

Three-dimensional spectral solution of the Schrödinger equation for arbitrary band structures

A. Trellakis and U. Ravaioli^{a)}

Beckman Institute and Department of Electrical and Computer Engineering, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801

(Received 16 May 2002; accepted for publication 28 June 2002)

We present a fast and robust method for the full-band solution of the Schrödinger equation on a grid, with the goal of achieving a more complete description of high energy states and realistic temperatures. Using fast Fourier transforms, the Schrödinger equation in the one band approximation can be expressed as an iterative eigenvalue problem for arbitrary shapes of the conduction band. The resulting eigenvalue problem can then be solved using Krylov subspace methods such as Arnoldi iteration. We demonstrate the algorithm by presenting an application, in which we compare nonparabolic effects in an ultrasmall metal–oxide–semiconductor (MOS) quantum cavity and a MOS quantum capacitor at room temperature. We show that for the cavity structure the nonparabolicity of the conduction band results in a significant lowering of high-energy electronic states and reshaping of the electron density, whereas the states and density in the MOS capacitor remain relatively unchanged. © 2002 American Institute of Physics.

[DOI: 10.1063/1.1502181]

I. INTRODUCTION

As electronic device dimensions approach the nanometer scale, quantum effects are expected to become increasingly important until they dominate electronic properties. Consequently there is considerable interest in the efficient numerical simulation of such structures, not only for exploring novel device architectures but also for maintaining reliability for scaled-down present-day devices. In this article we present a fast and robust method for the numerical solution of the full-band Schrödinger equation, with the goal of achieving a more complete description of high energy states and realistic temperatures. We then demonstrate the use of the algorithm by comparing nonparabolic effects on energy levels of an ultrasmall metal–oxide–semiconductor (MOS) quantum cavity and a MOS quantum capacitor at room temperature.

In order to understand why full-band algorithms are needed, we have to consider that many quantum simulations assume that the kinetic energy of carriers has a quadratic momentum dependence $E(\mathbf{p}) = \mathbf{p}^2/2m$. However, this model is only correct for the lowest carrier energies. At higher energies, the effective mass $m^*(E)$ becomes energy dependent, or, in other words, the kinetic energy contains cubic and higher order terms in momentum \mathbf{p} . If we increase the energy even further, even this approximation breaks down and a full-band description of the kinetic energy as a function of momentum $E = E(\mathbf{p})$ becomes necessary.

Therefore, practical numerical implementation of accurate quantum simulation of high energy quantum states poses a challenge that has not yet been addressed in previous studies. Atomistic models might be a possible solution,¹ however, the computational costs of these models are still pro-

hibitively high for general-purpose device simulation. Therefore, in order to simulate larger quantum structures it seems to be highly desirable to maintain a continuum description of the semiconductor that is corrected for high energy band structure effects. It is the goal in this article to outline the implementation of such a method, and to illustrate its feasibility by simulating two different MOS-type quantum structures.

II. COMPUTATIONAL METHOD

The proposed continuum full-band algorithm is formulated with the restriction of carrier motion to a single band. Interband transitions are ignored. This should be reasonable in most applications that involve silicon quantum structures, up to kinetic energies of about 1 eV. In this case, the kinetic energy of carriers in the band can be described by a function $E(\mathbf{p})$. This $E(\mathbf{p})$ might not be available in analytical form but only as an interpolated table, as is common for the output of band structure calculations. Nevertheless, independent of the origin of $E(\mathbf{p})$, the usual quantization rules apply and we may replace the momentum \mathbf{p} by the momentum operator $\hat{\mathbf{p}}$ to obtain the Hamiltonian. Thus, the Schrödinger equation for envelope wave functions that describe carrier motion under the influence of an external potential $V(\mathbf{x})$ can be written as

$$i\hbar \frac{\partial \Psi(\mathbf{x}, t)}{\partial t} = E(\hat{\mathbf{p}}) \Psi(\mathbf{x}, t) + V(\mathbf{x}) \Psi(\mathbf{x}, t), \quad \hat{\mathbf{p}} = \frac{\hbar}{i} \nabla \quad (1)$$

if $V(\mathbf{x})$ does not vary too rapidly within a lattice constant a , and if the dimensions of the device structures under consideration are much larger than a as well.²

Finding a solution to this equation is not as straightforward as for parabolic bands. This becomes quite obvious if we consider that now the kinetic energy operator $E(\hat{\mathbf{p}})$ may

^{a)}Electronic mail: ravaioli@uiuc.edu

contain the momentum $\hat{\mathbf{p}}$ up to very high and possibly infinite orders. Since $\hat{\mathbf{p}}$ is proportional to the gradient vector in the position representation, this means that we have to solve an equally high order differential equation. Numerical solution strategies based on finite differences or finite elements need to Taylor expand $E(\hat{\mathbf{p}})$ at least up to an order $n+1$ in \mathbf{p} to compute $E(\hat{\mathbf{p}})$ up to order n .

