates deep into both the substrate and the cap layers and (b) all the components of

DOI 10.1007/978-3-642-27737-5_343-2

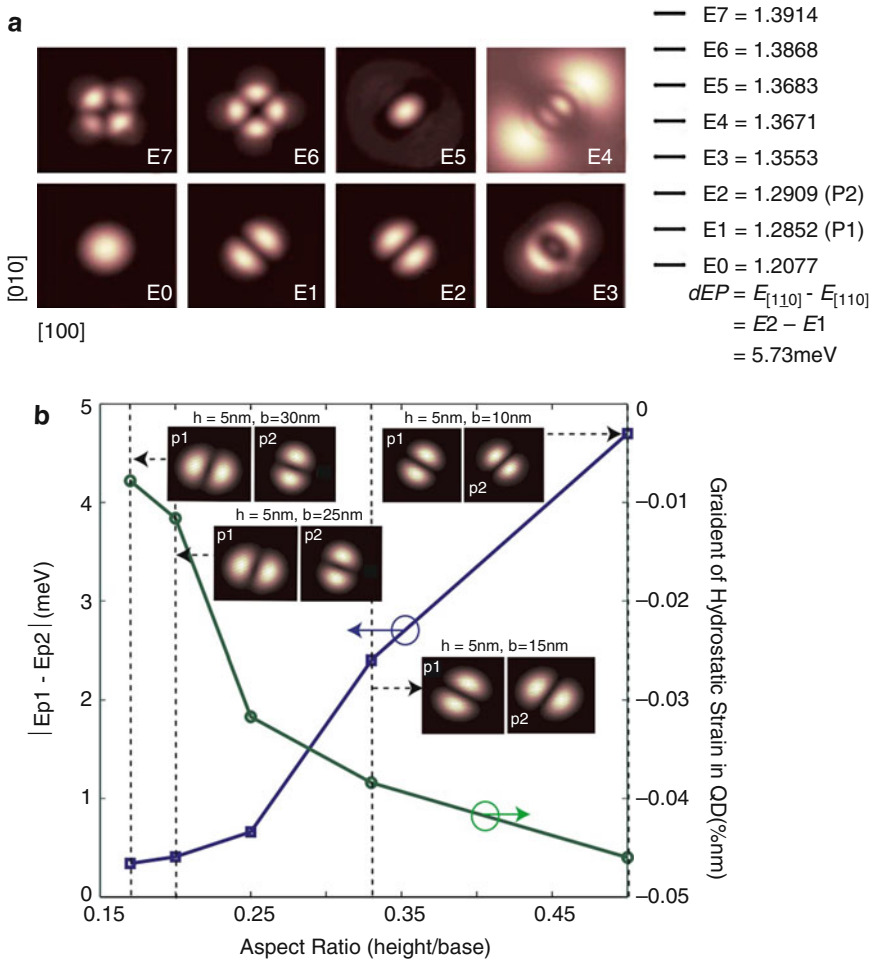


Fig. 15 (**a**) Conduction band wave functions and spectra (eV) for first eight energy levels in the dome-shaped quantum dot structure. Atomistic strain is included in the calculation. Note the optical anisotropy and non-degeneracy in the *P* energy level. The first state is oriented along [110] direction and the second state along [1$\underline{1}$0] direction. (**b**) Gradient in the hydrostatic strain along the [001] direction through the center of the dot and the resulting non-degeneracy and optical anisotropy in the *P* level as a function of the dot aspect ratio

biaxial stress have a nonzero slope inside the quantum dot region. The presence of the gradient in the trace of the hydrostatic strain introduces unequal stress in the zinc blende lattice structure along the depth, breaks the equivalence of the [110] and [1$\underline{1}$0] directions, and finally breaks the degeneracy of the first excited electronic state (the so-called *P* level). Figure 15a shows the wave function distribution for the first 8 (eight) conduction band electronic states within the device region for the dot (in a 2D projection). Note the optical anisotropy and non-degeneracy in the first excited (*P*) energy level. The first *P* state is oriented along the [110] direction and the second *P* state along the [1$\underline{1}$0] direction. The individual energy spectrum is also depicted in this figure which reveals the value of the *P* level splitting/non-degeneracy (defined as $E_{1\underline{1}0} - E_{110}$) to be about 5.73 meV.

As explained in Bester and Zunger (2005), the shape symmetry of a quantum dot is lowered due mainly to three reasons, all originating from the fundamental atomistic nature of the underlying crystal: (1) The *interface* between the dot material (InAs) and the barrier material (GaAs), even with a common anion (As atom), is not a reflection plane and hence anisotropic with respect to the anion. The direct neighbors above the anion plane (In atoms) that align in the [1$\underline{1}$0] direction are chemically different from the

DOI 10.1007/978-3-642-27737-5_343-2

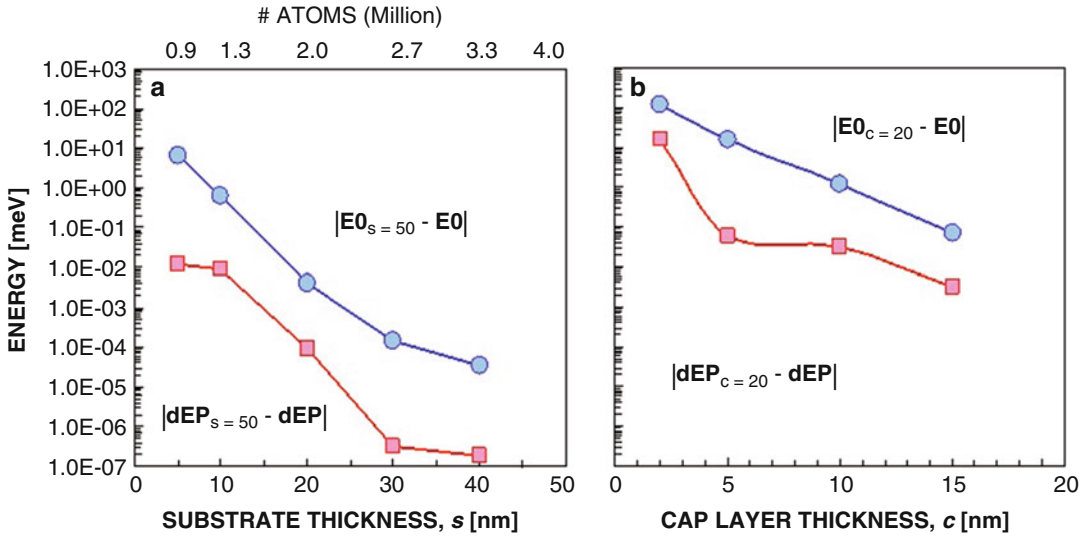


Fig. 16 (**a**) Substrate layer thickness dependence of the conduction band minimum and the *P* level splitting. Other structural parameters remain constant ($h = 5.65$ nm, $d = 11.3$ nm, $c = 10$ nm, and $D = 31.3$ nm). (**b**) The impact of cap layer thickness (with substrate $s = 30$ nm and other structural parameters remaining the same). Lanczos convergence tolerance $= 1 \times 10^{-7}$

neighbors under the anion plane (Ga atoms) that align in the [110] direction. This creates a short-range interfacial potential. It is important to note that these atomistic interfacial potentials originating from different facets do not necessarily compensate each other in dots where the base is larger than the top (e.g., pyramid, lens, truncated pyramid). (2) *Atomistic strain and relaxations* (originating from the atomic size difference between Ga and In atoms) result in a propagation of the interfacial potential further into the dot material and thus amplify the magnitude of the asymmetry. This component is not captured if the relaxation is performed using classic harmonic continuum–elasticity approach. Noticeable is the fact that symmetry breaking due to atomistic relaxations can even be observed in dots where the base is equal to the top (e.g., box, disk); however, the effect is magnified in dots of typical shape, where the base is larger than the top (e.g., pyramid, lens, truncated pyramid) due to the presence of a gradient in the magnitude of the strain tensor between top and bottom as already explained in Fig. 15a. In order to further characterize this effect, we have simulated dome-shaped dots with varying base diameters (from 10 to 30 nm) keeping the dot height constant (at 5 nm). Figure 15b shows the gradient in the hydrostatic strain and the resulting non-degeneracy in the *P* level as a function of the dot aspect ratio (height/base). Also, shown in the insets are the wave functions corresponding to the split *P* levels in each of these dots. Note that the non-degeneracy and the optical anisotropy diminish as the dot aspect ratio decreases (approaching a disk shape). (3) Finally, a long-ranged *piezoelectric field* develops in these dots in response to the strain-induced displacement field, which is fundamentally anisotropic. We will discuss this effect in detail in a subsequent section.

Need for a Deep Substrate and a Realistic Cap Layer. The strength of the NEMO 3D package lies particularly in its capability of simulating device structures with realistic boundary conditions. Our simulation results based on NEMO 3D show a significant dependence of the dot states and magnitude of level splitting on the substrate layer thickness, *s* (underneath the dot), and the cap layer thickness, *c* (above the dot). The strain in the QD system therefore penetrates deeply into the substrate and cannot be neglected. Figure 16 shows such observed dependency where *E0* is the ground state energy and *dEP* is the magnitude of the level splitting in the *P* electronic states due to the inclusion of atomistic strain and relaxation. The changes in both these quantities are calculated with respect to the largest *s* (50 nm) and *c* (20 nm), respectively, in Fig. 16a, b. The wave function orientation was found to remain unchanged

DOI 10.1007/978-3-642-27737-5_343-2

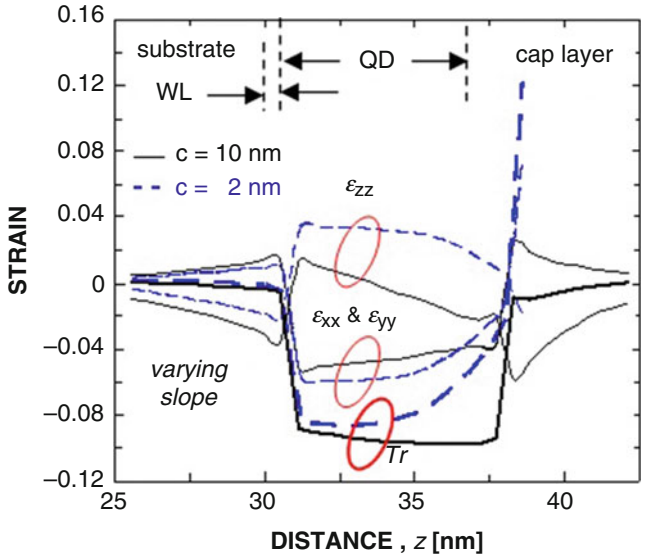


Fig. 17 The impact of cap layer thickness (with deep substrate, $s = 30$ nm, and $h = 5.65$ nm, $d = 11.3$ nm). Shown is the significant variation of gradient/slope in the strain profile within the quantum dot region. This results in a different splitting in the conduction band P energy level for the two different thicknesses of the cap layer

irrespective of the substrate depth and cap layer thickness. Figure 16a shows that it is indeed important to include enough of a substrate to capture the long-range strain, while Fig. 16b indicates opportunities to tune the eigen energy spectrum with different capping layer thicknesses.

Figure 17 reveals the reason of a strong dependency of the electronic ground state and the magnitude of non-degeneracy in P level on the cap layer thickness. Here, the hydrostatic strain profiles for two different cap layer thicknesses (2 and 10 nm) are plotted. The P level splitting in a device with 10-nm cap layer is found to be 5.73 meV and that for a 2-nm cap layer was 20.58 meV. The reason of the reduction in the splitting in the 10-nm cap layer device can be attributed mainly to the change in the gradient of hydrostatic strain inside the device region as depicted in Fig. 17.

Impact of Piezoelectric Fields. The presence of nonzero off-diagonal strain tensor elements leads to the generation of a piezoelectric field in the quantum dot structure, which is incorporated in the simulations as an external potential by solving the Poisson equation on the zinc blende lattice. Figure 18a, b shows the atomistic off-diagonal strain profiles in dome-shaped quantum dots with heights h of 2.8 and 5.65 nm, respectively. The off-diagonal strain tensors are higher in the larger diameter dot. The off-diagonal strain tensors are found to be larger in the dome-shaped dot. The off-diagonal strain tensors are used to calculate the first-order polarization in the underlying crystal (see Bester and Zunger 2005 for the governing equations) which gives rise to a piezoelectric charge distribution throughout the device region and then used to calculate the potential by solving the Poisson equation. The relevant parameters for the piezoelectric calculation are taken from Bester and Zunger (2005). Experimentally measured polarization constants of GaAs and InAs material (on unstrained bulk) values of -0.16 C/m^2 and -0.045 C/m^2 are used. The second-order piezoelectric effect (Bester et al. 2006a) is neglected here because of unavailability of reliable relevant polarization constants for InAs/GaAs quantum dot structures.

The calculated piezoelectric charge and potential surface plots in the XY and XZ planes are shown in Figs. 19 and 20, respectively, revealing a pronounced polarization effect induced in the structure. It is found that piezoelectric field alone favors the [1$\underline{1}$0] orientation of the P level. Also shown in Fig. 21 is the asymmetry in potential profile due to atomistic strain and inequivalence in the piezoelectric potential along [110] and [1$\underline{1}$0] directions at a certain height $z = 1$ nm from the base of the dot.

DOI 10.1007/978-3-642-27737-5_343-2

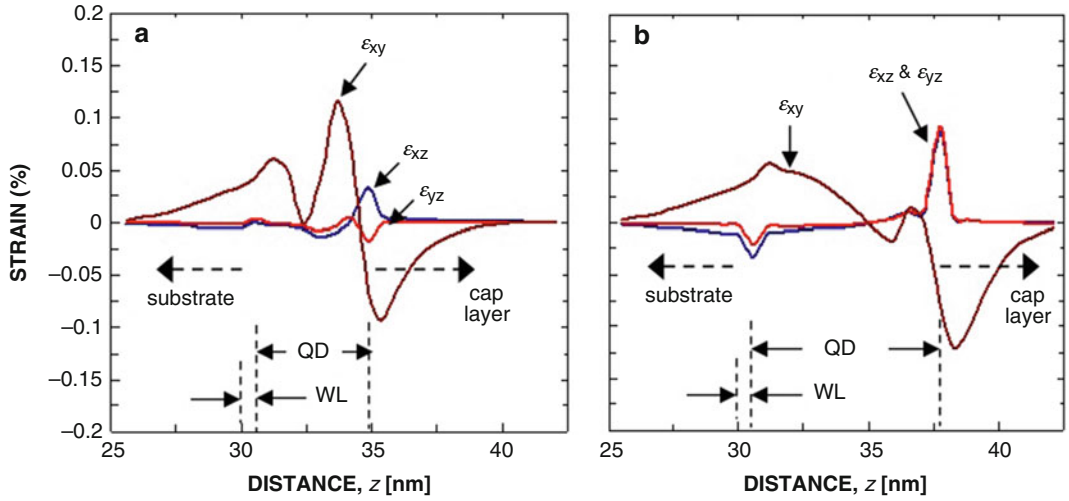


Fig. 18 Atomistic *off-diagonal* strain profile along the z (vertical) direction which in effect induces polarization in the quantum dot structure. (**a**) Diameter $d = 11.3$ nm and height $h = 2.8$ nm and (**b**) diameter $d = 11.3$ nm and height $h = 5.65$ nm. Note the increase in *off-diagonal* stain in (**b**)

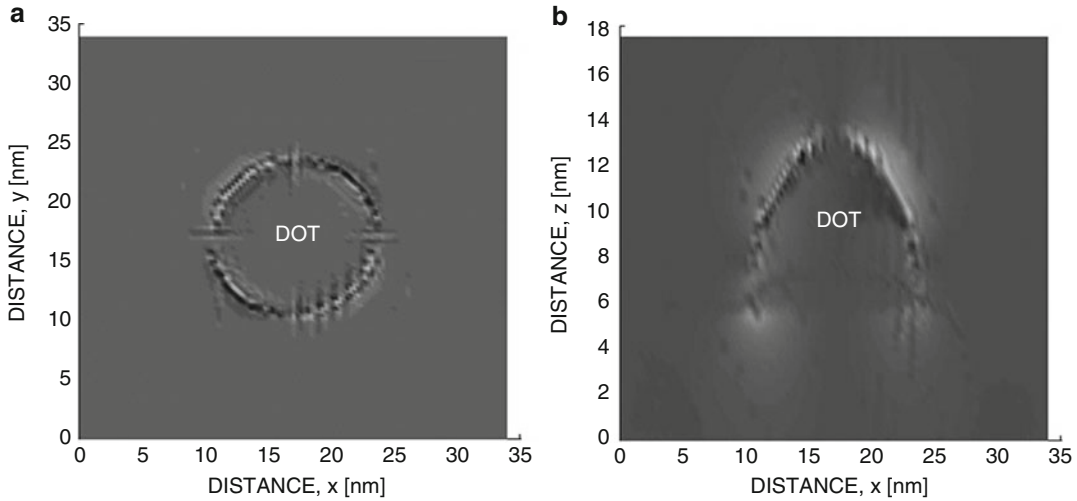


Fig. 19 Charge surface plot of a dome-shaped quantum dot (**a**) in the XY plane at $z = 1$ nm from the base of the dot and (**b**) in the XZ plane at $y = D_{\text{strain}}/2$. Charge is induced mainly in the vicinity of the boundary of the quantum dot ($d = 11.3$ nm and $h = 5.65$ nm)

Study of Varying-Sized Dots. The impact of atomistic strain and piezoelectric field on the ground state energy and magnitude of the P level energy splitting in dome-shaped quantum dots with varying diameter d and dot height h is shown in Figs. 22 and 23, respectively. The ground state energy for the strained system (without piezoelectricity), $E0$, decreases with an increase in both d and h because of an increase in the effective confinement volume. Figures 22a and 23a also show the change (absolute and relative to strain only) in the ground state energy due to the inclusion of piezoelectric potential in the strained system. The percentage change in the ground state energy is found to be monotonous in nature with an increase in dot diameter while the height dependency shows saturation beyond a certain value. Figures 22b and 23b show the change of three quantities related to the first excited P level, namely, split due to strain only (circle), split due to strain combined with piezoelectricity (square), and the contribution of the piezoelectric field only (triangle), as a function of diameter d and dot height h. The piezoelectric potential introduces

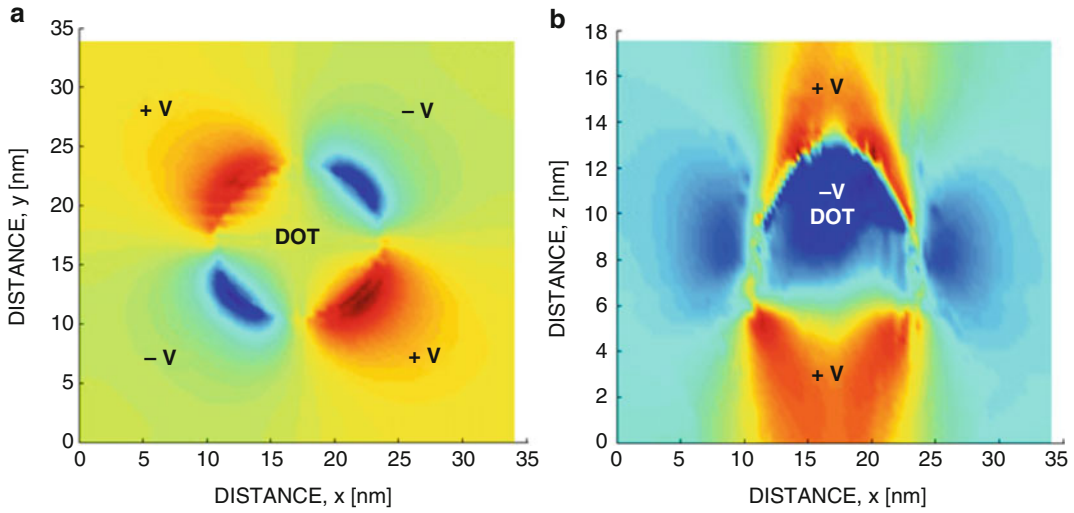


Fig. 20 Piezoelectric potential surface plot of a dome-shaped quantum dot (**a**) in the XY plane at $z = 1$ nm from the base of the dot and (**b**) in the XZ plane at $y = D_{strain}/2$. (**c**) Potential along [110] and [1̲10] directions at z = 1 nm from the base of the dot. Note the induced polarization in the potential profile and the unequal values of potential along the [110] and [1̲10] directions ($d = 11.3$ nm and $h = 5.65$ nm)

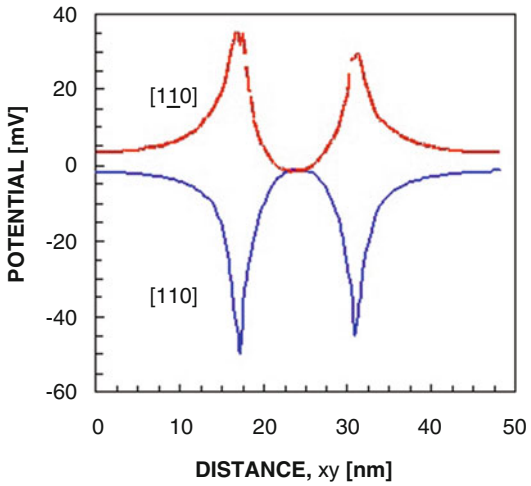


Fig. 21 Potential along [110] and [1̲10] directions at z = 1 nm from the base of the dot. Note the induced polarization in the potential profile and the unequal values of potential along the [110] and [1̲10] directions ($d = 11.3$ nm and $h = 5.65$ nm)

a global shift in the energy spectrum and is found to be strong enough to flip the optical polarization in certain-sized quantum dots. In those cases, the piezoelectric contribution (triangle) dominates over that resulting from the inclusion of atomistic strain alone in the simulations (circle) as can be seen in dots (see Fig. 22b; similar trend has also been found in Bester et al. 2006a) with diameters larger than 7 nm and (see Fig. 23b) height more than 3 nm. Figure 24 explains the reason behind this observation. Here, the piezoelectric potential profiles in dots with different height h are plotted along the z direction through the dot center. Note the increase in piezoelectric potential with dot height. The stronger piezoelectric potential induced in the larger dot results in the orientational flip in the P level electronic states.

Piezoelectricity-Induced Polarization Flip. Figure 25 shows the conduction band wave functions for the ground and first three excited energy states in the quantum dot structure with diameter of 11.3 nm and

DOI 10.1007/978-3-642-27737-5_343-2

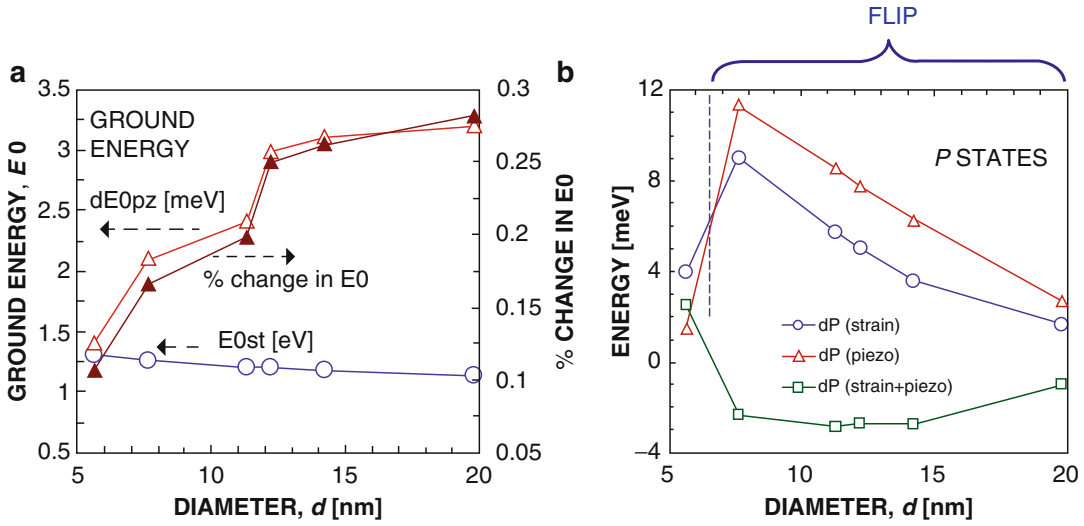


Fig. 22 Study of electronic structure with the variation of *dot diameter*, d, of the dome-shaped quantum dot. (**a**) Conduction band minimum/ground state in a strained system (*circle*) and change in the conduction band minimum due to induced piezoelectricity (*triangle*). (**b**) Split in the P level due to strain only (*circle*), split in the P level due to strain and piezoelectricity (*square*), and impact of piezoelectric potential *alone* (*triangle*) in the system (dot height, $h = 5.65$ nm)

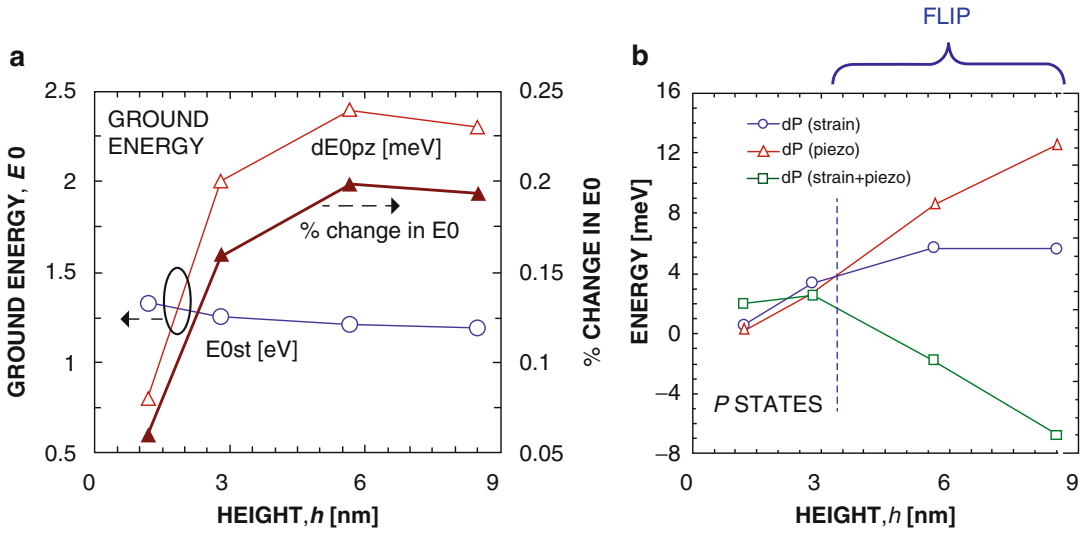


Fig. 23 Study of electronic structure with the variation of *dot height*, h, of the dome-shaped quantum dot. (**a**) Conduction band minimum/ground state in a strained system (*circle*) and change in the conduction band minimum due to induced piezoelectricity (*triangle*). (**b**) Split in the P level due to strain only (*circle*), split in the P level due to strain and piezoelectricity (*square*), and impact of piezoelectric potential *alone* (*triangle*) in the system (dot diameter, $d = 11.3$ nm)

height h of 5.65 nm. In Fig. 25a, strain and piezoelectricity are *not* included in the calculation. The weak anisotropy in the P level is due to the atomistic interface and material discontinuity. Material discontinuity mildly favors the [110] direction in the dot. In Fig. 25b, atomistic strain and relaxation is included resulting in a 5.73 meV split in the P energy levels. Strain favors the [110] direction. In Fig. 25c, piezoelectricity is included on top of strain inducing a split of -2.84 meV in the P energy level. The first P state is oriented along [1$\underline{1}$0] direction and the second state along [110] direction, and piezoelectricity alone induces a potential of 8.57 meV. Piezoelectricity thereby has not only introduced a global

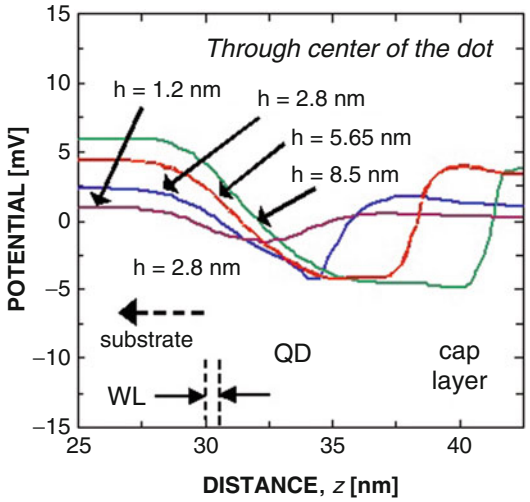


Fig. 24 Piezoelectric potential in dome-shaped quantum dots with $h = 2.8$ nm and $h = 5.65$ nm along the z direction through the center of the dots. Noticeable is the stronger polarization in the larger dot which results in a *flip* in the P level electronic states

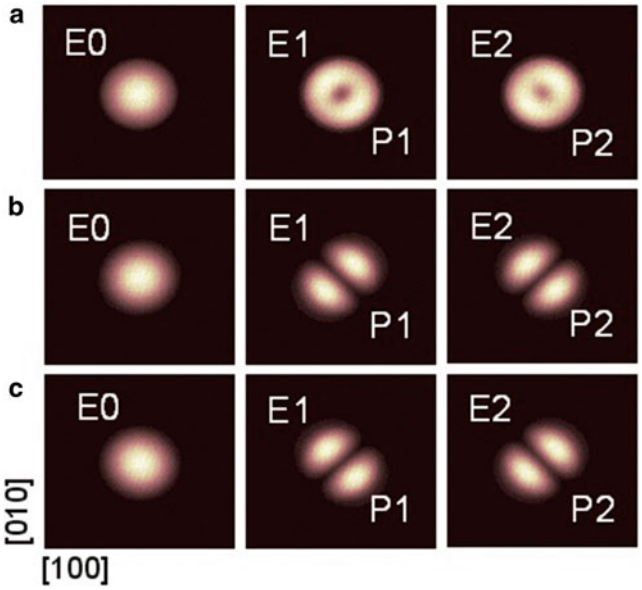


Fig. 25 Conduction band wave functions for first three energy levels in the quantum dot structure with diameter $d = 11.3$ nm and height $h = 5.65$ nm (**a**) without strain and piezoelectricity, $E_{[1\underline{1}0]} - E_{[110]} = 1.69$ meV; (**b**) with atomistic strain, $E_{[1\underline{1}0]} - E_{[110]} = 5.73$ meV; and (**c**) with strain and piezoelectricity, $E_{[1\underline{1}0]} - E_{[110]} = -2.84$ meV. Piezoelectricity *flips* the wave functions. An end-to-end computation involved about 4 M atoms and needed CPU time of about 8 h with 16 processors

shift in the energy spectrum but also *flipped* the orientation of the P states (Bester and Zunger 2005) in this case.

Stacked Quantum Dot System

Self-assembled quantum dots can be grown as stacks where the QD distance can be controlled with atomic-layer control. This distance determines the interaction of the artificial atomic states to form artificial molecules. The design of QD stacks becomes complicated since the structures are subject to inhomogeneous, long-range strain and growth imperfections such as nonidentical dots and interdiffused

DOI 10.1007/978-3-642-27737-5_343-2

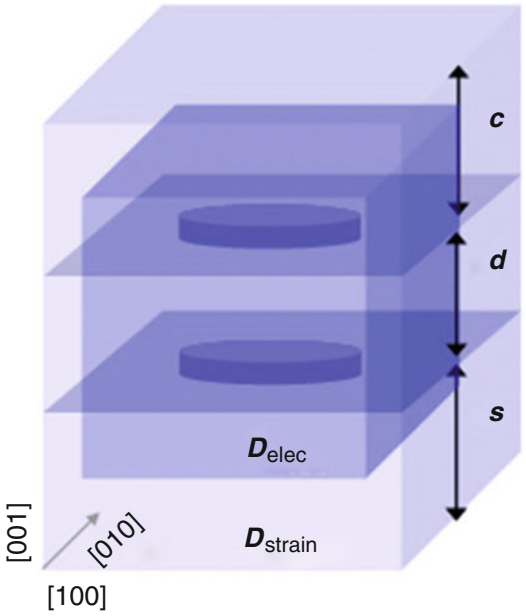


Fig. 26 Simulated InAs/GaAs double quantum dots with disk/cylindrical shape. The dots are of equal size with radius *r* of 7 nm and height *h* of 1.5 nm. The separation *d* is varied from 0.5 to 8 nm. Two simulation domains have been shown. The strain domain for 8-nm spacing between the dots contained about six million atoms

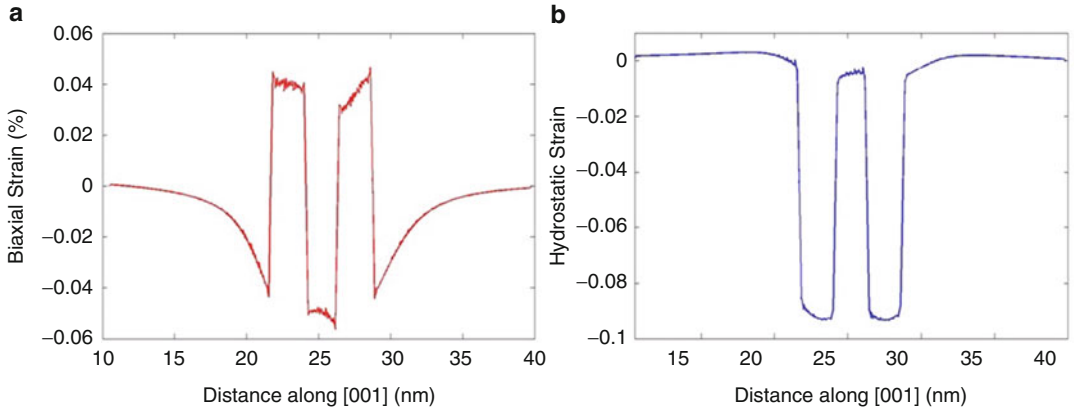


Fig. 27 Atomistic (**a**) *biaxial* $\{\varepsilon_{zz} - (\varepsilon_{xx} + \varepsilon_{yy})/2\}$ and (**b**) *hydrostatic* $\{\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}\}$ strain profile along the growth [001], *z* direction. Strain is seen to penetrate deep inside the substrate and the cap layer. Also, noticeable is the gradient in the trace of the hydrostatic strain curve (*Tr*) inside the dot region that results in optical polarization anisotropy and non-degeneracy in the electronic conduction band *P*. Atomistic strain thus lowers the symmetry of the dot

interfaces. Quantum dot stacks consisting of two QD layers are simulated next (see Fig. 26). The InAs quantum dots are disk shaped with a diameter of 7 nm and height of 1.5 nm positioned on a 0.6-nm-thick wetting layer. The substrate thickness under the first wetting layer is kept constant at 30 nm, and the cap layer on top of the topmost dot is kept at 20 nm for all the simulations. The strain simulation domain (D_{strain}) contains 8–10 M atoms, and the electronic structure domain (D_{elec}) contains 0.5–1.1 M atoms.

Figure 27 shows both the biaxial and hydrostatic strain profiles along the *z* direction. As in the single dot, we see a gradient in strain profile within the dot regions which results in strain-induced asymmetry. The hydrostatic component which is responsible for conduction band well is negative within the dot and approximately zero outside the dot and the regions in-between the dots. The biaxial component of strain

DOI 10.1007/978-3-642-27737-5_343-2

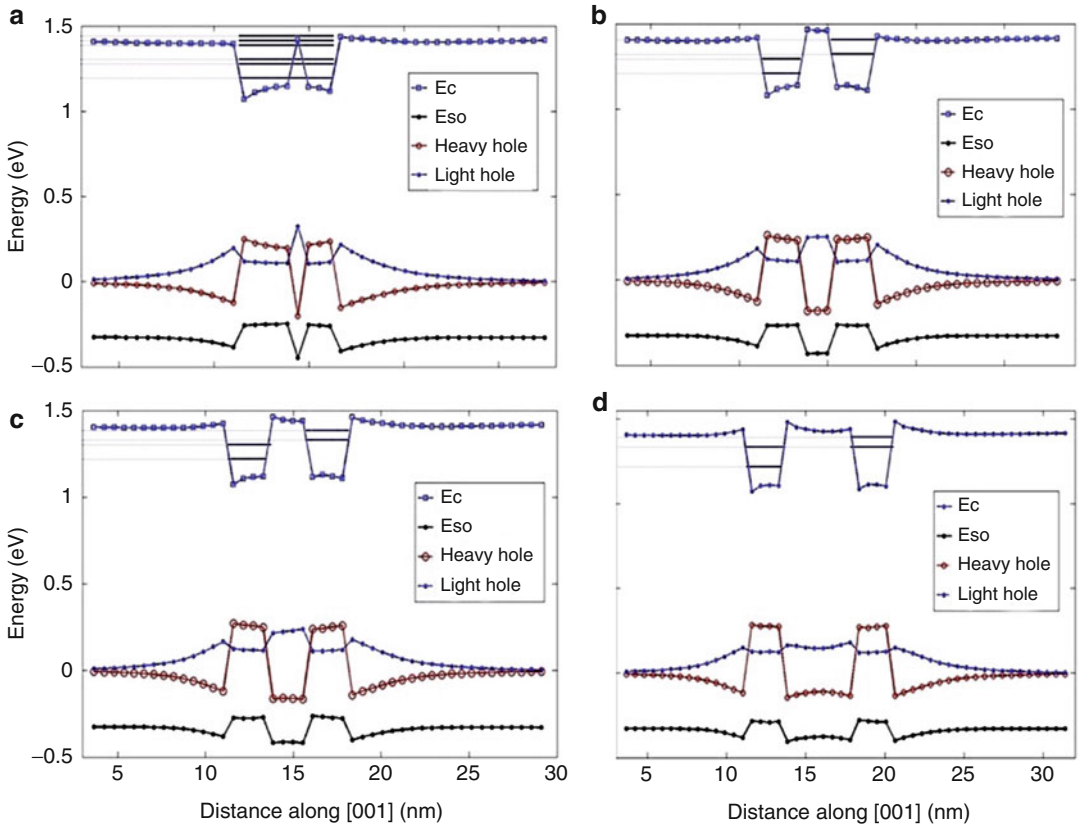


Fig. 28 Band-edge diagrams for double quantum dot systems for several inter-dot spacing: (**a**) 0.5 nm, (**b**) 1 nm, (**c**) 2 nm, and (**d**) 4 nm. Strain makes InAs conduction band potential wells shallower, enhancing the coupling between the dots. Noticeable is the effect on hole wells. Strain splits the light hole and heavy hole bands. Within the dot, heavy hole lies above the light hole edge. As strain coupling decreases, heavy hole well becomes more and more shallow (see **b** and **d**)

which has more effect on hole states is positive within dots and negative in-between the dots. The magnitude for both is approximately equal. Figure 28 shows the band-edge diagrams as a function of dot separations along the center of the dots in the growth direction [001]. Strain enhances the coupling between the dots. Hydrostatic component of strain makes the conduction band well shallower. Strain effects are more prominent on hole states where biaxial component of strain splits the light hole and heavy hole bands. Within the dot, heavy hole lies above the light hole edge, implying significant band mixing in the confining states. For very small separation like 0.5 nm, the well within dots is even shallower than the well in-between the dots. Figures 29 and 30 show the electron and hole-state energies, respectively, as a function of inter-dot separation. In a system without inhomogeneous strain, one would expect the identical dots to have degenerate eigenstate energies for large dot separations. Strain breaks the degeneracy even for large separations. As the dot separation is narrowed, the dots interact with each other mechanically through the strain field as well as quantum mechanically through wave function overlaps. Wave function plots in *XZ* plane have been shown in Fig. 29 for various dot separations. Noticeably, E2 for 4-nm separation is a *P*-like state, while it is *S*-like state in 6-nm separation. So there is a crossover between *p* to *s* for E2 as we increase separation between the dots. Also, E1 for 2-nm separation is confined more in the lower dot than the upper dot. This is caused by strain coupling which promotes confinement of the ground states in the lower dots in coupled quantum dot systems (Korkusinski and Klimeck 2006). The electronic states and wave functions in a coupled QD system are thus determined through a complicated interplay of strain, QD size, and wave function overlap. Only a detailed simulation can reveal that interplay.

DOI 10.1007/978-3-642-27737-5_343-2

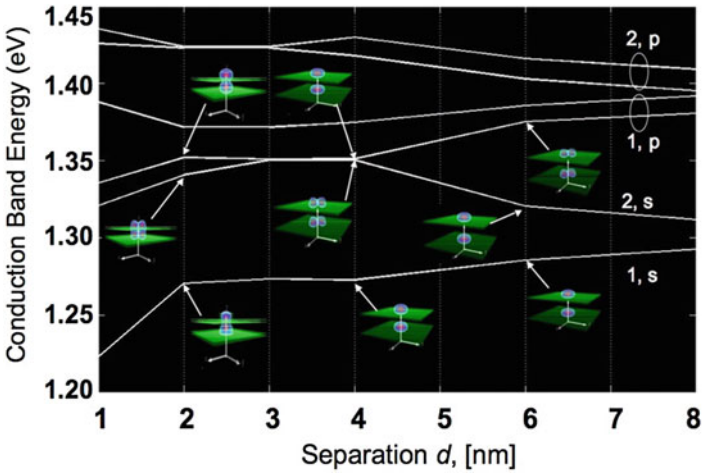


Fig. 29 Dependence of six lowest electron energy levels on separation distance *d* between the dots. For electron energy levels, the state names are mentioned as *s* or *p* orbital states. Here, 1 indicates bonding states, whereas 2 indicates antibonding states. Wave function plots in XZ plane have been shown for some dot separations. Noticeably, e2 for 4-nm separation is a P-like state, while it is an S-like state in 6-nm separation. So there is a crossover between p to s for e2 as we increase separation between the dots. Also, e1 for 4-nm separation is confined in lower dot more than upper dot. This is caused by strain coupling which tries to confine ground states in the lower dots in coupled quantum dot systems

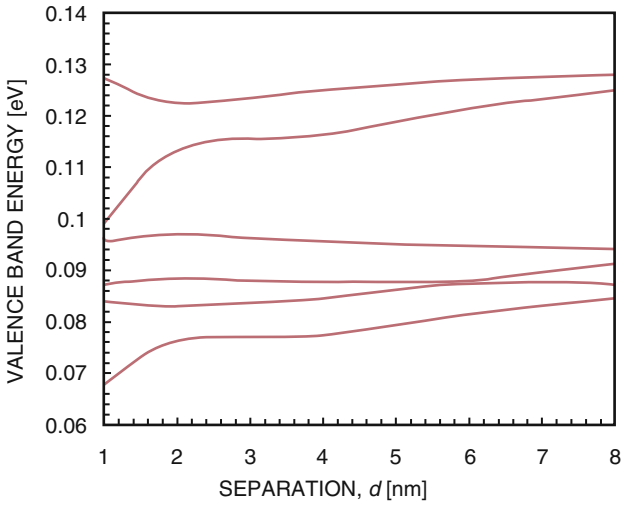


Fig. 30 Dependence of six lowest hole energy levels on separation distance *d* between the dots

Effects of Internal Fields in InN/GaN Quantum Dots

Objective. In the last decade, GaN and its related alloys especially InGaN have been the subject of intense experimental and theoretical research due mainly to their wide range of emission frequencies, high stability against defects, and potential for applications in various optoelectronic, solid-state lighting, and high-mobility electronic devices (Ponce and Bour 1997; Morkoç and Mohammad 1995). Since the heteroepitaxy of InN on GaN involves a lattice mismatch up to ∼11 %, a form of Stranski–Krastanov mode can be used for growing InN on GaN by molecular beam epitaxy (MBE). This finding gives rise to the possibility of growing InN quantum dots (QDs) on GaN substrates. Recent studies have shown that the strain between InN and GaN can be relieved by misfit dislocations at the hetero-interface after the deposition of the first few InN bilayers and before the formation of InN islands (Kalden et al. 2010;

DOI 10.1007/978-3-642-27737-5_343-2

Ke et al. 2006). Relaxed InN islands with controllable size and density can be formed by changing the growth parameters (such as temperature) in either MBE or metalorganic chemical vapor deposition (MOCVD) (Wang et al. 2009a). Relaxation of elastic strain at free surfaces in semiconductor dots (and nanowires) allows the accommodation of a broader range of lattice mismatch and band lineups in coherent nanostructures than is possible in conventional bulk and thin-film (quantum well) heterostructures, and, therefore, threading dislocations can be all but nonexistent in quantum dots (and nanowires). Furthermore, QDs used in the active region of optical devices provide better electron confinement (due to strongly peaked energy dependence of density of states) and thus a higher-temperature stability of the threshold current and the luminescence than quantum wells. It is also clear that high-quality bulk GaN is an ideal substrate material for nitride nanostructures. Pure GaN crystal is five times more thermally conductive than sapphire and optically transparent at visible and near-UV wavelengths.

Knowledge of the electronic bandstructure of nanostructures is the first and an essential step toward the understanding of the optical performance (luminescence) and reliable device design. Hexagonal group III nitride 2D quantum well (QW) heterostructures have experimentally been shown to demonstrate *polarized transitions* in quantized electron and hole states and non-degeneracy in the first excited state in various spectroscopic analyses (Waltereit et al. 2000). These observations suggest the existence of certain *symmetry-lowering mechanisms* (structural and electrostatic fields) in these low-dimensional nanostructures. While 0D QDs promise better performance, only very few and recent experimental results exist concerning the photoluminescence (PL) and electroluminescence (EL) of nitride QDs in the visible spectral region (Jarjour et al. 2008; Kalden et al. 2010), and experiments revealing *polarization anisotropy* in InN QDs are rare. Similar to the 2D QW structures, the optical properties of the QDs are expected, to a large extent, to be determined by an intricate interplay between the structural and the electronic properties and (since not yet been fully assessed experimentally) demand detailed theoretical investigations.

In this section, using mainly the NEMO 3D tool, we study the electronic bandstructure of wurtzite InN/GaN quantum dots having *three* different geometries, namely, box, dome, and pyramid (Ahmed et al. 2010; Yalavarthi et al. 2011). The main objectives are twofold: (1) to explore the nature and quantify the role of crystal atomicity, strain fields, and piezoelectric and pyroelectric potentials in determining the energy spectrum and the wave functions and (2) to address a group of related phenomena including shift in the energy state, symmetry lowering and non-degeneracy in the first excited state, strong band mixing in the overall conduction band electronic states, and strongly suppressed and optically anisotropic interband transitions.