While it is possible to use such high order schemes,^{3,4} this approach can become inefficient for large n . Also, depending on the functional form of $E(\mathbf{p})$, the convergence radius of its Taylor series might be too small for polynomial approximations. Finally, $E(\mathbf{p})$ may only be available in tabulated form, making it impossible to derive high order polynomial approximations. While there are solutions to these problems, they usually involve matrix representations of the kinetic energy operator that are quite complicated and possibly dense.⁵

Therefore, considering these issues and the fact that the kinetic energy operator $E(\hat{\mathbf{p}})$ is diagonal in momentum space, a spectral solution approach using Fourier transforms seems to be a natural choice for this problem. If we choose to perform our calculation in position space, we may insert Fourier transforms (FTs) around the kinetic energy operator, which then reduces to simple multiplication in momentum space. We then obtain for the Schrödinger equation,

$$i\hbar \frac{\partial \Psi(\mathbf{x}, t)}{\partial t} = \text{FT}^{-1} [E(\mathbf{p}) \text{FT} \psi(\mathbf{x}, t)] + V(\mathbf{x}) \psi(\mathbf{x}, t). \quad (2)$$

In fact, one could view

$$E(\hat{\mathbf{p}}) \psi(\mathbf{x}, t) = \text{FT}^{-1} [E(\mathbf{p}) \text{FT} \psi(\mathbf{x}, t)] \quad (3)$$

as a definition for the operator $E(\hat{\mathbf{p}})$ in position space. Alternatively, we may also have expressed the Schrödinger equation in momentum space by inserting Fourier transforms around the potential energy as

$$i\hbar \frac{\partial \Psi(\mathbf{p}, t)}{\partial t} = E(\mathbf{p}) \psi(\mathbf{p}, t) + \text{FT} [V(\mathbf{x}) \text{FT}^{-1} \psi(\mathbf{p}, t)]. \quad (4)$$

This would give us a Hamiltonian similar to the ones employed in density functional calculations, except that in those cases the kinetic energy is quadratic in the electron momentum due to the atomistic nature of these models.⁶ However, for both position and momentum space formulation, the numerically difficult problem of applying a high order kinetic energy operator on a wave function has now been reduced to the more manageable one of calculating Fourier transforms.

III. IMPLEMENTATION

Even though electronic structure calculations are usually formulated in momentum space,⁶ it is not clear whether a position space or a momentum space formulation leads to a numerically superior algorithm for our continuum model. However, since all numerical routines used by the authors assume position space, it seems prudent to maintain compatibility by implementing the position space, Eq. (2).

In order to obtain an efficient numerical implementation it is mandatory to use fast Fourier transforms, which can be calculated using only $\mathcal{O}(N \log N)$ computational steps on a

grid with N nodes, as compared to $\mathcal{O}(N^2)$ for other methods. Since the pointwise multiplication of two functions involves only $\mathcal{O}(N)$ steps, examining Eq. (2) shows that we can calculate the action of the Hamilton operator on a wave function with $\mathcal{O}(N \log N)$ steps. Having this fast vector-matrix multiplication available, fast iterative eigenvalue solvers like the Arnoldi method immediately become the candidates of choice. In fact, since the Hamilton matrix is only implicitly available (unless one tries to calculate it), iterative eigenvalue solvers are the only practical choice. We use an Arnoldi solver with Chebyshev preconditioner to solve this eigenvalue problem.⁷

The use of fast Fourier transforms imposes stringent restrictions on the choice of the computational grid. Most importantly, the grid lines have to be equidistant in each coordinate direction and, additionally, for many fast Fourier transforms commonly available, the number of grid lines has to be a power of 2 in each direction as well. These requirements can result in some unexpected difficulties. Too coarse a grid will sample only parts of momentum space and provides insufficient spatial resolution, which is expected. However, an excessively fine grid results in a computational first Brillouin zone that is larger than the first Brillouin zone of the lattice. This creates a question of how to deal with the high momentum computational states that are introduced this way. Mapping these states to physically equivalent low momentum states using the periodicity of the physical momentum space is not a viable option, since this approach would introduce spurious low energy modes into the solution. Instead, for the computation we assign very high energies to these modes, thereby eliminating their influence on lower energy bound states.

Finally, we would like to mention that the full-band approach implemented here is not a good substitute for a finite difference based solver in the case of parabolic energy bands, since the requirement of equidistant grids is very restrictive. Also, the Fourier based matrix-vector multiplication used in the full-band solver scales as $\mathcal{O}(N \log N)$ for N grid nodes, whereas the banded matrix-vector multiplication used in the conventional solver scales as $\mathcal{O}(N)$. Therefore, for large parabolic problems it is always faster to use a solver based on finite difference or finite element discretization. This is especially true in the one-dimensional case, where tridiagonal eigenvalue solvers are far superior to any iterative method.

IV. APPLICATION

As an application of our full-band algorithm, we compare the influence of conduction band valley nonparabolicity for the MOS quantum cavity shown in Fig. 1 with the nanoscale MOS capacitor shown in Fig. 2. The quantum cavity simulated here distinguishes itself from the quantum dot floating gates used for flash memories^{8,9} in that here a quantum box of undoped silicon with dimensions $8 \times 10 \times 8 \text{ nm}^3$ is in direct contact with a 30 nm thick epilayer of undoped silicon below, allowing the exchange of electrons. The silicon box is surrounded at the sides and at the top by SiO_2 , with the oxide thickness in the normal direction being 2 nm

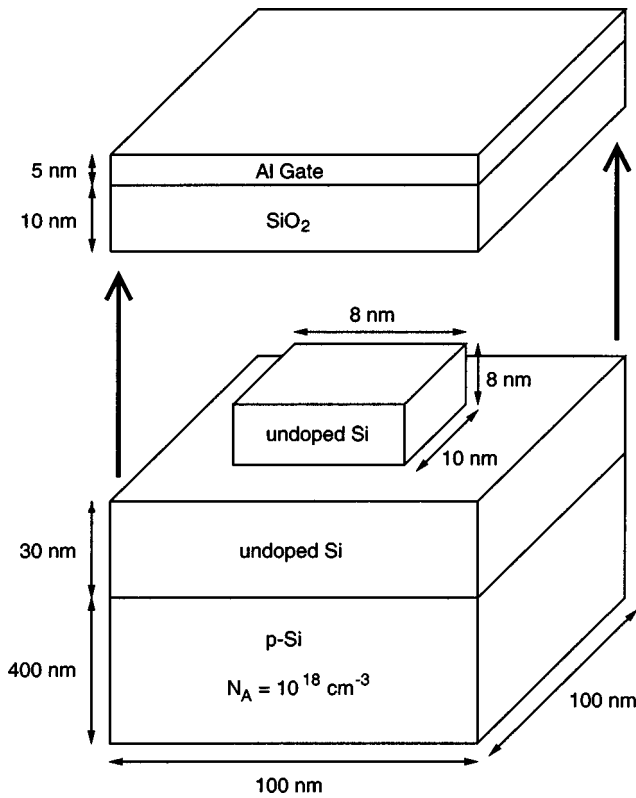


FIG. 1. MOS quantum cavity structure (expanded view). The top oxide and metal layers are lifted up to reveal the hidden rectangular silicon quantum cavity (center). Unlike a floating gate, the cavity is in direct contact with the substrate.

directly above the box and 10 nm everywhere else. The surface of the oxide is completely metallized by a 5 nm thick layer of Al. Vertical confinement towards the substrate is controlled by the electric field under the gate in conjunction with a highly *p*-doped deep substrate. Here, substrate doping is chosen as $N_A = 10^{18} \text{ cm}^{-3}$ in order to improve vertical confinement. The gate bias is 0.6 V, and the simulation temperature is 300 K.

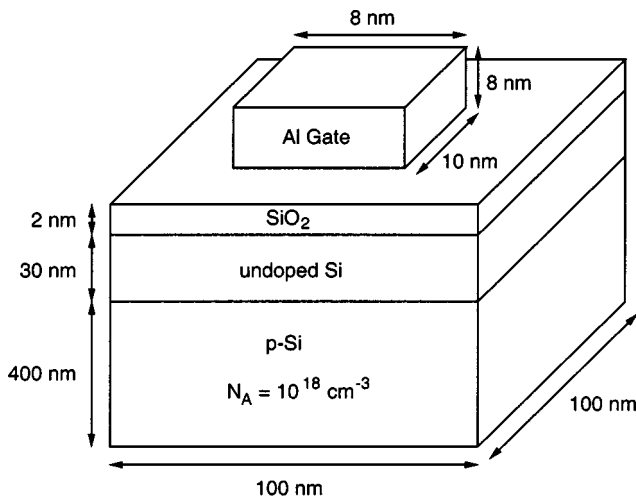


FIG. 2. Nanoscale MOS capacitor used for comparison with the quantum cavity. The small lateral dimensions of its gate give the capacitor the confinement characteristics of a quantum dot.