Simulation Results. Figure 31 shows the simulated quantum dots with box, dome, and pyramid geometries. The InN QDs grown in the [0001] direction and embedded in a GaN substrate used in this study have (unless otherwise stated) diameter/base length $d \sim 10.1$ nm and height $h \sim 5.6$ nm and are positioned on an InN wetting layer of one atomic-layer thickness. The simulation of strain is carried out in the large computational box, while the electronic structure computation is restricted to the smaller inner domain. All the strain simulations fix the atom positions on the bottom plane to the GaN lattice constant, assume periodic boundary conditions in the lateral dimensions, and open boundary conditions on the top surface. The strain parameters used in this work were validated through the calculation of Poisson ratio of the bulk materials. The inner electronic box assumes a closed boundary condition with passivated dangling bonds.

Figure 32 shows the topmost valence (HOMO) and first four conduction band wave functions (projected on the X–Y plane) for the quantum dots *without* strain relaxation. Here, both the InN dot and the GaN barrier assume the lattice positions of perfect wurtzite GaN. The topmost valence (HOMO) state in all three quantum dots has orbital *S* character retaining the geometric symmetry of the dots. The first electronic state is *S*-like, while the next three states are *P*-like and the split (non-degeneracy) in these

DOI 10.1007/978-3-642-27737-5_343-2

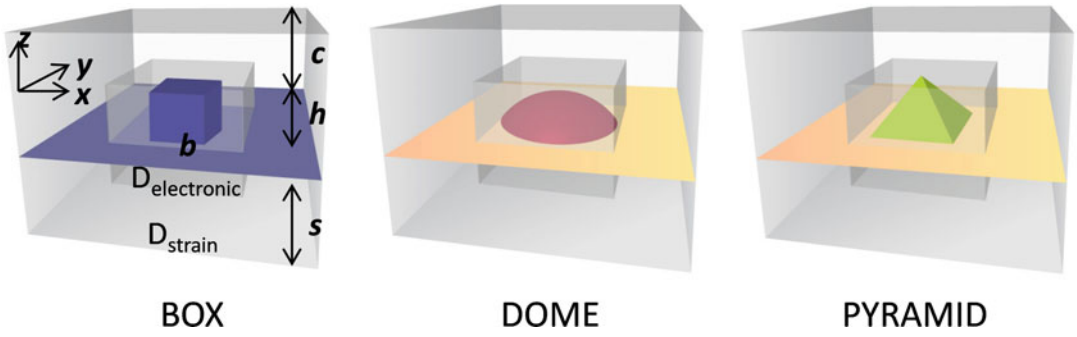


Fig. 31 Simulated InN/GaN quantum dots on a thin InN wetting layer. Two major computational domains are also shown: D_{elec}, central smaller domain for electronic structure (quantum) calculation, and D_{strain}, outer domain for strain calculation. In the figure, s is the substrate height, c is the cap layer thickness, h is the dot height, and d is the dot diameter/base length as appropriate

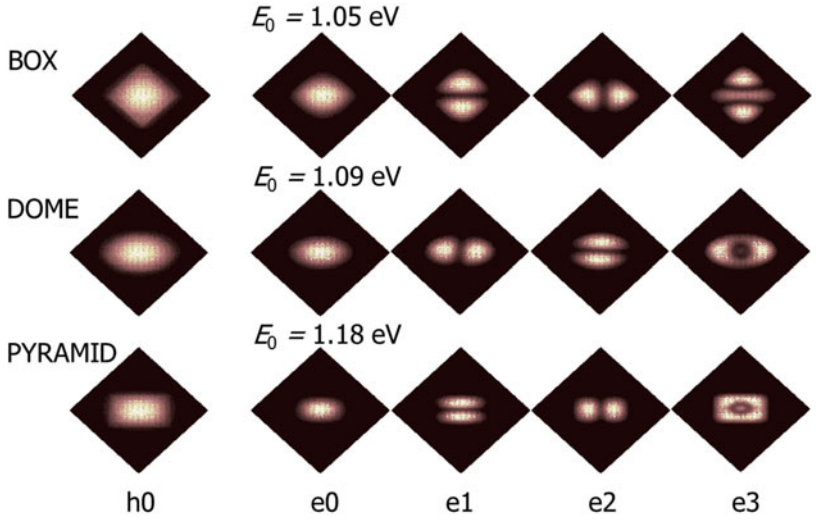


Fig. 32 Topmost valence (HOMO) and first four conduction band wave functions due to fundamental crystal and interfacial symmetry. Noticeable are the split and the anisotropy in the P level. Number of atoms simulated: 1.78 million (strain domain), 0.8 million (electronic domain)

levels originate from the crystal fields alone. The anisotropy in the P states assumes different orientations – for box and pyramid dots, the first P state is oriented along the [010] direction and the second along [100] direction, while the converse occurs in a dome. It is clear that the fundamental crystal atomicity and the interfaces (between the dot material InN and the barrier material GaN) lower the geometric shape symmetry even in the absence of strain relaxation. Therefore, the interface plane creates a short-range interfacial potential and cannot be treated as a reflection plane.

Next, we introduce atomistic strain relaxation in our calculations using the VFF method 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 equilibrium atomic positions are found by minimizing the total elastic energy of the system. However, piezoelectricity is neglected in this step. 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.

DOI 10.1007/978-3-642-27737-5_343-2

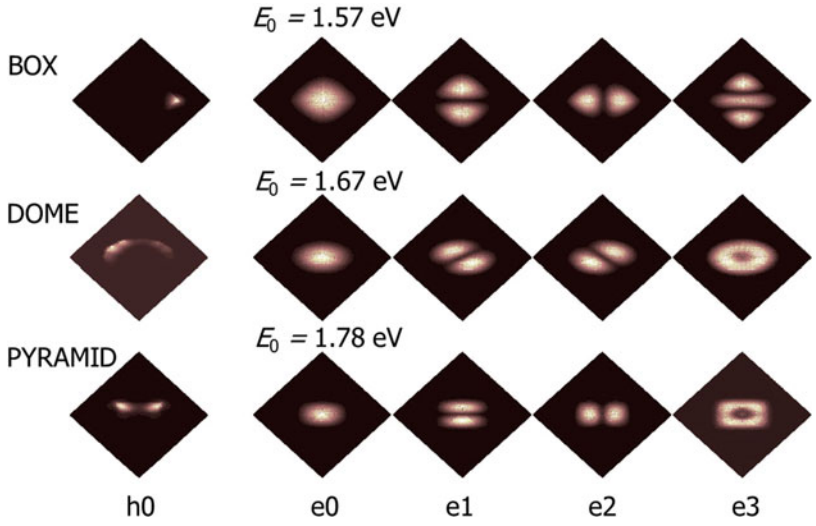


Fig. 33 Topmost valence (HOMO) and first four conduction band wave functions due to combined effects of atomicity and strain relaxation. Noticeable are the strongly displaced HOMO (hole) wave functions, deformed LUMO states, and split and the anisotropy in the P level

Strain modifies the effective confinement volume in the device, distorts the atomic bonds in length and angles, and hence modulates the confined states. From our calculations, atomistic strain was found to be anisotropic and long-ranged penetrating deep (∼20 nm) into both the substrate and the cap layers stressing the need for using *realistically extended* substrate and cap layers (multimillion-atom modeling) in the study of electronic structure of these reduced dimensional QDs. Figure 33 shows the wave function distributions for the topmost valence band and first 4 (four) conduction band electronic states in a 2D projection. Noticeable are the deformed LUMO (electronic) states and the pronounced optical anisotropy and non-degeneracy in the P levels. Strain introduces *uniform* orientational pressure (adds negative potential) in all three quantum dots with ΔP to be largest in a pyramid and minimum in a dome. Also, strain relaxation causes *blue* shift in the conduction band electronic states and results in strongly displaced HOMO (hole) wave functions. These observations will have significant implications on the optical polarization and performance of devices based on these nanostructures.

In pseudomorphically grown heterostructures, the presence of nonzero atomistic stress tensors results in a deformation in the crystal lattice and leads to a combination of piezoelectric and pyroelectric field, which has been incorporated in the Hamiltonian as an external potential (within a non-self-consistent approximation). The resulting piezoelectric and pyroelectric potential distributions along the growth direction are shown in Fig. 34. The potential (both the piezoelectric and pyroelectric), in accordance with the dot volume, has the largest magnitude in a box and is minimum in a pyramid. *The piezoelectric potential is significantly larger (∼10 times) than the pyroelectric counterpart and tends to oppose the latter*. This also suggests that for an appropriate choice of alloy composition and quantum dot size/geometry, spontaneous and piezoelectric fields may be caused to cancel out!

Figure 35 shows the topmost valence band and first 4 (four) conduction band wave functions for all three quantum dots including the strain relaxation and the piezoelectric potential. The piezoelectric potential introduces a global *red* shift in the energy spectrum and *opposes* the strain-induced field (without any significant modifications in the wave function orientations) in the box and pyramid dots. Figure 36 shows the same wave functions for all three quantum dots including the combined effects of *all four types* of internal fields, namely, interface, strain, piezoelectricity, and pyroelectricity in the calculations. The effect of pyroelectric potential was found to be *negligible* in all the three quantum dots.

DOI 10.1007/978-3-642-27737-5_343-2

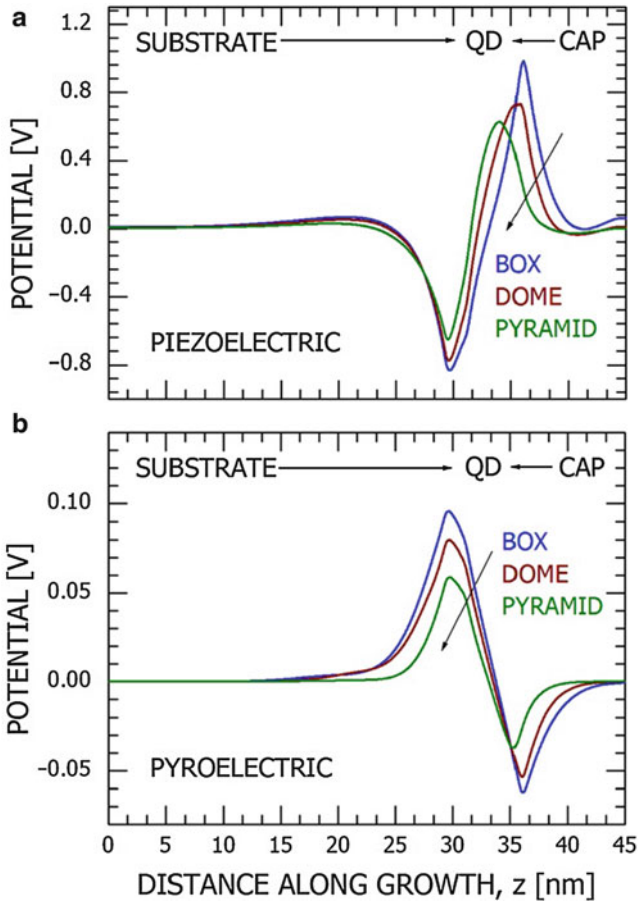


Fig. 34 Induced piezoelectric and pyroelectric potential distributions along the *z* (growth) direction. Note the spread/penetration in the surrounding material matrix

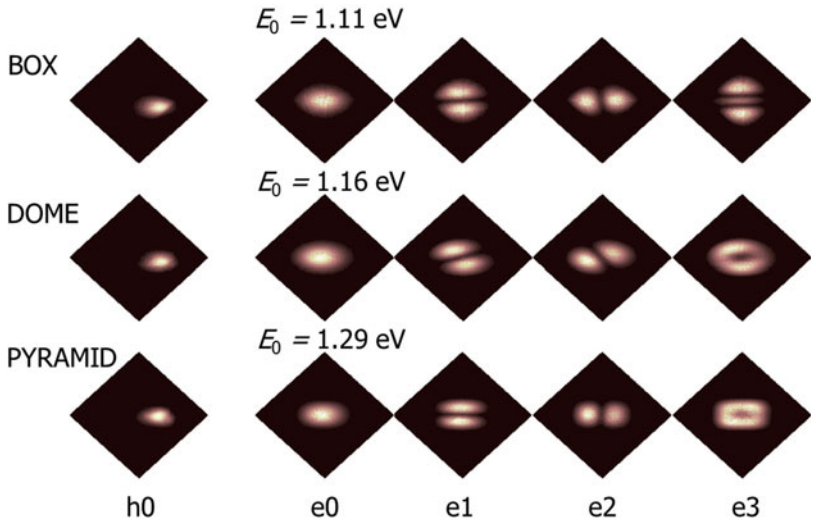


Fig. 35 Topmost valence band (HOMO) and first four conduction band wave functions including interfacial effects, strain, and piezoelectricity

DOI 10.1007/978-3-642-27737-5_343-2

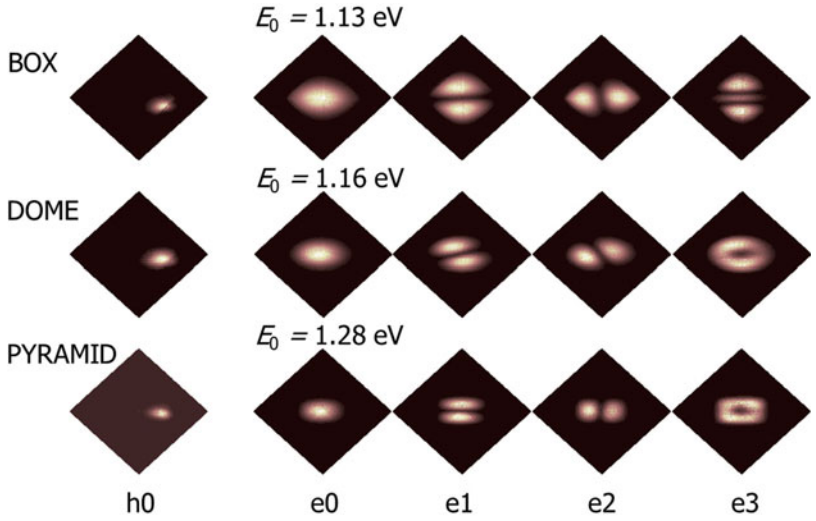


Fig. 36 Topmost valence band (HOMO) and first four conduction band wave functions including all four (4) competing internal fields originating from interfacial effects, strain, piezoelectricity, and pyroelectricity

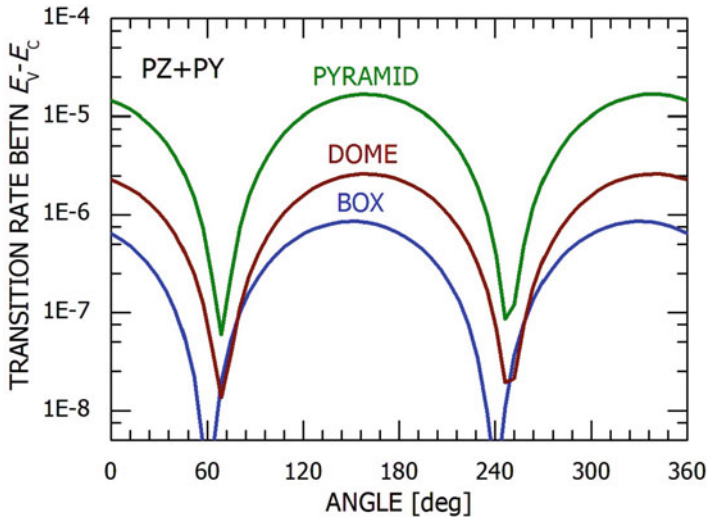


Fig. 37 Interband optical transition rates between ground hole (HOMO) and ground electronic states (LUMO) in all three quantum dots

Figure 37 shows the interband optical transition rates between ground hole (HOMO) and ground electronic states (LUMO) in all three quantum dots revealing significant suppression and strong polarization anisotropy (peak occurring at angles greater than zero) due to spatial irregularity (displacement) in the wave functions originating from the combined effects of all four internal fields. The true atomistic symmetry of the quantum dots, thus, influences the electronic bandstructure, and in general, the strengths of the optical transitions differ as a function of geometric shape of the quantum dot. The transition rates were found to be inversely proportional to volume of QD with values maximum in the pyramid and minimum for the box structures.

Importance of Second-Order Polarization in InAs/GaAs Quantum Dots

Objective. In the case of the InAs/GaAs quantum dots, the lattice mismatch is around 7 % and leads to a strong long-range strain field within the extended neighborhood of each quantum dot. Strain can be

DOI 10.1007/978-3-642-27737-5_343-2

atomistically inhomogeneous, involving not only biaxial components but also non-negligible shear components. Therefore, any spatial non-symmetric distortion in quantum dots (and other nanostructures) made of these materials will create piezoelectric fields, which will modify the electrostatic potential landscape. It is well known that the piezoelectric polarization is generally a *nonlinear* function of strain, the nonlinearity becoming important for large epitaxial strains. Recently, it has been shown (Bester et al. 2006a) that the piezoelectric polarization in strained InAs/GaAs systems has strong contributions from second-order effects that have so far been neglected. In this calculation, the piezoelectric tensor is given by (Bester et al. 2006a): $\tilde{e}_{\mu,j} = \tilde{e}^{0}_{\mu,j} + \sum_{k} \tilde{B}_{\mu,j,k}\eta_k$, where $\eta_j (j = 1, 6)$ is used to denote strain in the Voigt notation. Here, $\tilde{e}^{0}_{\mu,j}$ is the reduced proper piezoelectric tensor of the *unstrained* material, while $\tilde{B}_{\mu,j,k}$ is a fifth-rank tensor with Cartesian coordinates and μ is the strain index in Voigt notation j, k and represents the first-order change of the reduced piezoelectric tensor with strain. In Bester et al. (2006a), Bester et al. have found that, for [111]-oriented $In_xGa_{1-x}As$ quantum wells, the linear and the quadratic piezoelectric coefficients have the opposite effect on the field, and for large strains (large In concentration), the quadratic terms even dominate! Thus, the piezoelectric field turns out to be a rare example of a physical quantity for which the first-order and second-order contributions are of comparable magnitude.

In this section, we study the electronic properties of zinc blende InAs/GaAs quantum dots having three different geometries, namely, box, dome, and pyramid (as depicted in Fig. 6). In particular, for piezoelectricity, for the first time within the framework of $sp^3d^5s^*$ tight-binding theory, *four* different recently proposed polarization models (linear and nonlinear) have been considered in this study. In contrast to recent studies of similar quantum dots, our calculations yield a *nonvanishing* net piezoelectric contribution to the built-in electrostatic field.

Simulation Results. For the calculations of the piezoelectric polarization in InAs/GaAs QDs, we have considered *four* different models and followed the recipe in Bester et al. (2006a): (1) linear approximation using *experimental* (bulk) values for polarization constants (-0.045 C/m^2 for InAs and -0.16 C/m^2 for GaAs), (2) linear (first-order) approximation using microscopically determined values for polarization constants (-0.115 C/m^2 for InAs and -0.230 C/m^2 for GaAs), (3) second-order (quadratic) polarization using microscopically determined values for polarization constants ($\beta_{114} = -0.531$, $\beta_{124} = -4.076$, $\beta_{156} = -0.120$ for InAs and $\beta_{114} = -0.439$, $\beta_{124} = -3.765$, $\beta_{156} = -0.492$ for GaAs), and (4) a combination of the first- and the second-order effects using the abovementioned microscopically determined values for polarization constants.

The piezoelectric potential along the growth (z) direction using the four different models are shown in Fig. 38. From this figure, one can extract at least three important features: (1) piezoelectric potential has its largest *magnitude* in a pyramidal dot with the peak being located near the pyramid tip and the minimum in a box; (2) the spread of the potential is largest in a box and minimum in a pyramid; and (3) *within the quantum dot region*, the second-order effect has comparable/similar magnitude as the first-order contribution, and, indeed, the two terms oppose each other. However, noticeable is the fact that the first-order contribution, as compared to the quadratic term, penetrates *deeper* inside the surrounding material matrix. This particular effect, we believe, in contrast to the findings in Bester et al. (2006b), results in a nonvanishing and reasonably large *net* (first + second) piezoelectric potential within the region of interest. The fact that the first-order and the second-order terms oppose each other is also noticeable in Fig. 39, which depicts the surface plots of the piezoelectric potential distribution in the X–Y plane. Note that the first-order term has somewhat larger magnitude and spread than the quadratic term. Also, associated with both these two terms, noticeable is the asymmetry and inequivalence (in terms of potential magnitude and distribution) along the [110] and the [1$\underline{1}$0] directions.

Figure 40 shows the first 4 (four) conduction band wave functions for all three quantum dots including *both* the strain and the piezoelectric fields (fourth model) in the calculations. The piezoelectric potential

DOI 10.1007/978-3-642-27737-5_343-2

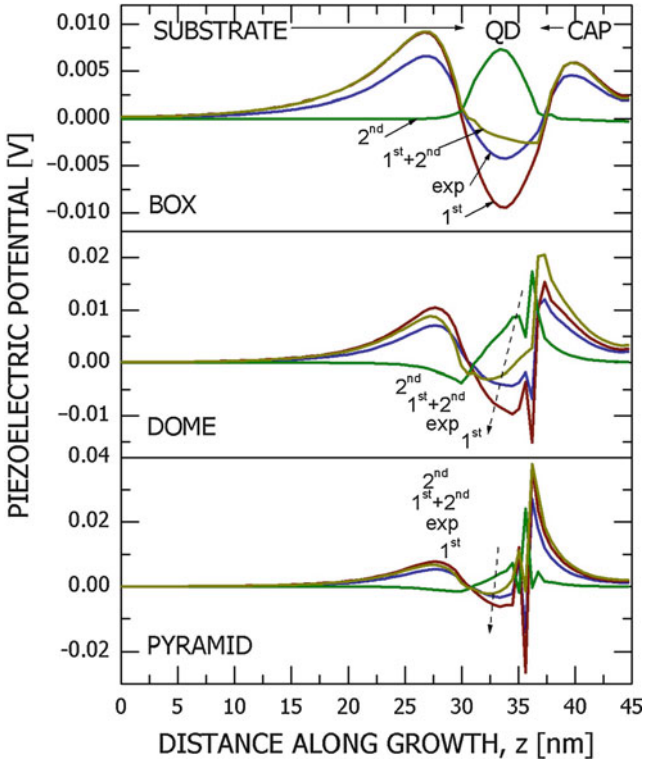


Fig. 38 Induced piezoelectric potential along the *z* (growth) direction in all three quantum dots. Four different models for the polarization constants have been used in the calculations: (1) linear and experimentally measured, (2) linear through ab initio calculations, (3) quadratic through ab initio calculations, and (4) combination of first- and second-order components. Also in this figure, note the varying spread/penetration of the potential in the surrounding material matrix as a function of dot shape (Data are taken from Sundaresan et al. 2010)

introduces a global shift in the energy spectrum and generally opposes the strain-induced field. In box- and dome-shaped dots, the net piezoelectric potential is found to be strong enough to fully offset the combined effects of interface and strain fields and, thereby, *flip* the optical polarization anisotropy. Also shown in this Figure are the splits in the P levels (ΔP) for all three quantum dots. In order to fully assess the piezoelectric effects, we have prepared Table 2 that quantifies the *individual* net contributions from crystal atomicity and interfaces, strain, and the various components of piezoelectric fields in the spilt of the P level. The net piezoelectric contribution is found to be largest in a box and minimum in a pyramid, which clearly establishes a direct correspondence between the piezoelectric potential and the volume of the quantum dot under study.