The MOS quantum capacitor used for the comparison (Fig. 2) is similar to the quantum dot simulated by Scholz *et al.*¹⁰ Here, a rectangular Al gate with dimensions $8 \times 10 \times 8 \text{ nm}^3$ is separated from the silicon substrate by 2 nm of gate oxide. The small lateral dimensions of the gate can be expected to give the capacitor the confinement characteristics of a quantum dot. The substrate consists of a 30 nm thick epilayer of undoped silicon situated next to the interface and a highly *p*-doped deep substrate ($N_A = 10^{18} \text{ cm}^{-3}$). The gate bias is 1.6 V, and the simulation temperature is again 300 K.

The physical model used to describe both structures consists of the full-band Schrödinger equation, Eq. (2), discussed in Sec. II,

$$E(\hat{\mathbf{p}})\psi_k + [V_h - e\phi + V_{xc}]\psi_k = E_k\psi_k, \quad (5)$$

coupled to a nonlinear Poisson equation

$$\nabla \cdot (\epsilon \nabla \phi) = -e(-n + p + N_D^+ - N_A^-), \quad (6)$$

where the unknowns are the wave functions ψ_k belonging to energy levels E_k and the electrostatic potential ϕ . Here, ϵ is the dielectric constant, n and p are the electron and hole concentrations, N_D^+ and N_A^- the ionized donor and acceptor concentrations, V_h the heterojunction step potential, and V_{xc} the exchange correlation potential in the local density approximation (see Ref. 11 for details concerning the numerical methods).

The nonparabolicity of the conduction band valleys in silicon is modeled using the following parameterizations described by Jacoboni and Reggiani:¹²

$$E_t(1 + \alpha E_t) = \frac{\hbar^2}{2m_t} \mathbf{k}_t^2, \quad E_1 = \frac{\hbar^2}{2m} \mathbf{k}_1^2. \quad (7)$$

Here, E_t , \mathbf{k}_t , and m_t , describe, respectively, the energy, wave number, and effective mass in the transversal direction of a valley in the silicon conduction band. Similarly, E_1 , \mathbf{k}_1 , and m_1 describe the corresponding quantities in the longitudinal direction. α is a parameter with the dimension of an inverse energy that describes the nonparabolicity of the valley. Following the suggestion in Ref. 12, we choose $\alpha = 0.7 \text{ eV}^{-1}$ for our calculation.

No special treatment of the silicon–oxide interfaces was performed. Since at these interfaces the conduction band structure changes, a more precise model than the one described above might be necessary in order to obtain correct results. However, the bound states of both structures can barely penetrate the oxide due to the more than 3 eV high conduction band step at the interface. Therefore, for the structures considered here we may safely assume that the kinetic energy $E(\hat{\mathbf{p}})$ is material independent without incurring major inaccuracies.

For other systems, as for example compound structures containing Si–SiGe or III–V type heterojunctions, the material dependence of $E(\hat{\mathbf{p}})$ may have a major influence on simulation results. There, it may be necessary to assume a position dependent band structure $E(\mathbf{p}, \mathbf{x})$, similar to the concept of a position dependent effective mass $m(\mathbf{x})$ that is generally used for parabolic band calculations.¹¹ However, for sharp material boundaries especially it is not clear if a position dependent band structure is a physically (or mathemati-