Figure 41 shows the influence of the three major types of internal fields on the single-particle conduction band ground states (*S* orbital) in the quantum dots. Due to size quantization, it is found that the ground energy increases as the volume of the quantum dot decreases. Noticeable is that the strain relaxation introduces pronounced *blue* shifts in the conduction band ground state, whereas the piezoelectricity causes a much smaller red shift therein.

Phosphorus (P) Impurity in Silicon

Physics of P Impurity. In a substitutional P impurity in Si, the four electrons from the outermost shell of P form bonds with the four neighboring Si atoms, while the fifth electron can ionize to the conduction band at moderate temperatures leaving a positively charged P atom with a Coulomb potential screened by the dielectric constant of the host. At low temperatures, this potential can trap an electron and form a

DOI 10.1007/978-3-642-27737-5_343-2

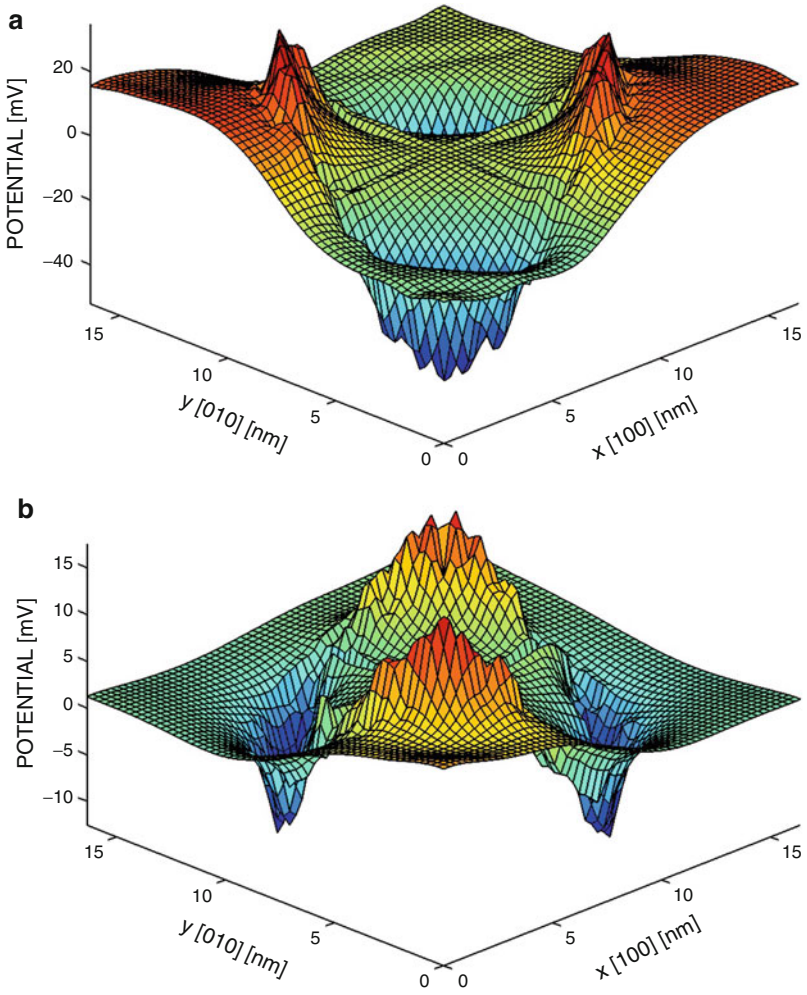


Fig. 39 Linear (**a**) and quadratic (**b**) contributions of the piezoelectric potential in the X–Y plane halfway through the dot height. Note the magnitude, orientation, and anisotropy in the induced potential

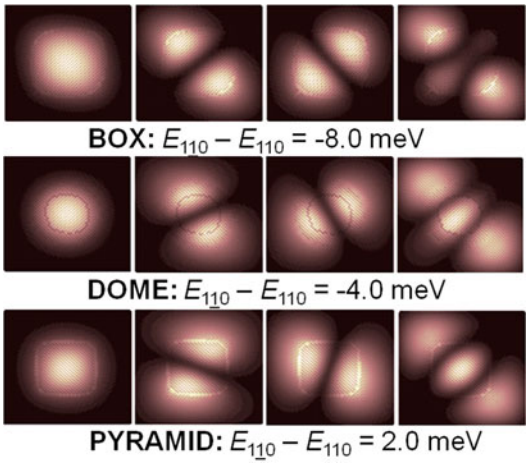


Fig. 40 First four-electronic wave functions and split in the P levels in all three quantum dots including atomicity/interfacial effects, strain, and piezoelectricity. Note the varying piezoelectric contributions, which can be attributed mainly to the volume of the quantum dot under study

DOI 10.1007/978-3-642-27737-5_343-2

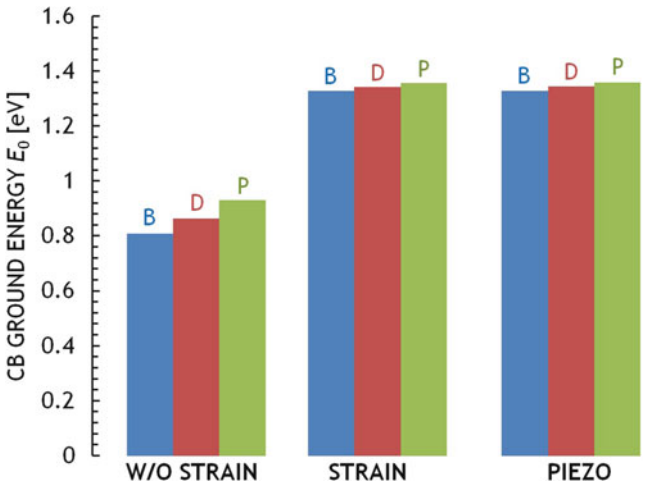


Fig. 41 Conduction band ground states in box (B)-, dome (D)-, and pyramid (P)-shaped quantum dots including interface effects (w/out strain), strain, and piezoelectricity (Data are taken from Sundaresan et al. 2010)

hydrogen-like system except that the sixfold degenerate conduction band valleys of Si give rise to a sixfold degenerate 1 S-type ground state. In practice, this sixfold degeneracy is lifted by strong coupling between the different valleys caused by deviations of the impurity potential from its Coulombic nature in the vicinity of the donor nucleus. If this so-called valley–orbit interaction is not taken into account, then the effective mass theory (EMT) predicts a P donor ground state binding energy of −33 meV as opposed to the experimentally measured value of −45.6 meV (Sellier et al. 2006). The influence of valley–orbit interaction is strongest for the six 1 s states and is negligible for the excited states, which are affected by the bulk properties of the host (Boykin et al. 2004a). The TB model considered here also models the excited states well by its accurate representation of the Si bandstructure. Hence, we limit our attention here to the effect of valley–orbit interaction on the 1 s states.

Study Approaches. Theoretical study of donors in Si dates back to the 1950s when Kohn and Luttinger (1995) employed symmetry arguments and variational envelope functions based on EMT to predict the nature of the donor spectrum and wave functions with a fair amount of success. Although many theorists who study donor-based nanodevices still use the Kohn–Luttinger variational envelope functions, recent approaches (Martins et al. 2004; Rahman et al. 2007; Wellard and Hollenberg 2005) have highlighted the need to consider a more extended set of Bloch states than the six-valley minima states and to go beyond the basic EMT assumptions for accurate modeling of impurities. For modeling high-precision donor electronics, it is very important to model the basic physics from a consistent set of assumptions and to obtain very accurate numbers in addition to correct trends. The model presented here serves these purposes well and can be used conveniently for large-scale device simulation.

Numerical Study of the Valley–Orbit Interaction. The inset of Fig. 42 shows the lowest six 1 s type energy states of a P donor in Si. When valley–orbit coupling is ignored, the six lowest states are degenerate in energy. When valley–orbit coupling is taken into account, the sixfold degenerate states split into a ground state of symmetry A1, a triply degenerate state of symmetry T, and a doubly degenerate state of symmetry E. Valley–orbit interaction, which arises due to the deviation of the impurity potential from its bulk-like Coulombic nature, is typically modeled by a correction term for the impurity potential in the vicinity of the donor site. The strength of this core-correcting potential determines the magnitude of the splitting of the six 1 s states and varies from impurity to impurity. Here, we consider a core-correcting cutoff potential U_0 at the donor site, reflecting a global shift of the orbital energies of the impurity. Figure 42 shows how the energy splitting is affected by the strength of U_0. For small U_0, the six 1 S-type states are degenerate in energy. As U_0 increases in magnitude, we obtain the singlet, triplet, and doublet

DOI 10.1007/978-3-642-27737-5_343-2

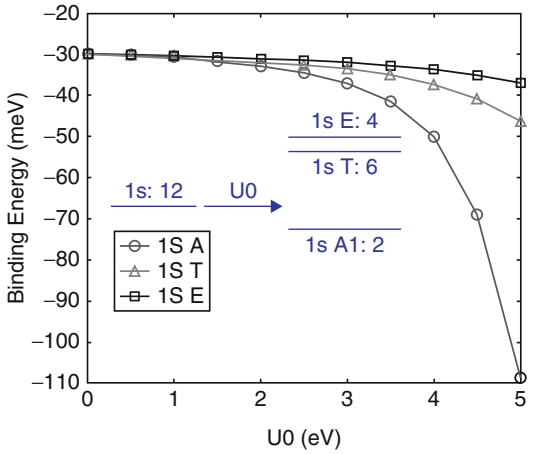


Fig. 42 Effect of central cell correction U_0 on energy splitting. (*Inset*) Group V donor 1 s states in Si splitting into three components due to valley–orbit interaction

components, as mentioned earlier. Since the triplet (and doublet) states remain degenerate irrespective of U_0, we only plot one state of the T (and E) manifold.

This single core-correction term was found to reproduce the donor eigenstates within a few meV and could be adjusted to match the donor ground state binding energy within a μeV. In general, however, the tight-binding parameters for Si can only reproduce the full bandstructure within a limited accuracy. To model high-precision donor electronics within a hundredth of a meV, as is needed in many quantum computing applications, additional core-correction terms are required. In semiempirical tight binding, it is only natural to adjust the on-site orbital energies of the P donor slightly from their Si counterparts to provide this additional correction. In Fig. 43, we show variation of the binding energy of the 1 s manifold as a function of the on-site orbital energies of the donor site. The four on-site energies considered are Es, Ep, Ed, and Es* corresponding to the s, p, d, and s* orbitals, respectively. The trends in the plots help us establish a recipe for optimizing the core correction for a donor species to reflect impurity eigenstates within the precision of 0.01 meV. For example, if the only donor ground state of A1 symmetry needs to have a higher binding energy, we can adjust either Es or Es*, each of which will push the A1 state deeper in energy without affecting the excited states (Fig. 43a, d). Figure 43c shows that both the triplet and the doublet state can be adjusted in energy by Ed without affecting the A1 state, while Fig. 43b shows that the triplet state alone is affected by Ep. On the other hand, U_0 reflects a global shift of all the on-site energies and can affect all the 1 s states, as already shown in Fig. 42b. In short, there are enough degrees of freedom to empirically adjust the core correction to obtain very exact eigenvalues. Once a set of these parameters (U_0, Es, Ep, Ed, Es*) is fixed, they can be used for a variety of applications like Stark shift, charge qubits, etc. without any additional modification. To model a generic impurity, it is recommended that U_0 be adjusted first so that the ground state binding energy is reproduced accurately. Then one can consider small deviations in a few of these on-site energies to fit the excited states accurately. In most cases, the parameters U_0, Ep, and Ed can be sufficient for accurate modeling. The plots here were obtained by the tight-binding $sp^3d^5s^*$ model without spin. Clearly, this is an empirical process that does not account fully for the different nature of the impurity atom in a host lattice. Additional mapping which includes the change of the impurity to host coupling matrices could be performed possible based on an input from an ab initio method.

Solution Methods: Lanczos and Block Lanczos. For a realistic simulation involving a few impurities, one needs to consider a lattice size of about seven million atoms. In atomistic tight binding with a 20 orbital nearest-neighbor model, this involves solving a Hamiltonian with 140 million rows and

DOI 10.1007/978-3-642-27737-5_343-2

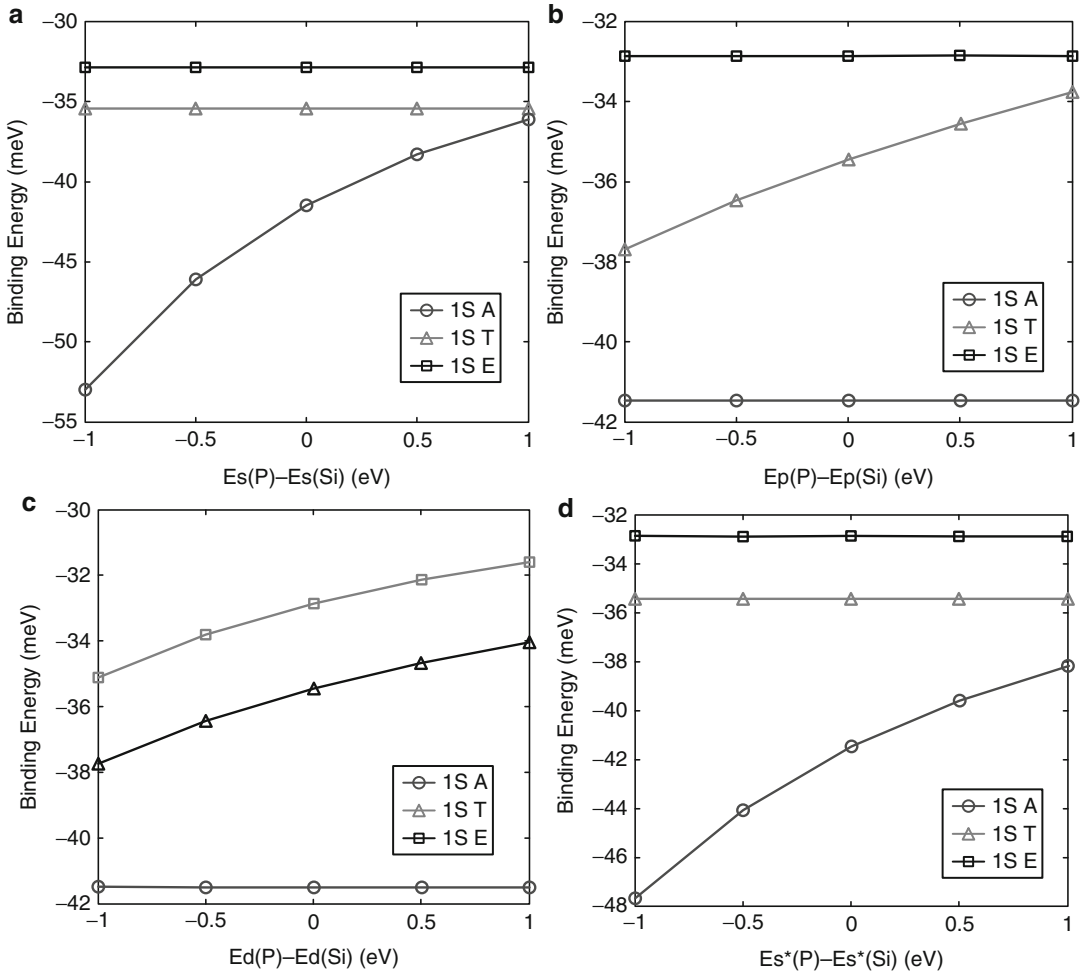


Fig. 43 Variation of 1 s binding energies with on-site orbital energies. The triplet (doublet) states remain degenerate. Hence, only one triplet (doublet) is shown

columns. Although this matrix is considerably sparse, solving for interior eigenvalues occurring near the conduction band poses a difficult problem. Compared to many other algorithms, the parallel Lanczos algorithm for eigen solution has proved very efficient. However, one drawback of the Lanczos algorithm is its inability to find degenerate and closely clustered eigenvalues with reliability. A blocked version of Lanczos resolves this problem at the cost of some additional computed time. Since there are many degenerate eigenstates present in the unperturbed impurity spectrum, the block Lanczos algorithm was a suitable solution method for the problem outlined here. Table 5 shows the comparison of eigenstates obtained from Lanczos, block Lanczos, and experimentally established values for single donors in bulk Si. While Lanczos fails to capture the degeneracy of the triplet, doublet, and spin states, block Lanczos resolves all the 12 eigenvalues reliably. The computational system considered here spans a domain of $30.5 \times 30.5 \times 30.5$ nm and contains about 1.4 million atoms. Closed boundary condition is applied in all three dimensions.

Typical Application: Donor-Based Charge Qubits. An impurity-based charge qubit involves a single electron bound to two ionized P donors in Si. A qubit can be encoded based on the localization of the electron in either of the two impurities (Hollenberg et al. 2004). When the Hamiltonian of such a system is solved, a set of bonding and antibonding states are obtained from the set of single impurity states. An important parameter in quantum computing applications is the tunnel coupling between the two lowest

DOI 10.1007/978-3-642-27737-5_343-2

Table 5 Comparison of the single-donor states relative to the conduction band minima of Si for Lanczos and block Lanczos algorithms. The Lanczos algorithm fails to capture degenerate states, while block Lanczos is able to resolve degeneracies at the expense of compute time. The eigenvalues were obtained by the $sp^3d^5s^*$ spin model and show spin degenerate eigenvalues as well. The slight deviation of the eigenvalues from the experimental values is due to the finite size (i.e., confinement effect) of the simulation domain of 30 nm^3

Experiment (Bae et al. 2007)	Lanczos	Block Lanczos (block size 6)	Symmetry
−45.59	−45.599	−45.599	1 s (A1)
−45.59		−45.599	1 s (A1)
−33.89	−33.932	−33.932	1 s (T)
−33.89		−33.932	1 s (T)
−33.89		−33.930	1 s (T)
−33.89		−33.930	1 s (T)
−33.89		−33.930	1 s (T)
−33.89		−33.930	1 s (T)
−32.58	−32.67	−32.670	1 s (E)
−32.58		−32.670	1 s (E)
−32.58		−32.670	1 s (E)
−32.58		−32.670	1 s (E)

eigenstates. This parameter depends on the separation of the two lowest eigenstates of the P_2^+ problem and is sensitive to relative donor placements and gate voltages. Figure 44 shows the tunnel coupling as a function of donor separation along [100] and [110] calculated in tight binding. The tunnel coupling tends to decay, as the impurities are located farther apart. While variation of tunnel coupling is found to be smooth along [100] direction, it is highly oscillatory along [110]. This is due to interference between Bloch parts of the impurity wave functions contributed by the Si crystal. These trends are already well established in literature (Hu et al. 2005) from effective mass theory. The impurity model in TB presented here is able to capture these effects with convenience.

Unlike EMT, the methodology developed here can consider a more extended Bloch structure of the host and incorporate many realistic device effects such as finite device sizes and interfaces under one framework and is convenient for large-scale device simulations. Treatment of such factors enables precise comparison with experimentally measured quantities, as was done in Rahman et al. (2007), where the hyperfine stark effect for a P donor was calculated in good agreement with experiment (Bradbury et al. 2006) and discrepancies with previous EMT (Friesen et al. 2005)-based calculations were resolved. Further work is under way to study CTAP (Greentree et al. 2004)-based architectures (Hollenberg et al. 2006) and charge qubits (Hollenberg et al. 2004; Hu et al. 2005; Koiller et al. 2006) and investigate donor–interface well hybridization in Si FinFET devices (Calderón et al. 2006; Lansbergen et al. 2008; Sellier et al. 2006).

Si on SiGe Quantum Well

Many quantum dot-based (Friesen et al. 2003) or impurity-based (Kalden et al. 2010) quantum computing architectures are proposed to be fabricated in Si/SiGe heterostructures. Since silicon has multiple degenerate values, it is critical to engineer these degeneracies out of the system to avoid dephasing of qubits. Miscut substrates (Fig. 45b) as opposed to flat substrates (Fig. 45a) are often used to ensure uniform growth of Si/SiGe heterostructures. However, a miscut modifies the energy spectrum of a QW. In a flat QW, the two degenerate valleys in strained Si split in energy and the valley minima occur at $\pm k_x = 0$. Valley splitting (VS) in a flat QW is a result of interaction among states in bulk z-valleys centered at $k_z = k_0$, where k_0 is position of the valley minimum in strained Si. The energy splitting between these two

DOI 10.1007/978-3-642-27737-5_343-2

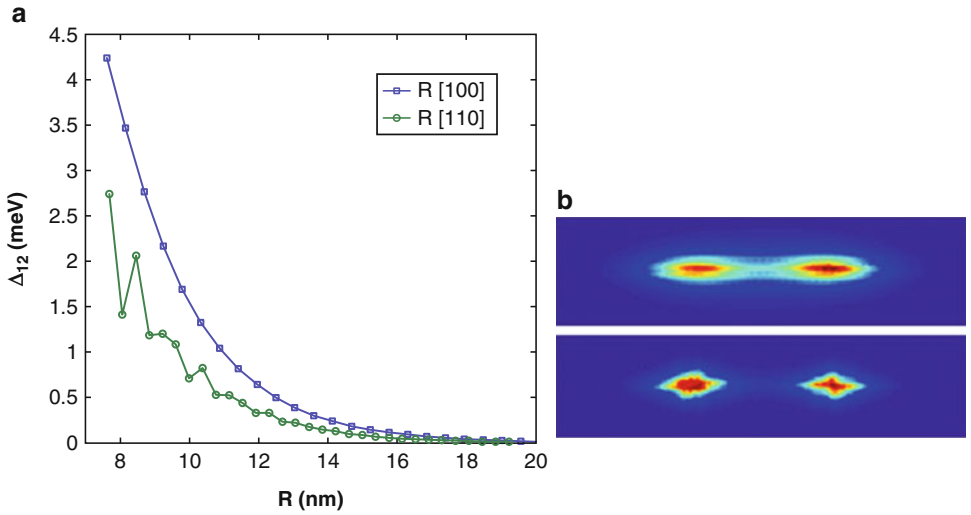


Fig. 44 (**a**) Variation of tunnel coupling for a P_2+ system with impurity separation along [100] and [110]. (**b**) Formation of molecular states for P_2+

lowest-lying valleys is called as valley splitting (VS). In quantum computing devices, VS is an important design parameter controlling the electron spin decoherence time (Boykin et al. 2004b, c, 2005). In a miscut QW, lowest-lying valleys are degenerate with minima at $\pm k_{x'}^0$ (Kharche et al. 2007). Thus, atomic-scale modulation of surface topology leads to very different electronic structures in flat and miscut QWs. As a consequence of this, flat and miscut QWs respond differently to the applied electric and magnetic fields. In the presence of lateral confinement in miscut QW, the two degenerate valleys in Fig. 45c interact and give rise to VS.