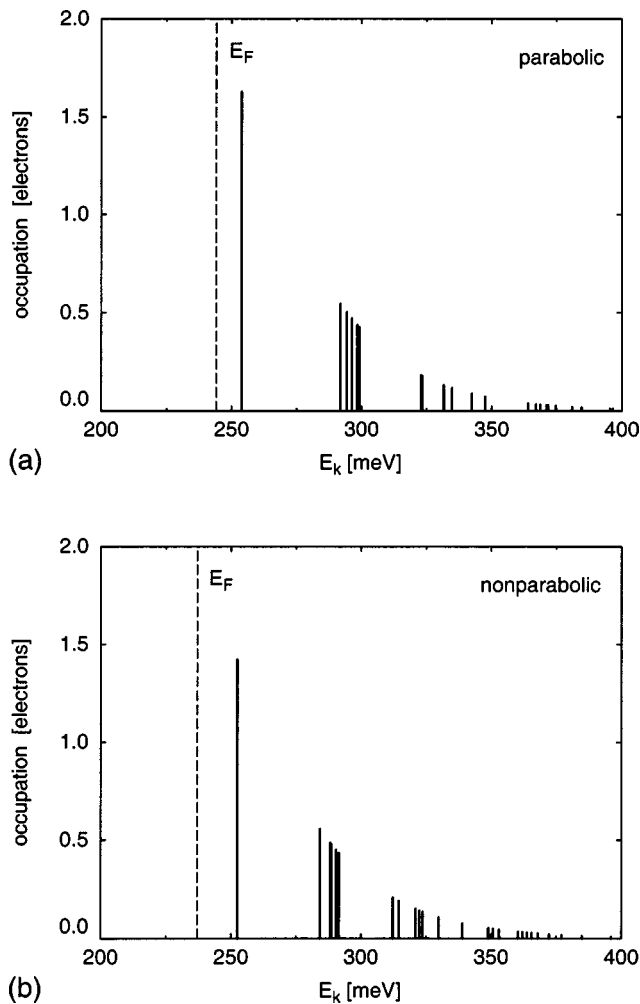


FIG. 3. Occupation numbers N_k of states E_k for the quantum cavity structure shown in Fig. 1 for parabolic (a) and nonparabolic (b) conduction band valleys. The Fermi level E_F is indicated by the dashed line and energies are normalized with respect to the bottom of the conduction band potential. The gate bias is 0.6 V, and the simulation temperature is 300 K.

cally) justifiable concept in the context of an envelope wave function description. Also, due to the noncommutativity of the momentum $\hat{\mathbf{p}}$ and position operator $\hat{\mathbf{x}}$ it is not clear how to properly quantize a kinetic energy of form $E(\mathbf{p}, \mathbf{x})$. These and related questions still require more detailed investigations.

Focusing our discussion initially on the quantum cavity, we examine its energy spectra for a parabolic and a nonparabolic valley simulation, as shown in Figs. 3(a) and 3(b). We find that the overall structure of the two spectra is very similar. The total occupation of the cavity is about 5 electrons in both cases, with about 1.5 electrons residing in the ground state, 2.5 electrons lying in a cluster of five states around 300 meV, and the states above 300 meV being mostly unoccupied (the fractionate occupations seen here are can be understood with in the context of thermodynamic equilibrium distributions such as the Fermi distribution).

However, we find that due to the conduction band nonparabolicity all states except the ground state are shifted towards lower energies with respect to the bottom of the conduction band. This shift is not too surprising if we consider

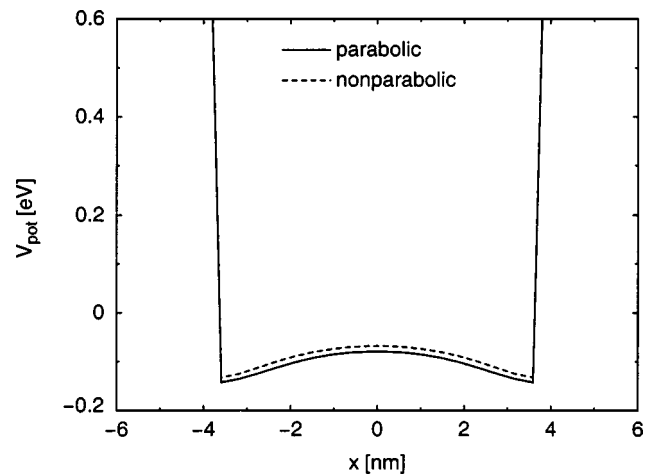


FIG. 4. Confinement potential V parallel to the Si-SiO₂ interface for the quantum cavity structure shown in Fig. 1 at a gate bias of 0.6 V. All plots are cut through the maximum of the electron density. The simulation temperature is 300 K.

that in Eq. (7) the effective electron mass increases with an increase in transverse momentum k_t within each conduction band valley. Since the states in all three energy ladders are occupied, as exemplified by the cluster of states located at 300 meV in Figs. 3(a) and 3(b), a significant influence of nonparabolicity on occupied electron states can be expected.

Despite this shift in energy, the total occupation of the cavity does not change much, since the Fermi level has been shifted downwards as well, and the total number of electrons is mainly determined by the strength of the electric field under the gate in the cavity. Instead, the electrons are redistributed over the energy spectrum, with the ground state occupation decreased and the occupation of the higher energy states increased.

Further insight can be gained if we examine the conduction band potential $V(\mathbf{x})$ and the electron density $n(\mathbf{x})$ in the cavity. Figure 4 shows the behavior of the potential parallel to the gate oxide interface, while Figs. 5(a)–5(c) show the density along all three coordinate axes. We find that nonparabolic effects result into an uplift of the conduction band potential with respect to the Fermi level inside the cavity. We also observe reshaping and widening of the electron density due to conduction band nonparabolicity.

The strong influence of nonparabolic effects on the electron density can be explained by the strong occupation of higher energy states for this structure, as observed earlier in the discussion of this. Since such high energy states have multiple density maxima they tend to have a significant effect on the shape of the overall electron density distribution. This effect is amplified by the particular upward warped lateral profile of the conduction band potential inside the cavity (Fig. 4). Such a profile tends to produce bipartite density distributions where the depth of the partition is very sensitive to the electron energy.

Turning our attention now to the MOS capacitor, we find that here nonparabolic effects are much less pronounced than for the quantum cavity. For example, if we compare the energy spectrum for the parabolic and the nonparabolic simulations [Figs. 6(a) and 6(b)], we find that the energy and

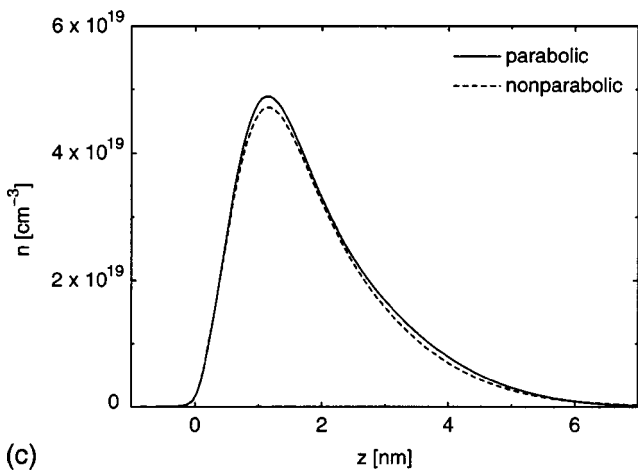
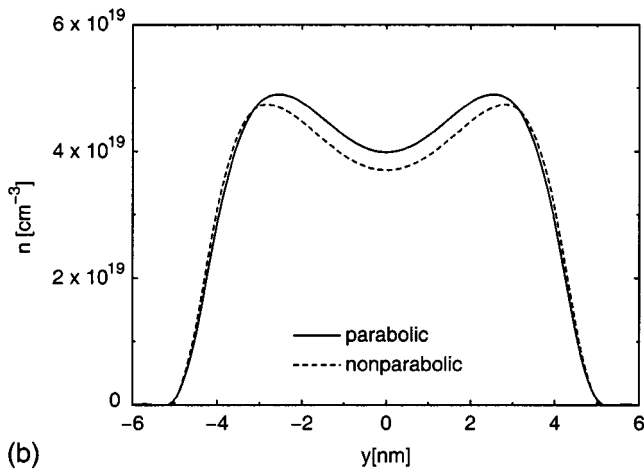
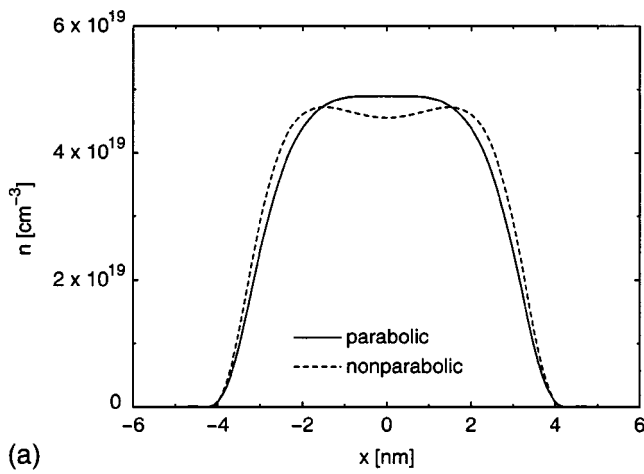


FIG. 5. Quantum electron density n parallel (x direction and y direction) and normal (z direction) to the Si-SiO₂ interface for the quantum cavity structure shown in Fig. 1. All plots are cut through a maximum of the electron density. The gate bias is 0.6 V, and the simulation temperature is 300 K.

occupation numbers of states with energies below 300 meV are almost unaffected by conduction band nonparabolicity. Closer examination shows that all these states belong to the one energy ladder where the longitudinal mass direction k_l is normal and the transverse mass directions k_t are parallel to the oxide interface. Energy levels with energies above 300 meV belong to the other two energy ladders. These high-energy states are strongly affected by nonparabolic effects