Traditional magnetic probe techniques such as Shubnikov–de Haas oscillations are used to measure energy spectrum of QWs. Valley and spin splittings are determined by electron valley resonance (EVR) (Goswami et al. 2007) and electron spin resonance (ESR) (Dobers et al. 1998) techniques. In these measurements, in plane (lateral) confinement of the Landau levels is provided by the magnetic field. Figure 45d shows the dependence of VS on applied magnetic field in flat and ideal 2^0 miscut QWs. Ideal miscut QWs refer to the miscut QWs with no step roughness. VS in flat QW is independent of magnetic field because in these QWs VS arises from z-confinement provided by the confining SiGe buffers (Kharche et al. 2007). In miscut QWs, however, VS is the result of the combined effect of the two confinements, the z-confinement provided by the SiGe buffers and the lateral confinement provided by the applied magnetic field. The two degenerate valleys centered at $\pm k_{x'}^0$ along x' direction in the miscut QWs (Fig. 45c) interact and split in the presence of magnetic field. At low magnetic fields, the dependence of VS in miscut QWs on the applied magnetic field is linear. In calculations of Fig. 45c, d, QWs are assumed to be perfect. Disorders such as step-roughness and alloy disorder in SiGe buffer which are inherently present in the experiments are completely ignored. As a result, calculated VS is nearly an order of magnitude lower than the experimentally measured values (Fig. 46d).

Miscut substrates undergo reconstruction to reduce the surface free energy which gives rise to the step roughness (Zandviet and Elswijk 1993; Fig. 46b, c). This type of step-roughness disorder is present at the Si/SiGe interface. Another type of disorder in Si/SiGe heterostructures is the random alloy disorder in SiGe buffer. These two disorders are always present in actual QW devices and thus need to be taken into account in VS computations. Schematic of an electronic structure computation domain is shown in Fig. 46a. QWs extend 15 nm along y direction to take into account the step-roughness disorder shown in Fig. 46c. x' confinement due to the magnetic field is incorporated through the Landau gauge

DOI 10.1007/978-3-642-27737-5_343-2

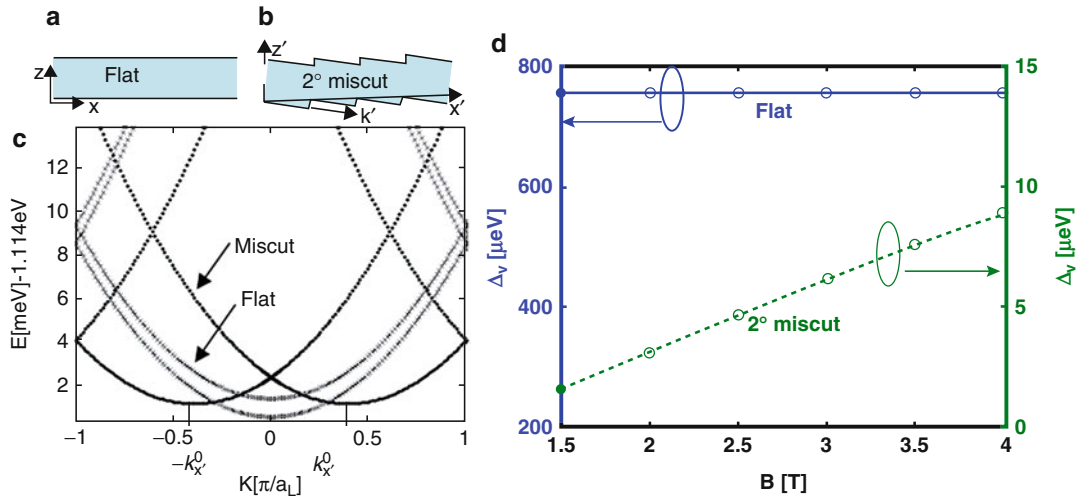


Fig. 45 (**a**) Schematic of a Si QW grown on [001] substrate. The crystal symmetry directions are along *x [100]* and *z [001]*. (**b**) Schematic of a 2° miscut QW unit cell. The unit cell is periodic along x' and y directions and confined in z' direction. Miscut angle is 2°. The step height is *a/4* which corresponds to one atomic layer, where a is lattice constant. (**c**) Bandstructure of 5.26-nm-thick flat/miscut QW along x/x' direction. Flat QW shows the presence of two non-degenerate valleys separated by an energy know as VS. Miscut QW shows the presence of two degenerate valleys centered at $\pm k_{x'}^0$. (**d**) VS in 10-nm-thick flat (001) and 2° miscut Si QWs. Electric field in z direction is 9 MV/m

$\left(\vec{A} = Bx\hat{y}\right)$. The resulting vector potential $\left(\vec{A}\right)$ is introduced into the tight-binding Hamiltonian through the gauge invariant Peierls substitution (Boykin et al. 2001; Boykin and Vogl 2001; Graf and Vogl 1995). Closed boundary conditions are used in x and z directions while y direction is assumed to be (quasi-) periodic. The confinement induced by closed boundary conditions in x' direction compete with the magnetic field confinement. The lateral extension of the strain and the electronic structure domain is set to 150 nm, which is about seven times larger than the maximum magnetic confinement length in a 2DEG at B = 1.5 T ($\simeq$21 nm). For the magnetic field ranges of 1.5–4 T, confinement is dominated by the magnetic field, and no lateral x-confinement effects due to the closed boundary conditions are visible in simulations. Modulation doping in Si/SiGe heterostructures induces built-in electric field. In the simulations performed here, constant electric field of 9 MV/m is assumed in the QW growth direction.

Figure 46d shows the computed VS in two miscut QWs. VS in ideal miscut QWs is an order of magnitude lower compared to the experimentally measured values. If the step-roughness disorder is included in the simulations, the computed VS is higher compared to that of an ideal miscut QW. In these calculations, surface roughness model of (Bester et al. 2006a) is used, and the uniform biaxial strain of $\varepsilon_{||} = 0.013$ which corresponds to $Si_{0.7}Ge_{0.3}$ buffer composition is assumed. This VS, however, is slightly smaller than the experimentally measured VS. This discrepancy can be answered by adding SiGe buffers in the electronic structure simulation domain. Three nanometer of SiGe buffer is included on top and bottom of the Si QW to take into account the wave function penetration into the finite barrier QW buffers. Strain computation domain has the same x and y dimensions as the electronic structure domain. To take into account the long-range nature of strain (Korkusinski et al. 2005a) 40 nm of SiGe buffer is included on both sides of Si QW. z-dimension of the strain domain is 90 nm. Valence force field (VFF) model of Keating (Keating 1966) is employed to calculate the relaxed geometries. The VS computed by taking both step and alloy disorder into account is found to match closely to the experimentally measured values.

The time required to compute the ten million atom strain calculation on 20 CPUs of an Intel x86-64 dual core Linux cluster is about 9 h. The subsequent two million atom electronic structure calculation requires 10 h.

DOI 10.1007/978-3-642-27737-5_343-2

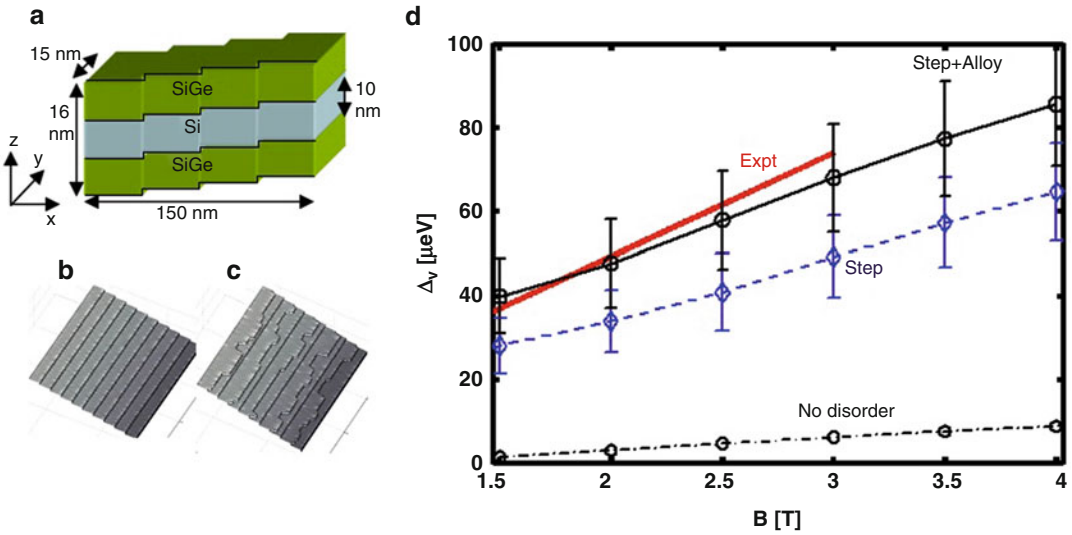


Fig. 46 (**a**) Schematics of the simulation domain. (**b**) Ideal steps on a miscut substrate. (**c**) Step disorder resulting from the surface reconstruction on the miscut substrate. (**d**) VS of the first Landau level in a 10-nm-thick strained Si QW. The VS labeled as "No disorder" is shown for comparison, and it is same as that of in Fig. 1d. VS increases due to the step disorder. When alloy disorder in SiGe buffer is included along with the step disorder, the computed VS matches the experimentally measured values. Error bars represent the standard deviation in VS. In the calculations of VS labeled as "No disorder" and "Step disorder," uniform biaxial strain of $\varepsilon_{||} = 0.013$ is assumed

SiGe Nanowires

Semiconductor nanowires are being actively investigated as the potential candidates for the end of the semiconductor technology roadmap devices. They are also attractive for sensing applications due to their high surface-to-volume ration. Several researchers have recently demonstrated the nanowire field-effect transistors (FETs) fabricated from pure elemental or compound semiconductors like Si (Cui et al. 2001), Ge (Greentree et al. 2004), and GaAs (Persson et al. 2004) as well as semiconductor alloys like SiGe (Kim et al. 2007) and their III–V counterparts. For the device design at the nanoscale, it is important to understand and to be able to predict transport properties of nanowires. Atomistic disorders such as alloy disorder, surface roughness, and inhomogeneous strain strongly influence the electronic structure and the charge transport in nanoscale devices. To simulate nanodevices, traditional effective mass approaches should be abandoned (Wang et al. 2005), and more accurate atomistic approaches should be adopted. Here, SiGe alloy nanowires are studied from two different perspectives. First, the electronic structure where bandstructure of a nanowire is obtained by projecting out small cell bands from a supercell eigenspectrum (Boykin et al. 2007c; Boykin and Klimeck 2005) and, second, the transport where transmission coefficients through the nanowire are calculated using an atomistic wave function (WF) approach (Boykin et al. 2007a; Luisier et al. 2006).

SiGe random alloy nanowires have two types of disorders: atom disorder due to random alloying and inhomogeneous strain disorder due to different Si–Si, Ge–Ge, and Si–Ge bond lengths. These disorders break the translational symmetry in semiconductor alloy nanowires. Thus, one runs into the problem of choosing a unit cell for the bandstructure calculation. Disorder can be taken into account by simulating larger repeating units (supercells) containing many small cells (Fig. 45). The nanowire bandstructure obtained from the supercell calculation is folded. The one-dimensional version of the zone-unfolding method (Boykin et al. 2007c; Boykin and Klimeck 2005) is used to project out the approximate eigenspectrum of the nanowire supercell on the small cell Brillouin zone. The probability sum rule (Boykin et al. 2007c) is used to extract the approximate bandstructure of the alloy nanowire from

DOI 10.1007/978-3-642-27737-5_343-2

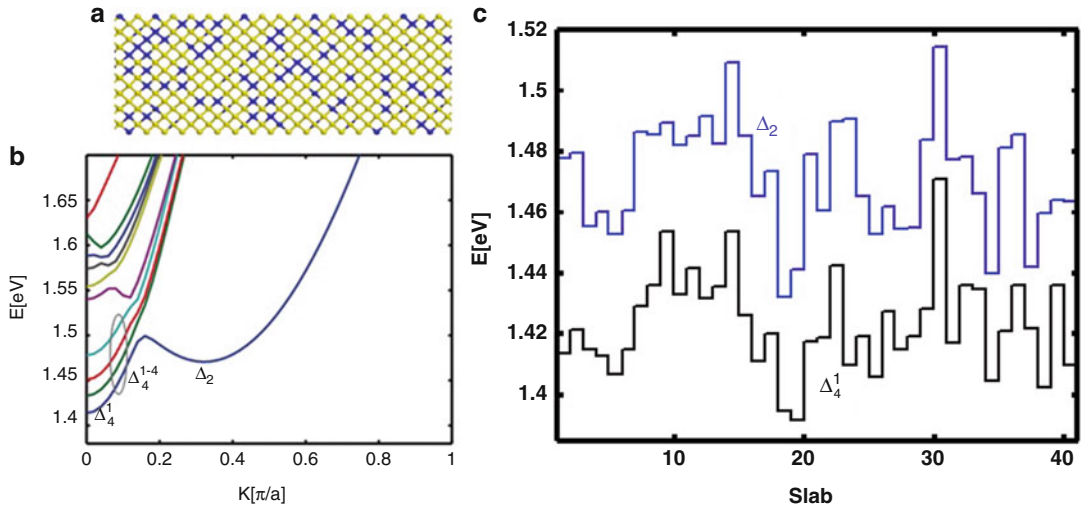


Fig. 47 (**a**) Atomistically resolved disorder in the $Si_{0.8}Ge_{0.2}$ 40 × 6 × 6 nanowire. (**b**) Conduction bandstructure of the first slab assuming that the slab is repeated infinitely. Δ_4 valleys are split into four separate bands. Δ_2 valley bands are doubly degenerate. (**c**) Band-edge minima of lowest energy Δ_4 and Δ_2 valleys along length of the nanowire

the projected probabilities. The small cell bandstructure obtained by this method captures the effect of SiGe alloy disorder on the electronic structure.

The nanowire geometry is specified in terms of conventional zinc blende (cubic) unit cells as $n_x \times n_y \times n_z$ where n_i is the number of cubes in direction i. The wire dimensions are 40 × 6 × 6 (22.3 × 3.3 × 3.3 nm), i.e., it is constructed from 40 1 × 6 × 6 slabs along [100] crystallographic direction. Figure 47a depicts a sliver cut through the center of the SiGe nanowire indicating the atomistically resolved disorder of the wire. Only the central 5-nm-long portion of this 22-nm-long wire is shown for visualization purpose. All electronic structure and transport calculations have been done in 20-band $sp^3d^5s^*$ tight-binding model with spin–orbit coupling. The bulk tight-binding and strain Si and Ge parameters are taken from Boykin et al. (2004a, 2007b). Relaxed wire geometries are calculated from valence force field approach.

The unfolding procedure to compute an approximate bandstructure from the large supercell calculation requires many eigenvectors. In practice, these eigenstates are closely spaced in energy, and Lanczos algorithm requires about 50,000 iterations to resolve 575 states in the energy range of interest. Such calculations require about 5.5 h on 30 cores of an Intel x86-64 dual core Linux cluster machine.

Figure 47b shows the conduction bandstructure of the first slab out of 40 slabs along nanowire length. Δ_4 valleys are split into four separate bands, while Δ_2 valley bands are doubly degenerate. Local band-edge plots of the lowest Δ_4 and Δ_2 valley minima are shown in Fig. 47c. This so-called local bandstructure of each slab is calculated assuming that this slab repeats infinitely along the nanowire. Due to fluctuations in atomic arrangements along the nanowire length, one expects to see the different bandstructures for each slab as shown in Fig. 48a. Variations of band edges along the nanowire length cause reflections which lead to the formation of the localized states and peaks in transmission plots.

The NEMO 3D team is currently developing with Mathieu Luisier at ETH Zurich a new 3D quantum transport simulator (Luisier et al. 2006). Here, we show a comparison of a 3D disordered system transport simulation with a NEMO 3D electronic structure calculation. The transmission coefficient (Fig. 48b) shows the noisy behavior due to random SiGe alloy disorder and inhomogeneous strain disorder in the wire. Steps in the transmission plot can be roughly related to the unfolded bandstructure (Fig. 46a) from supercell calculations. Four separate Δ_4 valley bands appear as a single band with a finite energy spread in the projected bandstructure. These four bands turn on near 1.44 eV which corresponds to the conduction

DOI 10.1007/978-3-642-27737-5_343-2

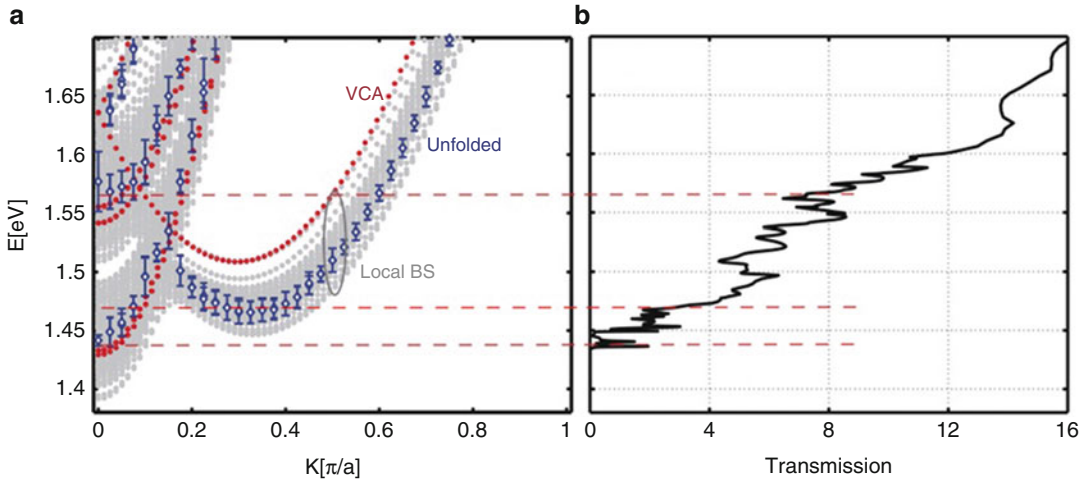


Fig. 48 (**a**) Bandstructures of 40 × 6 × 6 Si0.8Ge0.2 alloy nanowire in local bandstructure (*gray*), VCA (*red*), and zone-unfolding (*blue*) formulations. (**b**) Transmission through 40 × 6 × 6 wire. Steps in transmission are identified as resulting from new bands appearing in projected bandstructure. Note that atomistic, narrow 1D wires result automatically into 1D localization

band transmission turn on. Two Δ_2 valley bands turn on near 1.47 eV which leads to a step in the transmission. Four more channels due to higher Δ_4 valley sub-bands turn on near 1.57 eV. These transmission features cannot be related to the conventional virtual crystal approximation (VCA) bandstructure shown in Fig. 46a. Peaks in the transmission plot can be related to the local density of states in the wire (Kharche et al. 2008).

Projected supercell bandstructures and atomistic transport calculations are found to be complimentary and mutually supporting. Both methods provide better insight into the transport through the disordered nanowires.

Graphene

The high bulk mobility of graphene makes it a potential material for the active regions of future nanotransistors (Chen et al. 2008; Novoselov et al. 2004, 2005a, b; Zhang et al. 2005). Because bulk graphene has no gap, it is less useful than armchair graphene nanoribbons (AGNRs), especially those under 10-nm wide, which can have measurable gaps (Li et al. 2008; Nakada et al. 1996; Wang et al. 2008, 2009b). Specifically, field-effect transistors with AGNR channels are expected to have improved ON/OFF ratios. Accurate, computationally efficient models for both bulk graphene and AGNRs are thus of interest.

An ideal graphene sheet or AGNR is perfectly flat, resulting in a bandstructure which presents both challenges and opportunities from the perspective of device modeling and design. In a minimal tight-binding model which neglects the miniscule spin–orbit coupling in carbon (sp^3), the bands separate into two noninteracting sets: $\pi : \{p_z\}$ and $\sigma : \{s, p_x, p_y\}$. The bands involved in transport are the π-bands, and this recognition led Wallace (Wallace 1947) to develop the simplest possible model for graphene, one with a single p_z orbital per atom. Since there are two atoms per unit cell, this model is a two-band model. It has obvious attractions in terms of computational efficiency. Unfortunately, it is not especially accurate for bulk and for AGNRs it is worse. In bulk, the bands are perfectly and artificially symmetric from the Dirac point (band crossing at K) to the M point (Castro Neto et al. 2009). Perhaps its most serious drawback for AGNRs is that it does not permit a realistic hydrogen passivation approach. If hydrogen is modeled with a single s orbital, then there is no coupling at all to the π-bands and hence no passivation. AGNR bandstructure calculations in the single p_z orbital model are thus for unpassivated devices, which are not nearly as useful as passivated AGNRs. An accurate passivation model is essential, because the gaps of

the three different AGNR families $(3n, 3n+1, 3n+2)$ exhibit very different behavior versus width (Boykin et al. 2011). For example, those in the $3n + 2$ family have zero gap when unpassivated. The single p_z orbital model is thus not well-suited for accurate, realistic AGNR modeling.

These problems are overcome by expanding the basis for the π-bands to include the d orbitals: $\{d_{yz}, d_{zx}\}$. These orbitals only couple to each other and the p_z orbitals when the very small spin–orbit interaction in carbon is neglected (Boykin et al. 2011). Because the resulting basis remains small, it is still computationally inexpensive, yet it gives an impressive increase in accuracy. For passivation, the hydrogen basis set also includes all three orbitals, which permits a realistic, accurate passivation model. Including the d orbitals for hydrogen is quite reasonable, since in atomic hydrogen the $n = 2$ and $n = 3$ levels are separated by less than 2 eV. More importantly, when the hydrogen basis consists of only a single p_z orbital, it has not proven possible to fit the bands of all three families of passivated AGNRs. This difficulty most likely arises because in such a model the d components of the wave function, unlike the p_z components, suffer an abrupt termination and are effectively unpassivated (Boykin et al. 2011).

Figure 49a shows the bands of graphene as calculated using density functional theory with GW corrections, implemented in the ABINIT package (Gonze et al. 2005, 2009; Hybertsen and Louie 1986). Figure 49b shows a close-up of the bands from K to M as calculated with DFT + GW (symbols), the p/d model (solid lines), and the single p_z model (dotted lines). The parameters of the p/d model are fit to the DFT + GW results. The K points of all three models are aligned for comparison (there is no absolute energy in DFT). The p/d model accurately reproduces the DFT + GW bands, while the single p_z model is artificially symmetric. The reproduction of the band edges at M is particularly bad, and as discussed in (Boykin et al. 2011), artificial symmetries such as this are mathematically guaranteed by the one-orbital-per-atom basis.

The advantages of the p/d model with its realistic hydrogen passivation approach are even more apparent when it is used to calculate the bands of AGNRs. The gaps of the three families $(3n, 3n+1, 3n+2)$ have very different behavior, as shown in Fig. 50a; the width designation is that of (Yalavarthi et al. 2011). Without passivation, in the single p_z model, the gaps of the $3n + 2$ family are zero, contrary to DFT calculations (Boykin et al. 2011; Yalavarthi et al. 2011). In addition, the other two families' gaps are significantly off. As discussed in (Boykin et al. 2011), a model with only a single p_z orbital per hydrogen atom will not fit the behavior of all three passivated AGNR families. On the other hand, using the full p/d basis for hydrogen allows an accurate fit of all three families across a range of widths. It has been found that hydrogen passivation parameters fit to only the AGNR-7, -8, and -9 series provide excellent reproduction of the gaps for a broad range of widths (Boykin et al. 2011).

Figure 50b–d shows the bands of the series AGNR-11, -12, and -13 as calculated with DFT-LDA (symbols), the p/d model (solid lines), and the single p_z model (dotted lines). As discussed previously, the DFT-LDA and p/d model are hydrogen passivated while the single p_z model is not (Boykin et al. 2011). In the passivated AGNRs, the hydrogen atoms are at the ideal carbon–carbon bond angle, and all carbons are in their ideal positions, so that there is no edge disorder. The local density approximation (LDA) is used with DFT because the gap enhancement with GW corrections for a nanoribbon in vacuum is very large (Yalavarthi et al. 2011), due to reduced screening and enhanced Coulomb effects. These effects are significantly reduced for AGNRs on substrates, precisely the case in any useful device, so that the LDA gaps tend to be more representative of those in actual devices. In the figures, the uppermost valence band edges of all AGNRs are aligned for comparison (Boykin et al. 2011). Rather remarkably, the p/d results accurately reproduce the DFT-LDA bands throughout the zone, even though the hydrogen passivation parameters were fit to only the gaps of the AGNR-7, -8, and -9 series. In the single p_z model, note the incorrect gaps and especially the erroneous zero gap for AGNR-11. The single p_z model is therefore unsuited to accurate AGNR modeling. On the other hand, the p/d model produces highly accurate

DOI 10.1007/978-3-642-27737-5_343-2

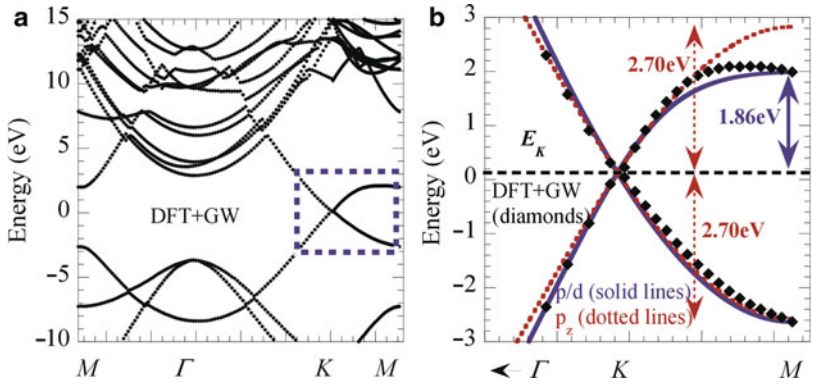


Fig. 49 (**a**) DFT + GW bands of bulk graphene. The critical part of the Brillouin zone from *K* to *M* is highlighted by the box with a blue dashed line border. (**b**) Bands of bulk graphene as calculated with DFT + GW (*black solid symbols*), the *p/d* model (*solid blue lines*), and the single p_z model (*dotted red lines*). The *p/d* model is parameterized to fit the DFT + GW bands, and it correctly reproduces the asymmetry at *M*. The single p_z model, on the other hand, is artificially symmetric. The *K* points of all three models are aligned (Data are taken from Boykin et al. 2011)

passivated AGNR bands at only a modest computational expense and is thus an attractive approach for simulating nanoFETs with AGNR channels.

Expanding the basis of the π-bands to include *d* orbitals allows for reproduction of the ab initio bands, which is impossible with the commonly employed single p_z model. More importantly, the *p/d* model accommodates a realistic passivation treatment which is critical for accurate AGNR modeling since most nanoribbons used in devices will be passivated, yet is still computationally efficient. The *p/d* model is thus well-suited for graphene-based nanodevice modeling.

Polarization Properties of Single and Bilayer InAs/GaAs Quantum Dot Systems

As we have seen in Section A and Section B, the polarization properties of QDs strongly depend on the orientation of the electron and hole wave functions, which are determined by the asymmetric nature of the interface between the QD and the surrounding GaAs buffer, by strain and by piezoelectric fields. In this section, we theoretically and experimentally explore the wavelength and polarization properties of independent layers (hereafter referred to as single quantum dot layers (SQD)) and two closely stacked layers (bilayers) of InAs/GaAs QDs. In contrast to a recent study of single InAs QD layers (Jayavel et al. 2004), our experimental and theoretical results (Usman et al. 2011a) indicate an increase in the TE/TM ratio when the upper QD layer in the bilayer is covered by an InGaAs strain-reducing capping layer (SRCL). This increase is due to biaxial strain-induced heavy hole (HH)–light hole (LH) splitting, which increases in the presence of an InGaAs SRCL. Our experimental measurements and theoretical calculations indicate that an InGaAs SRCL can only redshift the optical wavelength and does not reduce the TE/TM ratio, required to achieve an isotropic polarization response.

Quantum Dot Geometries. The cross-sectional transmission electron microscopy (TEM) images in Fig. 51a, b indicate that the seed layer QD in the bilayer samples has a diameter of ∼20 nm and a height of ∼7 nm. The upper QD with a GaAs cap (Fig. 51a) has a diameter of ∼30 nm and a height of ∼8 nm. The InGaAs cap tends to increase the height of the upper QD due to the reduced out-diffusion of indium from the QDs during capping, resulting in slightly taller QDs with a height of ∼10 nm, as can be seen in Fig. 51b. Schematics of the model QDs in the SQD and bilayers are shown in Fig. 51c–e. We consider in Fig. 1c an InAs SQD embedded in a GaAs buffer, in Fig. 1d a bilayer consisting of two InAs QDs embedded in a GaAs matrix, and in Fig. 1e a bilayer consisting of two InAs QDs embedded in a GaAs matrix but with the upper QD capped with a 4-nm $In_{0.26}Ga_{0.74}As$ SRCL. The dimensions of the SQD are same as those of the QDs in the seed layer of the bilayers.

DOI 10.1007/978-3-642-27737-5_343-2

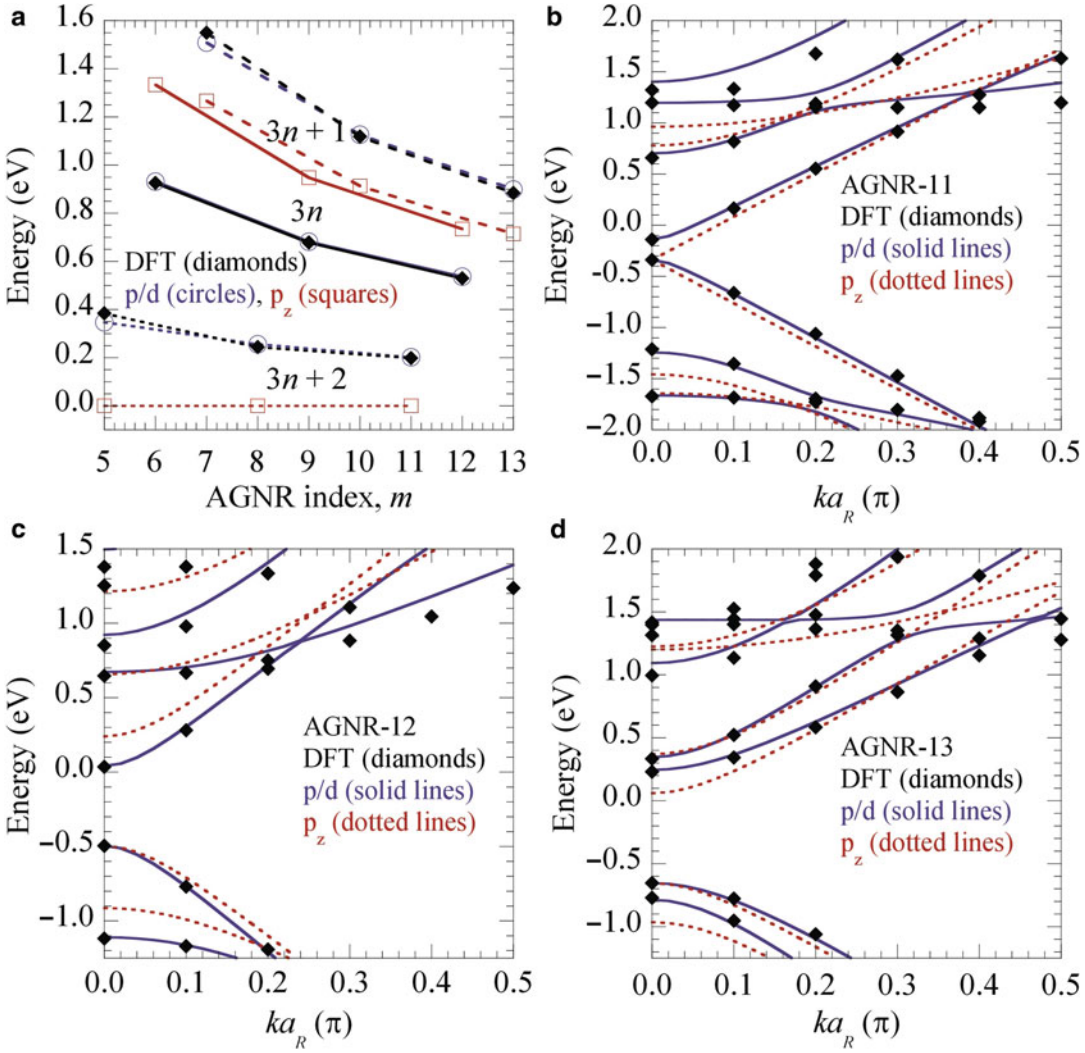


Fig. 50 (**a**) Gaps of AGNRs as calculated with DFT-LDA (*black solid diamonds*), the *p/d* model (*blue open circles*), and the single p_z model (*red open squares*). Lines are a guide to the eye. The DFT-LDA and *p/d* nanoribbons are hydrogen passivated, while those in the single p_z model are not (see text). (**b**–**d**) Bands of AGNRs (11, 12, and 13) as calculated with DFT-LDA (*black solid diamonds*), the *p/d* model (*blue solid lines*), and the single p_z model (*red dotted lines*). The uppermost valence band edge for each model is aligned to facilitate comparison (Data are taken from Boykin et al. 2011)

Only the Upper QD Is Optically Active. In bilayer QDs, the upper QD is slightly larger than the lower one (Xie et al. 1995). Strain due to the upper QD tends to increase the energy of the seed QD energy levels and thus the lowest electron and highest hole energy levels are confined in the upper QD. Due to a large separation between the QD layers (10 nm), the electron and hole wave functions do not form hybridized molecular states. Such hybridized states can be observed for closely stacked QDs, separated by ∼6 nm or less (Usman et al. 2011a) as well as being observed by applying an external electric field (Usman et al. 2009). In both our bilayers, the first three electron and hole energy levels are confined to the upper QD. The first electron energy level in the lower QD is the fourth energy level, E4, and the sixth energy level, E6, in the GaAs and InGaAs SRCL-capped bilayer, respectively. In these bilayers, the upper QD serves as an optically active layer for ground state optical emission. The lower QD does not contribute to optical emission for reasonably low carrier occupation, in agreement with previous photoluminescence (PL) measurements (cf Ru et al. 2003).

DOI 10.1007/978-3-642-27737-5_343-2

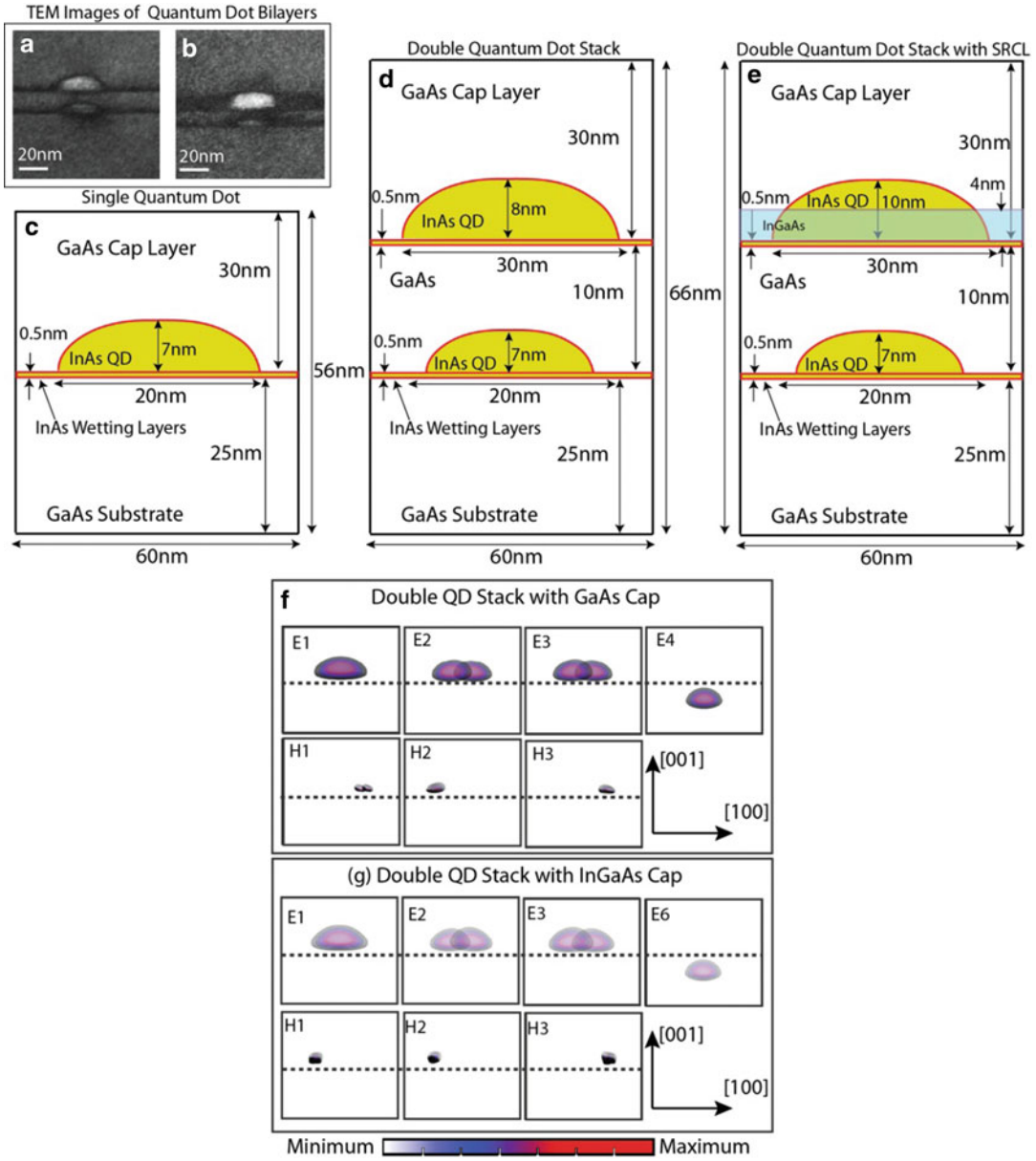


Fig. 51 (**a**, **b**) TEM Images of the bilayer QD stacks without (**a**) and with (**b**) the SRCL capping. The SRCL overgrowth clearly tends to increase the size of the QD. (**c**) Geometry of a single InAs QD lying on top of a 0.5-nm InAs wetting layer inside GaAs buffer. (**d**) Geometry of a two-InAs-QD vertical stack surrounded by GaAs buffer. Both QDs are lying on the top of 0.5-nm wetting layers. The separation between the wetting layers is 10 nm. The upper QD in the stack is larger than the lower QD due to strain-driven self-assembly process. (**e**) Geometry of a two-InAs-QD vertical stack surrounded by GaAs buffer. Both QDs are lying on the top of 0.5-nm wetting layers. The separation between the wetting layers is 10 nm. The upper QD is first covered by a 4-nm $In_{0.26}Ga_{0.74}As$ strain-reducing layer before depositing the GaAs capping layer. The upper QD has slightly larger size than the upper QD of the stack without SRCL (part (**d**)) due to reduced In segregation effect. (**f**) Wave function plots of first four-electron and first three-hole energy levels of the two QD stack (part (**e**)). The *blue* and the *red* colors show the intensity of magnitude (*blue* is lowest and *red* is highest). The *dotted horizontal line* is marked to guide the eyes and separates the upper and the lower QDs. All of the first three electrons and first three holes are in the upper QD indicating that the upper QD is optically active, whereas the lower QD remain inactive. (**g**) Wave function plots of four-electron and three-hole energy levels of the two QD stack with the SRCL cap (part (**e**)). The *blue* and the *red* colors show the intensity of magnitude (*blue* is lowest and *red* is highest). The *dotted horizontal line* is marked to guide the eyes and separates the upper and the lower QDs. All of the first three electrons and first three holes are in the upper QD indicating that the upper QD is optically active whereas the lower QD remains inactive and serves to control the optical properties of the upper QD through the strain coupling. The first electron state in the lower QD is E_6 (Reprinted with permission from (Usman et al. 2011a) copyright 2011, AIP Publishing LLC)

Redshift of Optical Emissions to 1,300 nm and Beyond. Figure 52 compares the room-temperature electroluminescence (RT EL) spectra obtained from the SQD to that measured on a GaAs-capped bilayer. The dotted lines are the electron–hole transition energies computed from the NEMO 3D simulations, which closely match the experimental data. A redshift of ∼77 nm is observed between the SQD and GaAs-capped bilayer samples. This is because of the larger height of the upper QD in the bilayer as well as the strain coupling between the two layers of bilayer, which tend to reduce the optical gap and redshifts the optical emission. Figure 52 also compares the RT EL spectra measured on the bilayers with and without the InGaAs SRCL. An InGaAs SRCL relaxes the hydrostatic strain and reinforces the biaxial strain, causing a reduction in the optical gap and a further ∼122-nm redshift in the emission wavelength. The short-wavelength peak around 1,400 nm observed in the EL spectrum obtained from the InGaAs SRCL bilayer is comprised of a combination of excited-state emission from the QDs (i.e., more evident in this spectrum due to the reduced QD density in this sample) and ground state emission from a population of smaller QDs that are present due to the unoptimized growth conditions. This Section only considers the ground state emission from the main QD population in each sample, so the number of modeled states has not been extended to cover this peak.

Physics of the InGaAs SRCL-Induced Redshifts. In Fig. 53, we compare the local band edges of the lowest conduction band (CB) and two highest valence bands (HH and LH) for the two bilayers (Fig. 51b, c). The InGaAs SRCL reduces the hydrostatic strain and increases the biaxial strain in the upper QD layer (Usman et al. 2011c). The CB band edge therefore shifts to a lower energy while the HH band edge shifts to a higher energy, resulting in a ∼50 meV reduction of the band gap energy, inducing an ∼88 meV change in the optical gap. As a result, the ground state optical emission redshifts by ∼122 nm, and ground state optical emission above 1,300 nm can be achieved.

Hole Energy Levels Aligned Along (110) and (−110). Figure 54 shows a top view of the spatial distribution of the lowest conduction band energy level (E1) and three highest hole energy levels (H1, H2, and H3) for the SQD sample (top row) and the GaAs-capped bilayer (bottom row). E1 possesses a symmetric S-type wave function. However, the hole states are oriented along [110] or [−110], due to strain and piezoelectric field-induced symmetry lowering. Figure 54 shows a plot of the wave function for one state corresponding to the Kramer's doublet. The other degenerate state will have a wave function concentrated at the opposite edge of the QD. As mentioned previously, optical transition rates are calculated by summing the contribution from both degenerate states. It should also be noted that the first three-hole wave functions, plotted in Fig. 54, are concentrated at the interface of the QD rather than at the QD center. This is due to the large aspect ratios (height/base) of these QDs, which results in an increase in biaxial strain at the QD/GaAs matrix interface, creating HH traps (pockets).

Due to the anisotropy of the hole wave functions along [110] and [−110], the interband optical transitions between the electron and hole states, namely, E1↔H1, E1↔H2, and E1↔H3, will be strongly polarization dependent. However, polarization-resolved photoluminescence collected from the surface of equivalent unprocessed samples indicates the emission is isotropic in the plane of the QDs (i.e., TE(110)/TE(−110) ∼ 1.0 +/− 0.1). We demonstrate here that in order to achieve isotropic polarization dependence in the plane of QDs, more than one hole energy level should be included in the calculation of the ground state optical transition strength. This is because hole energy levels are very closely packed, as can be seen in Fig. 52. The difference between the highest and the lowest transition energies calculated here (E1-H3)-(E1-H1) is ∼16 meV, ∼12.5 meV, and ∼14 meV for the SQD, the GaAs-capped, and InGaAs SRCL-capped bilayer, respectively, indicating that the first three-hole energy levels in our QDs are within $\sim 0.5 k_B T$ (∼12.9 meV) at RT. It can therefore be concluded that in our QDs, the topmost valence band states are very closely packed implying that at RT, multiple hole levels can contribute to the measured transition intensity.