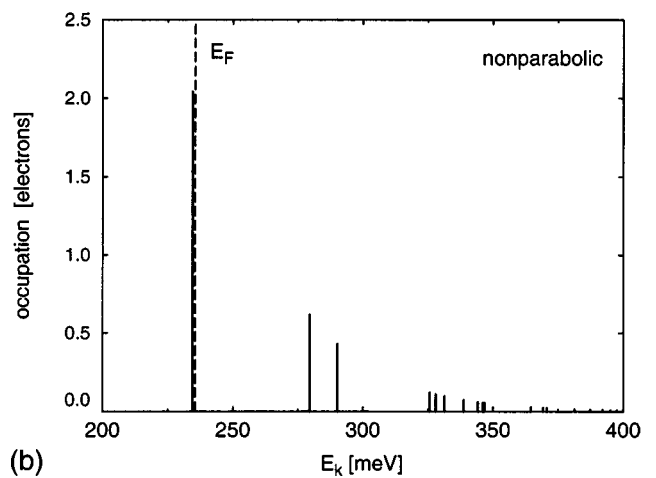
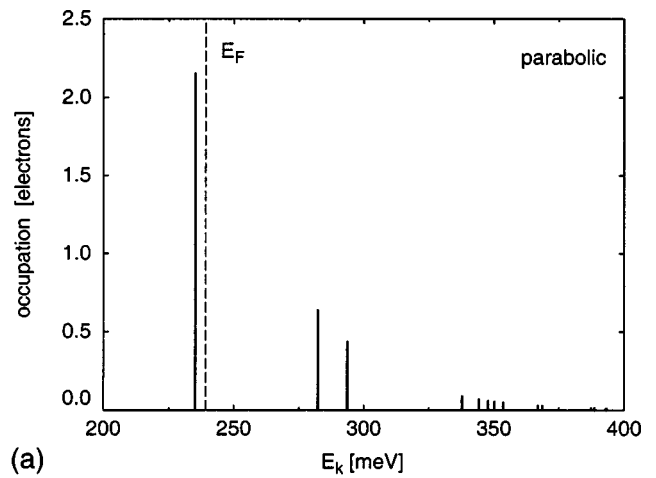


FIG. 6. Occupation numbers N_k of states E_k for the MOS capacitor shown in Fig. 2 for parabolic (a) and nonparabolic (b) conduction band valleys. The Fermi level E_F is indicated by the dashed line and energies are normalized with respect to the bottom of the conduction band potential. The gate bias is 1.6 V, and the simulation temperature is 300 K.

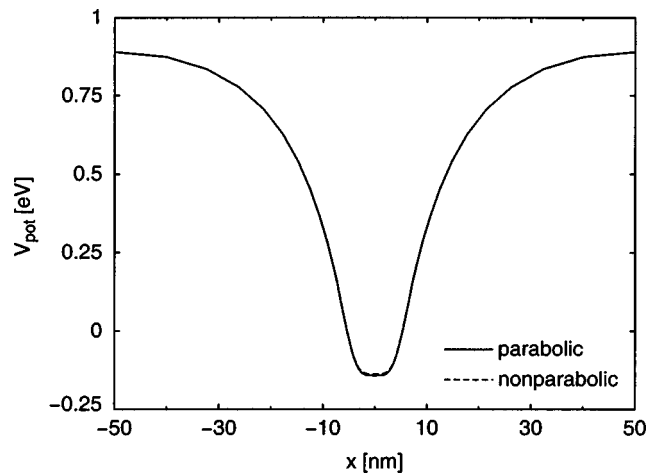


FIG. 7. Confinement potential V parallel to the Si-SiO₂ interface for the MOS capacitor shown in Fig. 2 at a gate bias of 1.6 V. All plots are cut through the maximum of the electron density. Note that the potentials for the parabolic and the nonparabolic cases are virtually identical.