DOI 10.1007/978-3-642-27737-5_343-2

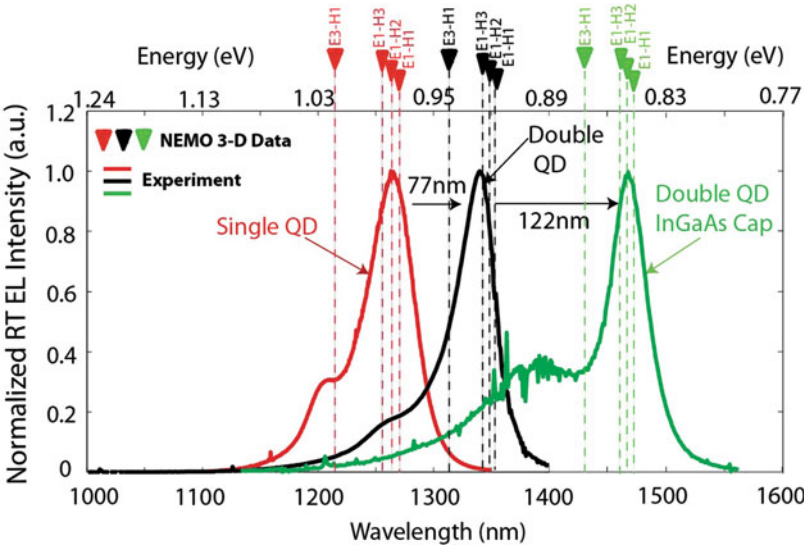


Fig. 52 Comparison of the ground state optical emissions from the single InAs QD and the double InAs QD stack without SRCL and with SRCL. The *red* (single QD), *black* (QD stack without the SRCL), and *green* (QD stack with the SRCL) curves show the room-temperature electroluminescence spectra measured in the experiment. The *vertical dotted lines* are the electron–hole transition energies calculated from the NEMO 3D simulations. A *redshift* of ∼77 nm is observed for the bilayers without the SRCL compared to the single InAs QDs. The InGaAs SRCL further *red shifts* the spectrum by ∼122 nm. The noise in the *green* color close to 1,400 nm wavelength is due to water absorption (Reprinted with permission from (Usman et al. 2011a) copyright 2011, AIP Publishing LLC)

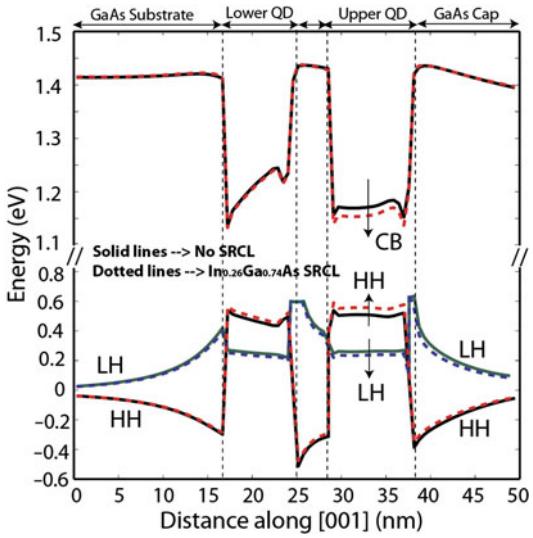


Fig. 53 The comparison of the lowest conduction band edge (CB) and the highest two valence band edges (HH and LH) for the bilayer QD stacks with (*dotted lines*) and without (*solid lines*) the SRCL. The InGaAs SRCL shifts the CB edge to lower energies and the HH band edge to the higher energies, thus reducing the band gap and increasing the optical emission wavelength. The SRCL also shifts the HH and LH band edges in the opposite directions and increases the HH–LH separation. The overall reduction in the band gap due to the SRCL is ∼50 meV that induces a shift of ∼88 meV in the optical gap. The shift of HH band toward higher energies results in an increase in the TE/TM ratio (Reprinted with permission from (Usman et al. 2011a) copyright 2011, AIP Publishing LLC)

DOI 10.1007/978-3-642-27737-5_343-2

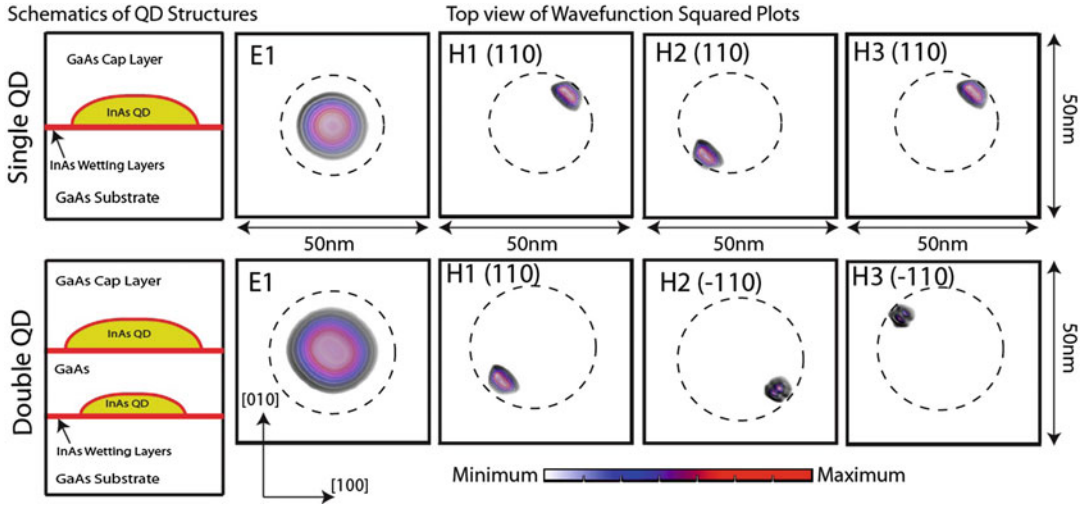


Fig. 54 *Top view* of the wave function plots of the lowest electron energy level (E1) and the highest three-hole energy levels (H1, H2, and H3) for the single QD (*upper row*) and the double QD stack without the SRCL (*lower row*). The color intensity of the plots indicates the magnitude of the wave function data: the *red* color corresponding to the maximum value and the *blue* color corresponding to the minimum value. *Dashed circles* are drawn to highlight the boundaries of the QD regions. Leftmost column show the schematic diagram of the system. The lowest electron energy level is of S-type character and shows a symmetric distribution of the charge density. The hole energy levels are strongly affected by the strain and piezoelectricity and tend to align along the [110] or [−110] direction (Reprinted with permission from (Usman et al. 2011a) copyright 2011, AIP Publishing LLC)

Multiple Holes Are Required to Achieve Isotropic In-Plane Polarization. Figure 55 shows the optical intensity computed from the NEMO 3D simulations as a polar plot for (a) the SQD and (b) GaAs-capped bilayer. The direction of the incident light is along [001]. The optical transition intensities are plotted as a function of the angle, θ, measured between [100] and [010]. In Fig. 55a, we observe a slightly anisotropic polarization dependence (i.e., TE(110)/TE(−110) ∼1.18) as the first three-hole energy levels are oriented along [110] (top row of Fig. 54), resulting in a slightly weaker TE(−110) mode. Figure 55b shows nearly isotropic polarization emission for the GaAs-capped bilayer. This is because H1 is oriented along [110], while H2 and H3 are both oriented along [−110]. This orthogonal distribution of hole wave functions tends to cancel out the in-plane polarization sensitivity, and hence the resulting optical spectrum becomes nearly polarization insensitive (i.e., TE(110)/TE(−110) ∼1.07). We conclude that any theoretical study of the polarization-resolved ground state optical emission should include multiple hole energy levels in order to achieve polarization insensitivity in the plane of the QDs. Previous studies (Sheng and Leburton 2002) only considered the topmost valence band state.

TE/TM Ratio Analysis. For telecommunication applications, a polarization-insensitive response is desirable for some edge-emitting devices. Self-assembled QDs have low aspect ratios, typically 0.1–0.3, and strong carrier confinement along [001], resulting in very anisotropic optical properties, where the TE optical mode is dominant and the TM optical mode can be very weak. Large aspect ratio (>0.6) columnar QDs and multilayer QD stacks have been designed to achieve isotropic in-plane polarization. Here, we calculate the polarization-dependent optical transitions for our SQD sample and compare it to the response of our GaAs-capped and InGaAs SRCL-capped bilayers shown in Fig. 51c–e. We compare these results to the experimentally measured TE/TM ratios obtained from cleaved-edge PV spectral measurements.

SQD to GaAs-Capped Bilayer: TE/TM Ratio Decreases. Figure 56 compares the optical intensities computed from the NEMO 3D simulator and represented as a polar plot for the SQD, the GaAs-capped

DOI 10.1007/978-3-642-27737-5_343-2

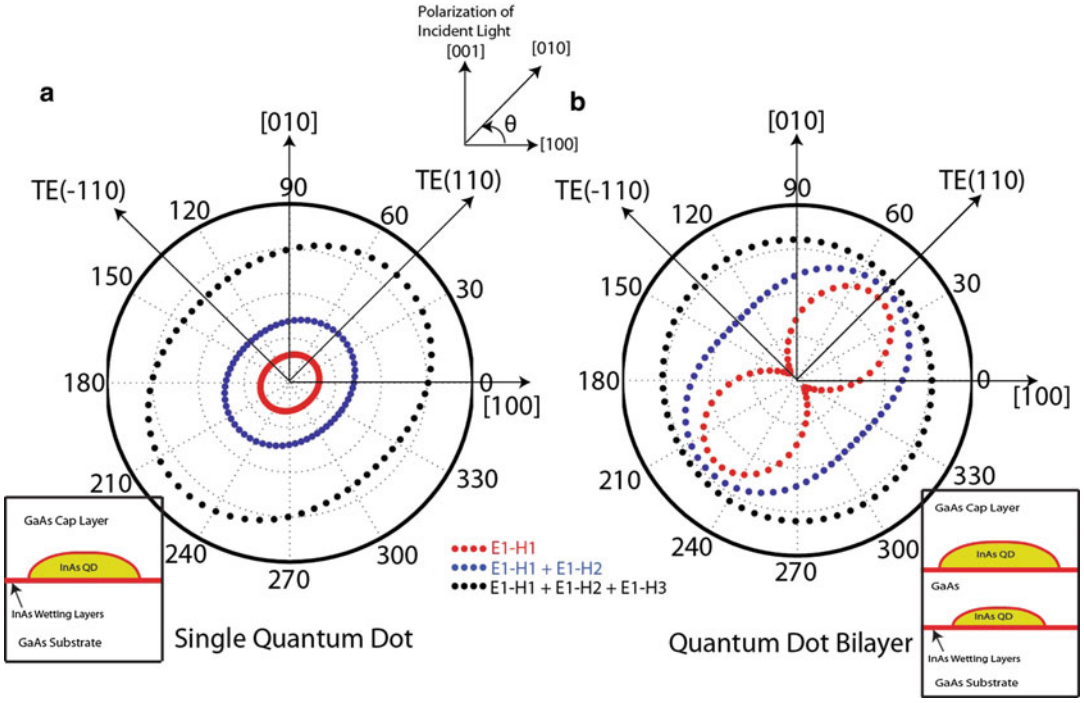


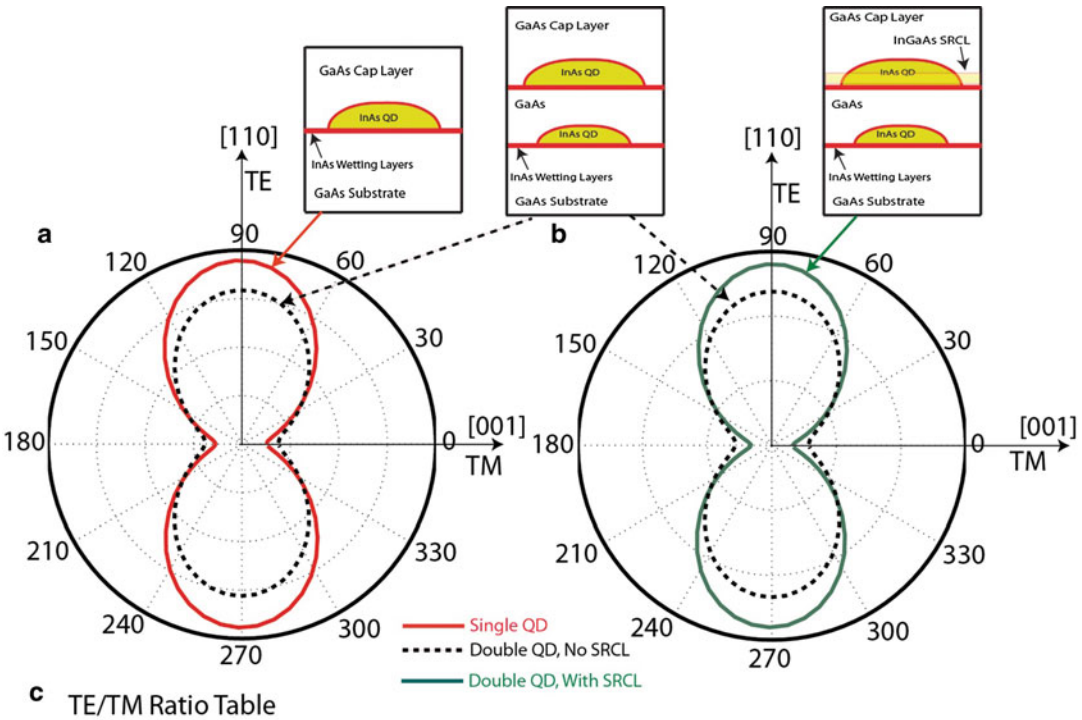
Fig. 55 Optical intensity model results represented as polar plots for (**a**) the single InAs QD and (**b**) the bilayer InAs QD stack without the SRCL. The schematics of the quantum dot geometry are inserted as insets. The direction of the polarization of the incident light is assumed along the [001] direction. The optical transition intensities are plotted as a function of the angle θ between the [100] and the [010] directions. *Red dots*: only E1↔H1 transition is plotted. *Blue dots*: the sum of E1↔H1 and E1↔H2 transitions is plotted. *Black dots*: the sum of E1↔H1, E1↔H2 and E1↔H3 transitions is plotted (Reprinted with permission from (Usman et al. 2011a) copyright 2011, AIP Publishing LLC)

bilayer, and the InGaAs SRCL-capped bilayer. The direction of the incident light is along [−110]. Interband optical transition intensities are calculated as a function of the angle θ between [001] and [110], and each curve represents the sum of the optical intensities for E1↔H1, E1↔H2, and E1↔H3. In Fig. 56a, we see that the TE/TM ratio decreases in going from an SQD to a GaAs-capped bilayer. As in a bilayer, the strain of the seed QD influences the growth of the upper QD, resulting in a larger upper QD. The taller upper QD reduces the [001] carrier confinement, and hence, the TM optical mode is enhanced, leading to a reduction in the TE/TM ratio. It should be noted that although TE/TM ratio decreases for the bilayer QD system as compared to the SQD, it is still considerably higher than the desired isotropic value, i.e., TE/TM = 1.0.

GaAs-Capped Bilayer to InGaAs SRCL-Capped Bilayer: TE/TM Ratio Increases. Figure 56b compares the polar plots of the bilayer with and without an InGaAs SRCL. An increase in the polarization sensitivity occurs when the upper QD layer is covered by an InGaAs SRCL. The reason for this increase can be understood by considering the valence band-edge diagrams in Fig. 3, which show that due to the biaxial strain reinforcement, an InGaAs SRCL shifts the HH and LH band edges in opposite directions, leading to an increase in the splitting between the two valence bands. As a result, the topmost valence band states will have enhanced HH character resulting in the calculated and measured increase in the TE optical mode leading to an increase in the TE/TM ratio.

The results for our InGaAs SRCL-capped bilayer are in contrast to a recent experimental study (Jayavel et al. 2004) where an InGaAs SRCL is shown to decrease the TE/TM ratio for a single InAs QD. That result was attributed to the significant increase in the QD aspect ratio, from 0.235 to 0.65, due to the influence of

DOI 10.1007/978-3-642-27737-5_343-2



c TE/TM Ratio Table

	Experiment	NEMO 3-D
Single QD No SRCL	4.0 +/- 0.1	~7.3
Double QD No SRCL	3.5 +/- 0.1	~4.06
Double QD With SRCL	5.2 +/- 0.1	~8.7

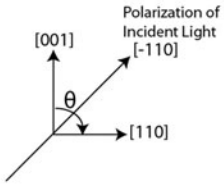


Fig. 56 (**a**, **b**) Optical intensity model results represented as polar plots are shown for the single InAs QD (*red solid curve*), the bilayer InAs QD stack without SRCL (*black dotted curve*), and the bilayer InAs QD stack with SRCL (*green solid line*). The schematics of the quantum dot geometry are inserted in as small insets. The direction of the polarization of the incident light is assumed to be along the [−110] direction. The interband optical transition intensities are calculated as a function of the angle θ between the [110] and the [001] directions. Each curve represents the sum of the intensities of the E1↔H1, E1↔H2, and E1↔H3 transitions. (**c**) The comparison of the TE(110)/TM(001) ratios for the three QD systems from experimental PV measurements and NEMO 3D calculations (Reprinted with permission from (Usman et al. 2011a) copyright 2011, AIP Publishing LLC)

the InGaAs cap in preserving the QD height during the capping process. For our bilayers, the relatively low capping temperature for QDs in the second layer will also preserve the QD height, and we have not observed a significant enhancement in the QD aspect ratio due to the presence of an InGaAs capping layer. The TEM images in Fig. 51a, b clearly indicate that an InGaAs SRCL does not drastically change the aspect ratio of our QDs, in contrast to the InGaAs-capped SQDs reported by Jayavel et al. (2004).

The table in Fig. 56c summarizes the TE/TM ratios from the experimental PV measurements and the NEMO 3D-based calculations for the three QD systems. The calculated values are in reasonable agreement with the experimental measurements for all three cases and show similar trends.

Encyclopedia of Complexity and Systems Science
DOI 10.1007/978-3-642-27737-5_343-2

Summary and Future Directions

NEMO 3D is introduced 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 3D 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 3D is not limited to research computing alone; a first educational version including visualization capabilities has been released on http://www.nanoHUB.org and has been used by hundreds of users for thousands of simulations. The full version of NEMO 3D will soon be available for device engineers, material scientists, educators, and students through the nanoHUB, powered by the NSF TeraGrid. Tool documentation, tutorials, and case studies will be posted on nanoHUB as supplemental material. We will generate and deliver tutorials on parallelization and software development through the nanoHUB.

NEMO 3D demonstrates the capability to model a large variety of relevant, realistically sized nanoelectronic devices. The impact of atomistic strain and piezoelectricity on the electronic structure in dome-shaped InAs/GaAs quantum dots is explored. Under the assumptions of realistic boundary conditions, strain is found to be long-ranged and penetrate around 25 nm into the dot substrate thus stressing the need for using large dimensions of these surrounding layers and at least three million atoms in the simulations. The true symmetry of the quantum dots is found to be lower than the geometric shape symmetry because of the fundamental atomistic nature of the underlying zinc blende crystal lattice. Atomistic strain is found to induce further optical polarization anisotropy favoring the [110] direction and pronounced non-degeneracy in the quantum dot excited states, magnitude (few meV) of which depends mainly on the dot size and surrounding material matrix. First-order piezoelectric potential, on the other hand, favors the [1$\underline{1}$0] direction, reduces the non-degeneracy in the *P* states, and is found to be strong enough to *flip* the optical polarization in certain-sized quantum dots (Bester and Zunger 2005). Regarding the second-order piezoelectricity, in contrast to some recent small-scale numerical experiments, our calculations yield a nonvanishing and reasonably large *net* piezoelectric potential, which can be attributed to the fact that the potential from the linear term, as compared to the quadratic counterpart, penetrates *deeper* into the surrounding material matrix. Simulations of QD stacks exemplify the complicated mechanical strain and quantum mechanical interactions on confined electronic states. Molecular states can be observed when the dots are in close proximity. Also, as for the two closely stacked layers (bilayers) of InAs/GaAs QDs, in contrast to a recent study of single InAs QD layers, our experimental and theoretical results indicate an increase in the TE/TM ratio when the upper QD layer in the bilayer is covered by an InGaAs strain-reducing capping layer (SRCL). Simulations of SiGe-buffered Si QWs indicate the importance of band-to-band interactions that are naturally understood in the NEMO 3D basis. Valley splitting is computed as a function of magnetic field matching experimental data. Simulations of disordered SiGe-alloyed nanowires indicate the critical importance of the treatment of atomistic disorder. Typical approaches of a smoothed out material (VCA) or considerations of bandstructure in just individual slices clearly fail to represent the disordered nanowire physics. A semiempirical tight-binding model for Group V donors in silicon is presented. The dependence of valley–orbit interaction on on-site cutoff potential and orbital energies is explored. A block-based Lanczos algorithm was demonstrated as a robust and reliable method of finding eigenvalues and vectors of the resulting system. The technique outlined here enables high-precision modeling of impurity-based quantum electronics with relative ease and accuracy. Simulations on graphene demonstrate that expanding the basis of the π-bands to include *d* orbitals (the *p/d* model) allows for reproduction of the ab initio bands, which is impossible with the commonly employed single p_z model.

DOI 10.1007/978-3-642-27737-5_343-2

All these NEMO 3D calculations underline the importance to represent explicitly the atomistically resolved physical system with a physics-based local orbital representation. Such million atom systems result in system sizes of tens of millions, and end-to-end 52 million atom simulations representing one *billion* degree of freedom systems were presented. The complexity of the system demands the use of well qualified, tuned, optimized algorithms and modern HPC platforms. Building and maintaining such a code is not a light undertaking and requires a significant group community effort.

Integrated circuit design faces a crisis – the 40-year process of transistor downscaling has led to atomic-scale features, making devices subject to unavoidable manufacturing irregularities at the atomic scale and to heat densities comparable to a nuclear reactor. *A new approach to design that embraces the atomistic, quantum mechanical nature of the constituent materials is necessary to develop more powerful yet energy miserly devices*. We are in the process of developing a general-purpose simulation engine. It will model not only the electronic bandstructure but also the out-of-equilibrium electron *transport* in realistically extended devices using fully quantum mechanical (QM) models in an atomistic material description containing millions of atoms. The research will enable discovery of new technologies for faster switching, smaller feature size, and reduced heat generation. Using this new approach, designers can directly address questions of quantization and spin, tunneling, phonon interactions, and heat generation. It is widely accepted that the nonequilibrium Green function formalism (NEGF) QM statistical mechanics theory, in conjunction with an atomistic basis, can answer these questions. It is also widely perceived that the problem is computationally hard to solve. A generalized approach to trilevel parallelism in voltage, energy, and space is highly desired. Another task addresses the bottleneck of calculating open boundary conditions (BCs) for large cross sections for realistically large structures. The BCs can be reused for each voltage point and each charge self-consistent iteration. With a view to achieving these goals, the necessary levels of parallelism to tackle the problem on 200,000+ CPUs have been designed and demonstrated to scale well. Computer scientists and HPC experts embedded in the team will guide the implementation and explore performance, execution reliability, and alternative hardware and algorithms. The new simulation code named OMEN (with nonequilibrium Green function and 3D atomistic representation) will be an open-source project and disseminated through the nanoHUB.

Acknowledgment

The work has been supported by the Indiana 21st Century Fund, Army Research Office, Office of Naval Research, Semiconductor Research Corporation, ARDA, and the National Science Foundation. The work described in this publication was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The development of the NEMO 3D tool involved a large number of individuals at JPL and Purdue, whose work has been cited. Drs. R. Chris Bowen, Fabiano Oyafuso, and Seungwon Lee were key contributors in this large effort at JPL. The authors acknowledge an NSF TeraGrid award DMR070032. Shaikh Ahmed acknowledges NSF Grant Nos. CCF-1218839 and EECS-1102192. Access to the Blue Gene was made available through the auspices of the Computational Center for Nanotechnology Innovations (CCNI) at Rensselaer Polytechnic Institute. Access to the Oak Ridge National Lab XT3/4 was provided by the National Center for Computational Sciences project. We would also like to thank the Rosen Center for Advanced Computing at Purdue for their support. NanoHUB computational resources were used for part of this work.

DOI 10.1007/978-3-642-27737-5_343-2

Bibliography

Primary Literature

Agnello PD (2002) Process requirements for continued scaling of CMOS – the need and prospects for atomic-level manipulation. IBM J Res Dev 46:317–338

Ahmed S, Usman M, Heitzinger C, Rahman R, Schliwa A, Klimeck G (2007) Atomistic simulation of non-degeneracy and optical polarization anisotropy in zincblende quantum dots. In: The 2nd annual IEEE international conference on nano/micro engineered and molecular systems (IEEE-NEMS), Bangkok

Ahmed S, Islam S, Mohammed S (2010) Electronic structure of InN/GaN quantum dots: multimillion atom tight-binding simulations. IEEE Trans Electron Devices 57:164–173

Arakawa Y, Sasaki H (1982) Multidimensional quantum well laser and temperature dependence of its threshold current. Appl Phys Lett 40:939

Bae H, Clark S, Haley B, Klimeck G, Korkusinski M, Lee S, Naumov M, Ryu H, Saied F (2007) Electronic structure computations of quantum dots with a billion degrees of freedom. Supercomputing 07, Reno

Bester G, Zunger A (2005) Cylindrically shaped zinc-blende semiconductor quantum dots do not have cylindrical symmetry: atomistic symmetry, atomic relaxation, and piezoelectric effects. Phys Rev B 71:045318. Also see references therein

Bester G, Wu X, Vanderbilt D, Zunger A (2006a) Importance of second-order piezoelectric effects in zinc-blende semiconductors. Phys Rev Lett 96:187602

Bester G, Zunger A, Wu X, Vanderbilt D (2006b) Effects of linear and nonlinear piezoelectricity on the electronic properties of InAs/GaAs quantum dots. Phys Rev B 74:081305

Bowen R, Klimeck G, Lake R, Frensley W, Moise T (1997) Quantitative resonant tunneling diode simulation. J Appl Phys 81:207

Boykin T, Klimeck G (2005) Practical application of zone-folding concepts in tight-binding. Phys Rev B 71:115215

Boykin T, Vogl P (2001) Dielectric response of molecules in empirical tight-binding theory. Phys Rev B 65:035202

Boykin T, Bowen R, Klimeck G (2001) Electromagnetic coupling and gauge invariance in the empirical tight-binding method. Phys Rev B 63:245314

Boykin T, Klimeck G, Bowen R, Oyafuso F (2002) Diagonal parameter shifts due to nearest-neighbor displacements in empirical tight-binding theory. Phys Rev B 66:125207

Boykin T, Klimeck G, Oyafuso F (2004a) Valence band effective mass expressions in the sp^3d5s^* empirical tight-binding model applied to a new Si and Ge parameterization. Phys Rev B 69:115201

Boykin T, Klimeck G, Eriksson M, Friesen M, Coppersmith S, Allmen P, Oyafuso F, Lee S (2004b) Valley splitting in strained Si quantum wells. Appl Phys Lett 84:115

Boykin T, Klimeck G, Eriksson M, Friesen M, Coppersmith S, Allmen P, Oyafuso F, Lee S (2004c) Valley splitting in low-density quantum-confined heterostructures studied using tight-binding models. Phys Rev B 70:165325

Boykin T, Klimeck G, Allmen P, Lee S, Oyafuso F (2005) Valley-splitting in V-shaped quantum wells. J Appl Phys 97:113702

Boykin T, Luisier M, Schenk A, Kharche N, Klimeck G (2007a) The electronic structure and transmission characteristics of disordered AlGaAs nanowires. IEEE Trans Nanotechnol 6:43

Boykin T, Kharche N, Klimeck G (2007b) Brillouin zone unfolding of perfect supercells composed of non-equivalent primitive cells. Phys Rev B 76:035310

Boykin T, Kharche N, Klimeck G, Korkusinski M (2007c) Approximate bandstructures of semiconductor alloys from tight-binding supercell calculations. J Phys Condens Matter 19:036203

Boykin TB, Luisier M, Klimeck G, Jiang X, Kharche N, Zhou Y, Nayak SK (2011) Accurate six-band nearest-neighbor tight-binding model for the π-bands of bulk graphene and graphene nanoribbons. J Appl Phys 109:10434

Bradbury F et al (2006) Stark tuning of donor electron spins in silicon. Phys Rev Lett 97:176404

Calderón MJ, Koiler B, Hu X, Das Sarma S (2006) Quantum control of donor electrons at the Si-SiO_2 interface. Phys Rev Lett 96:096802

Canning A, Wang LW, Williamson A, Zunger A (2000) Parallel empirical pseudopotential electronic structure calculations for million atom systems. J Comp Phys 160:29

Castro Neto AH, Guinea F, Peres NMR, Novoselov KS, Gein AK (2009) The electronic properties of graphene. Rev Mod Phys 81:109

Chen P, Piermarocchi C, Sham L (2001) Control of exciton dynamics in nanodots for quantum operations. Phys Rev Lett 87:067401

Chen J-H, Jang C, Xiao S, Ishigami M, Fuhrer MS (2008) Intrinsic and extrinsic performance limits of graphene devices on SiO_2. Nat Nanotechnol 3:206

Colinge JP (2004) Multipole-gate SOI MOSFETs. Solid State Electron 48:897–905

Cui Y, Lauhon L, Gudiksen M, Wang J, Lieber C (2001) Diameter-controlled synthesis of single-crystal silicon nanowire. Appl Phys Lett 78:2214

Debernardi A et al (2006) Computation of the Stark effect in P impurity states in silicon. Phys Rev B 74:035202

Dobers M, Klitzing K, Schneider J, Weimann G, Ploog K (1998) Electrical detection of nuclear magnetic resonance in GaAs-$Al_xGa_{1-x}As$ heterostructures. Phys Rev Lett 61:1650

Eriksson M, Friesen M, Coppersmith S, Joynt R, Klein L, Slinker K, Tahan C, Mooney P, Chu J, Koester S (2004) Spin-based quantum dot quantum computing in Silicon. Quantum Inf Process 3:133

Fafard S, Hinzer K, Raymond S, Dion M, Mccaffrey J, Feng Y, Charbonneau S (1996) Red-emitting semiconductor quantum dot lasers. Science 22:1350

Friesen M, Rugheimer P, Savage D, Lagally M, van der Weide D, Joynt R, Eriksson M (2003) Practical design and simulation of silicon-based quantum-dot qubits. Phys Rev B 67:121301

Friesen M et al (2005) Theory of the Stark effect for P donors in Si. Phys Rev Lett 94:186403

Gonze X et al (2005) A brief introduction to the ABINIT software package. Z Kristallogr 220:558

Gonze X et al (2009) ABINIT: first-principles approach to material and nanosystem properties. Comput Phys Commun 180:2582

Goswami S, Slinker KA, Friesen M, McGuire LM, Truitt JL, Tahan C, Klein LJ, Chu JO, Mooney PM, van der Weide DW, Joynt R, Coppersmith SN, Eriksson MA (2007) Controllable valley splitting in silicon quantum devices. Nat Phys 3:41

Graf M, Vogl P (1995) Electromagnetic fields and dielectric response in empirical tight-binding theory. Phys Rev B 51:4940

Greentree A, Cole J, Hamilton A, Hollenberg L (2004) Coherent electronic transfer in quantum dot systems using adiabatic passage. Phys Rev B 70:235317

Greytak A, Lauhon L, Gudiksen M, Lieber C (2004) Growth and transport properties of complementary germanium nanowire field-effect transistors. Appl Phys Lett 84:4176

Grundmann M, Stier O, Bimberg D (1995) InAs/GaAs pyramidal quantum dots: strain distribution, optical phonons, and electronic structure. Phys Rev B 52:11969

Hollenberg L et al (2004) Charge-based quantum computing using single donors in semiconductors. Phys Rev B 69:113301

Hollenberg L et al (2006) Two-dimensional architectures for donor-based quantum computing. Phys Rev B 74:045311

Hu X et al (2005) Charge qubits in semiconductor quantum computer architecture: tunnel coupling and decoherence. Phys Rev B 71:235332

Hybertsen MS, Louie SG (1986) Electron correlation in semiconductors and insulators: band gaps and quasiparticle energies. Phys Rev B 34:5390

Jancu J, Scholz R, Beltram F, Bassani F (1998) Empirical spds* tight-binding calculation for cubic semiconductors: general method and material parameters. Phys Rev B 57:6493

Jarjour A, Taylor R, Oliver R, Kappers M, Humphreys C, Tahraoui A (2008) Electrically driven single InGaN/GaN quantum dot emission. Appl Phys Lett 93:233103

Jayavel P, Tanaka H, Kita T, Wada O, Ebe H, Sugawara M, Tatebayashi J, Arakawa Y, Nakata Y, Akiyama T (2004) Control of optical polarization anisotropy in edge emitting luminescence of InAs/GaAs self-assembled quantum dots. Appl Phys Lett 84:1820

Kalden J, Tessarek C, Sebald K, Figge S, Kruse C, Hommel D, Gutowski J (2010) Electroluminescence from a single InGaN quantum dot in the green spectral region up to 150 K. Nanotechnology 21:015204

Kane B (1998) A silicon-based nuclear spin quantum computer. Nature 393:133

Ke W, Fu C, Chen C, Lee L, Ku C, Chou W, Chang W-H, Lee M, Chen W, Lin W (2006) Photoluminescence properties of self-assembled InN dots embedded in GaN grown by metal organic vapor phase epitaxy. Appl Phys Lett 88:191913

Keating P (1966) Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure. Phys Rev 145:637

Kharche N, Prada M, Boykin T, Klimeck G (2007) Valley-splitting in strained Silicon quantum wells modeled with 2 degree miscuts, step disorder, and alloy disorder. Appl Phys Lett 90:092109

Kharche N, Luisier M, Boykin T, Klimeck G (2008) Electronic structure and transmission characteristics of SiGe nanowire. J Comput Electron 7:350

Kim C, Yang J, Lee H, Jang H, Joa M, Park W, Kim Z, Maeng S (2007) Fabrication of $Si_{1-x}Ge_x$ alloy nanowire field-effect transistors. Appl Phys Lett 91:033104

Klimeck G, Boykin T, Chris R, Lake R, Blanks D, Moise T, Kao Y, Frensley W (1997) Quantitative simulation of strained InP-based resonant tunneling diodes. In: Proceedings of the 1997 55th IEEE device research conference digest, Colorado vol 92

Klimeck G, Bowen R, Boykin T, Cwik T (2000a) sp^3s^* tight-binding parameters for transport simulations in compound semiconductors. Superlattices Microstruct 27:519–524

Klimeck G, Bowen R, Boykin T, Salazar-Lazaro C, Cwik T, Stoica A (2000b) Si tight-binding parameters from genetic algorithm fitting. Superlattices Microstruct 27:77–88

Klimeck G, Oyafuso F, Boykin T, Bowen R, Allman P (2002) 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 3:601

Klimeck G, Boykin T, Luisier M, Kharche N, Schenk A (2006) A Study of alloyed nanowires from two perspectives: approximate dispersion diagrams and transmission coefficients. In: Proceedings of the 28th international conference on the physics of semiconductors ICPS 2006, Vienna

Klimeck G, Ahmed S, Bae H, Kharche N, Clark S, Haley B, Lee S, Naumov M, Ryu H, Saied F, Prada M, Korkusinski M, Boykin T (2007a) Atomistic simulation of realistically sized nanodevices using NEMO 3-D: part I – models and benchmarks. IEEE Trans Electron Devices 54:2079

Klimeck G, Ahmed S, Kharche N, Korkusinski M, Usman M, Prada M, Boykin T (2007b) Atomistic simulation of realistically sized nanodevices using NEMO 3-D: part II – applications. IEEE Trans Electron Devices 54:2090

Klimeck G, Mannino M, McLennan M, Qiao W, Wang X (2008) https://www.nanohub.org/simulation_tools/qdot_tool_information
Kohn W, Luttinger J (1995) Theory of donor states in silicon. Phys Rev 98:915
Koiller B, Hu X, Das Sarma S (2006) Electric-field driven donor-based charge qubits in semiconductors. Phys Rev B 73:045319
Korkusinski M, Klimeck G (2006) Atomistic simulations of long-range strain and spatial asymmetry molecular states of seven quantum dots. J Phys Conf Ser 38:75–78
Korkusinski M, Klimeck G, Xu H, Lee S, Goasguen S, Saied F (2005) Atomistic simulations in nanostructures composed of tens of millions of atoms: importance of long-range strain effects in quantum dots. In: Proceedings of 2005 NSTI conference, Anaheim
Korkusinski M, Saied F, Xu H, Lee S, Sayeed M, Goasguen S, Klimeck G (2005) Large scale simulations in nanostructures with NEMO3-D on Linux clusters. Linux Cluster Institute conference, Raleigh
Lanczos C (1950) An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. J Res Natl Bur Stand 45:255–282
Lansbergen GP, Rahman R, Wellard CJ, Woo I, Caro J, Collaert N, Biesemans S, Klimeck G, Hollenberg LCL, Rogge S (2008) Gate induced quantum confinement transition of a single dopant atom in a Si FinFET. Nat Phys 4:656
Lazarenkova O, Allmen P, Oyafuso F, Lee S, Klimeck G (2004) Effect of anharmonicity of the strain energy on band offsets in semiconductor nanostructures. Appl Phys Lett 85:4193
Lee S, Kim J, Jönsson L, Wilkins J, Bryant G, Klimeck G (2002) Many-body levels of multiply charged and laser-excited InAs nanocrystals modeled by empirical tight binding. Phys Rev B 66:235307
Lee S, Lazarenkova O, Oyafuso F, Allmen P, Klimeck G (2004a) Effect of wetting layers on the strain and electronic structure of InAs self-assembled quantum dots. Phys Rev B 70:125307
Lee S, Oyafuso F, Allmen P, Klimeck G (2004b) Boundary conditions for the electronic structure of finite-extent, embedded semiconductor nanostructures with empirical tight-binding model. Phys Rev B 69:045316
Li X et al (2008) Chemically derived, ultrasmooth graphene nanoribbon semiconductors. Science 319:1229
Liang G, Xiang J, Kharche N, Klimeck G, Lieber C, Lundstrom M (2006) Performance analysis of a Ge/Si core/shell nanowire field effect transistor. cond-mat 0611226
Loss D, DiVincenzo DP (1998) Quantum computation with quantum dots. Phys Rev A 57:120
Luisier M, Schenk A, Fichtner W, Klimeck G (2006) Atomistic simulation of nanowires in the $sp^3d^5s^*$ tight-binding formalism: from boundary conditions to strain calculations. Phys Rev B 74:205323
Martins A et al (2004) Electric-field control and adiabatic evolution of shallow donor impurities in silicon. Phys Rev B 69:085320
Maschhoff K, Sorensen D (1996) A portable implementation of ARPACK for distributed memory parallel architectures. Copper Mountain conference on iterative methods, Copper Mountain, 9–13 Apr 1996
Maximov M, Shernyakov Y, Tsatsul'nikov A, Lunev A, Sakharov A, Ustinov V, Egorov A, Zhukov A, Kovsch A, Kop'ev P, Asryan L, Alferov Z, Ledentsov N, Bimberg D, Kosogov A, Werner P (1998) High-power continuous-wave operation of a InGaAs/AlGaAs quantum dot laser. J Appl Phys 83:5561
Michler P, Kiraz A, Becher C, Schoenfeld W, Petroff P, Zhang L, Hu E, Imamoglu A (2000) A quantum dot single-photon turnstile device. Science 290:2282–2285
Moore G (1975) Progress in digital integrated electronics. IEDM Tech Dig (21):11–13
Moreau E, Robert I, Manin L, Thierry-Mieg V, Gérard J, Abram I (2001) Quantum cascade of photons in semiconductor quantum dots. Phys Rev Lett 87:183601
Morkoç H, Mohammad SN (1995) High-luminosity blue and blue-green gallium nitride light-emitting diodes. Science 267:51

Nakada K, Fujita M, Dresselhaus G, Dresselhaus MS (1996) Edge state in graphene ribbons: nanometer size effect and edge shape dependence. Phys Rev B 54:17954

Naumov M, Lee S, Haley B, Bae H, Clark S, Rahman R, Ryu H, Saied F, Klimeck G (2007) Eigenvalue solvers for atomistic simulations of electronic structureswith NEMO-3D. In: 12th international workshop on computational electronics, Amherst, 7–10 Oct

Novoselov KS et al (2004) Electric field effect in atomically thin carbon films. Science 306:666

Novoselov KS et al (2005a) Two-dimensional gas of massless Dirac fermions in graphene. Nature 438:197

Novoselov KS et al (2005b) Two-dimensional atomic crystals. Proc Natl Acad Sci U S A 102:10451

Oberhuber R, Zandler G, Vogl P (1998) Subband structure and mobility of two–dimensional holes in strained Si/SiGe MOSFET's. Phys Rev B 58:9941–9948

Oyafuso F, Klimeck G, Allmen P, Boykin T, Bowen R (2003a) Strain effects in large-scale atomistic quantum dot simulations. Phys Status Solidi B 239:71

Oyafuso F, Klimeck G, Bowen R, Boykin T, Allmen P (2003b) Disorder induced broadening in multimillion atom alloyed quantum dot systems. Phys Status Solidi C 4:1149

Persson A, Larsson M, Steinström S, Ohlsson B, Samuelson L, Wallenberg L (2004) Solid phase diffusion mechanism for GaAs NW growth. Nat Mater 3:677

Petroff P (2003) Single quantum dots: fundamentals, applications, and new concepts. Springer, Berlin

Ponce FA, Bour DP (1997) Nitride-based semiconductors for blue and green light-emitting devices. Nature 386:351

Prada M, Kharche N, Klimeck G (2007) Electronic structure of Si/InAs composite channels. In: MRS spring conference, symposium G: extending Moore's law with advanced channel materials, San Francisco, 9–13 Apr 2007

Pryor C, Kim J, Wang L, Williamson A, Zunger A (1998) Comparison of two methods for describing the strain profiles in quantum dots. J Appl Phys 83:2548

Qiao W, Mclennan M, Kennell R, Ebert D, Klimeck G (2006) Hub-based simulation and graphics hardware accelerated visualization for nanotechnology applications. IEEE Trans Vis Comput Graph 12:1061–1068

Rahman A, Klimeck G, Lundstrom M (2005) Novel channel materials for ballistic nanoscale MOSFETs bandstructure effects. In: 2005 I.E. international electron devices meeting, Washington, DC, pp 601–604

Rahman R et al (2007) High precision quantum control of single donor spins in silicon. Phys Rev Lett 99:036403

Ramdas A et al (1981) Spectroscopy of the solid-state analogues of the hydrogen atom: donors and acceptors in semiconductors. Rep Prog Phys 44:1297–1387

Reed M (1993) Quantum dots. Sci Am 268:118

Reed M, Randall J, Aggarwal R, Matyi R, Moore T, Wetsel A (1988) Observation of discrete electronic states in a zero-dimensional semiconductor nanostructure. Phys Rev Lett 60:535

Ru ECL, Howe P, Jones TS, Murray R (2003) Strain-engineered InAs/GaAs quantum dots for long-wavelength emission. Phys Rev B 67:165303

Sameh A, Tong Z (2000) The trace minimization method for the symmetric generalized eigenvalue problem. J Comput Appl Math 123:155–175

Sellier H et al (2006) Transport spectroscopy of a single dopant in a gated silicon nanowire. Phys Rev Lett 97:206805

Semiconductor Industry Association (2001) International technology roadmap for semiconductors. http://public.itrs.net/Files/2001ITRS/Home.htm

Sheng W, Leburton JP (2002) Interband transition distributions in the optical spectra of InAs/GaAs self-assembled quantum dots. Appl Phys Lett 80:2755

Slater J, Koster G (1954) Simplified LCAO method for the periodic potential problem. Phys Rev 94:1498–1524

Stegner A et al (2006) Electrical detection of coherent P spin quantum states. Nat Phys 2:835

Stier O, Grundmann M, Bimberg D (1999) Electronic and optical properties of strained quantum dots modeled by 8-band *k.p* theory. Phys Rev B 59:5688

Sundaresan S, Islam S, Ahmed S (2010) Built-In Electric Fields in InAs/GaAs Quantum Dots: Geometry Dependence and Effects on the Electronic Structure. In: Technical Proceedings of IEEE Nanotechnology Materials and Devices Conferences (NMDC), California, USA pp. 30–35

Sze S, May G (2003) Fundamentals of semiconductor fabrication. Wiley, New York

Usman M, Ahmed S, Korkusinski M, Heitzinger C, Klimeck G (2006) Strain and electronic structure interactions in realistically scaled quantum dot stacks. In: Proceedings of the 28th international conference on the physics of semiconductors ICPS 2006, Vienna

Usman M, Ryu H, Woo I, Ebert DS, Klimeck G (2009) Moving toward nano-TCAD through multimillion-atom quantum-dot simulations matching experimental data. IEEE Trans Nanotechnol 8:330

Usman M, Heck S, Clarke E, Spencer P, Ryu H, Murray R, Klimeck G (2011a) Experimental and theoretical study of polarization-dependent optical transitions in InAs quantum dots at telecommunication-wavelengths (1300–1500 nm). J Appl Phys 109:104510

Usman M, Inoue T, Harda Y, Klimeck G, Kita T (2011b) Experimental and atomistic theoretical study of degree of polarization from multilayer InAs/GaAs quantum dot stacks. Phys Rev B 84:115321

Usman M, Tan M, Ryu R, Krenner H, Ahmed S, Boykin T, Klimeck G (2011c) Quantitative excited state spectroscopy of a single InGaAs quantum dot molecule through multi-million-atom electronic structure calculations. Nanotechnology 22:315709

Vasileska D, Khan H, Ahmed S (2005) Quantum and coulomb effects in nanodevices. Int J Nanosci 4:305–361

Vrijen R et al (2000) Electron-spin-resonance transistors for quantum computing in silicon-germanium heterostructures. Phys Rev A 62:012306

Wallace PR (1947) The band theory of graphite. Phys Rev 71:622

Waltereit P, Brandt O, Trampert A, Grahn HT, Menniger J, Ramsteiner M, Reiche M, Ploog KH (2000) Nitride semiconductors free of electrostatic fields for efficient white light-emitting diodes. Nature 406:865–868

Wang L, Zunger A (1994) Solving Schrödinger's equation around a desired energy: application to silicon quantum dots. J Chem Phys 100:2394

Wang J, Rahman A, Ghosh A, Klimeck G, Lundstrom M (2005) Performance evaluation of ballistic silicon nanowire transistors with atomic-basis dispersion relations. Appl Phys Lett 86:093113

Wang X, Ouyang Y, Li X, Wang H, Guo J, Dai H (2008) Room-temperature all-semiconducting sub-10-nm graphene nanoribbon field-effect transistors. Phys Rev Lett 100:206803

Wang H, Jiang D, Zhu J, Zhao D, Liu Z, Wang Y, Zhang S, Yang H (2009a) Kinetically controlled InN nucleation on GaN templates by metalorganic chemical vapour deposition. J Phys D 42:145410

Wang X, Li X, Zhang L, Yoon Y, Weber P, Wang H, Guo J, Dai H (2009b) N-doping of graphene through electrothermal reactions with ammonia. Science 324:768

Wellard C, Hollenberg L (2005) Donor electron wave functions for phosphorus in silicon: beyond effective-mass theory. Phys Rev B 72:085202

Welser J, Hoyt J, Gibbons J (1992) NMOS and PMOS transistors fabricated in strained silicon/relaxed silicon-germanium structures. IEDM Tech Dig 1000–1002 doi:10.1109/IEDM.1992.307527

DOI 10.1007/978-3-642-27737-5_343-2

Williamson A, Wang L, Zunger A (2000) Theoretical interpretation of the experimental electronic structure of lens-shaped self-assembled InAs/GaAs quantum dots. Phys Rev B 62:12963–12977
Wong HS (2002) Beyond the conventional transistor. IBM J Res Dev 46:133–168
Xie Q, Madhukar A, Chen P, Kobayashi NP (1995) Vertically self-organized InAs quantum box islands on GaAs (100). Phys Rev Lett 75:2542
Yalavarthi K, Gaddipati V, Ahmed S (2011) Internal fields in InN/GaN quantum dots: geometry dependence and competing effects on the electronic structure. Phys E 43:1235–1239
Yang L, Park C-H, Son Y-W, Cohen ML, Louie SG (2007) Quasiparticle energies and band gaps in graphene nanoribbons. Phys Rev Lett 99:186801
Zandviet H, Elswijk H (1993) Morphology of monatomic step edges on vicinal Si(001). Phys Rev B 48:14269
Zhang Y, Tan Y-W, Stormer HL, Kim P (2005) Experimental observation of the quantum Hall effect and Berry's phase in graphene. Nature 438:201
Zheng Y, Rivas C, Lake R, Alam K, Boykin T, Klimeck G (2005) Electronic properties of silicon nanowires. IEEE Trans Electron Devices 52:1097–1103
Zhirnov VV, Cavin RK III, Hutchby JA, Bourianoff GI (2003) Limits to binary logic switch – a Gedanken model. Proc IEEE 91:1934–1939
Zhu W, Han JP, Ma T (2004) Mobility measurement and degradation mechanisms of MOSFETs made with ultrathin high-k dielectrics. IEEE Trans Electron Devices 51:98–105

Books and Reviews

Bimberg D, Grundmann M, Ledentsov N (1999) Quantum dot heterostructures. Wiley, New York
Datta S (2005) Quantum transport: atom to transistor. Cambridge University Press, Cambridge
Harrison P (2005) Quantum wells, wires and dots: theoretical and computational physics of semiconductor nanostructures, 2nd edn. Wiley-Interscience, Hoboken