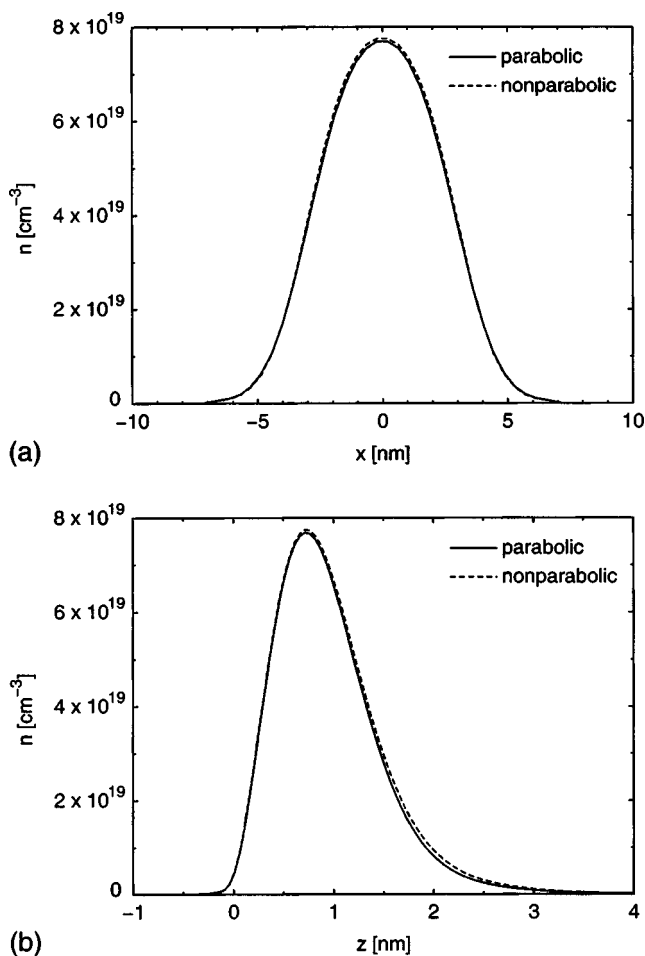


FIG. 8. Quantum electron density n parallel (a) and normal (b) to the Si-SiO₂ interface for the MOS capacitor shown in Fig. 2. All plots are cut through the maximum of the electron density. The gate bias is 1.6 V, and the simulation temperature is 300 K.

but they reside too far above the Fermi level to be occupied at room temperature.

The reduced importance of nonparabolic effects in the MOS capacitor is also reflected in its conduction band potential (Fig. 7) and electron density [Figs. 8(a) and 8(b)]. Here we find that the conduction band forms a largely parabolic potential well parallel to the oxide interface, with the electron density concentrated in a single maximum under the gate. These profiles are quite similar to the ones found for narrow MOS conduction channels,¹³ which is not too surprising considering the similarity between narrow channel structures and the MOS capacitor. In a way, the MOS capacitor behaves rather like a patch of two-dimensional electron

gas than a typical quantum dot structure, with the lateral dimensions of the confining potential being much wider than the typical thickness of an inversion layer (several nm). With most electrons residing in states where the transversal mass direction is parallel to the oxide interface, transversal wave vectors \mathbf{k}_t will remain small enough in a wide potential well to remain essentially unaffected by nonparabolic effects.

V. CONCLUSIONS

We have demonstrated a numerical method for the solution of the full-band Schrödinger equation in the one band approximation. The simulation algorithm is based on the use of fast Fourier transforms to express the full-band Hamiltonian operator as an iterative eigenvalue problem for arbitrary shapes of the energy band. The resulting eigenvalue problem is then solved using a standard iterative solver, such as an Arnoldi iteration. The applicability of the method is demonstrated for examples of a MOS quantum cavity and a nanoscale MOS capacitor at room temperature. For these structures, we have shown that the nonparabolicity of the conduction band results in significant lowering of high-energy electronic states and, in the case of the quantum cavity, also in reshaping of the electron density. Finally we have also shown that the energy spectrum and electron density of the cavity structure is much more sensitive to nonparabolic effects than those of the MOS capacitor.

ACKNOWLEDGMENTS

This work was partially supported by a DURINT Nanotechnology contract from the Army Research Office, by the Distributed Center for Advanced Electronics Simulation (DesCARTES) through National Science Foundation Grant No. ECS 98-02730, and by a graduate fellowship from the Computational Science and Engineering (CSE) program at the University of Illinois.

¹L.-W. Wang and A. Zunger, *J. Phys. Chem.* **98**, 2158 (1994).

²W. A. Harrison, *Solid State Theory* (Dover, New York, 1979), pp. 71–74.

³J. R. Chelikowsky, N. Troullier, and Y. Saad, *Phys. Rev. Lett.* **72**, 1240 (1994).

⁴J. R. Chelikowsky *et al.*, *Phys. Rev. B* **50**, 11355 (1994).

⁵H. Wei and T. Carrington Jr., *J. Chem. Phys.* **101**, 1343 (1994).

⁶M. C. Payne *et al.*, *Rev. Mod. Phys.* **64**, 1045 (1992).

⁷A. T. Galick, T. Kerkhoven, and U. Ravaioli, *IEEE Trans. Microwave Theory Tech.* **40**, 699 (1992).

⁸G. Iannaccone, A. Trellakis, and U. Ravaioli, *J. Appl. Phys.* **84**, 5032 (1998).

⁹J. J. Welser *et al.*, *IEEE Electron Device Lett.* **18**, 278 (1997).

¹⁰M. Scholze *et al.*, *VLSI Design* **9**, 231 (1998).

¹¹A. Trellakis and U. Ravaioli, *Comput. Methods Appl. Mech. Eng.* **181**, 437 (2000).

¹²C. Jacoboni and L. Reggiani, *Rev. Mod. Phys.* **55**, 645 (1983).

¹³A. Trellakis and U. Ravaioli, *J. Appl. Phys.* **86**, 3911 (1999).