Full Scale Self Consistent Simulation of Quantum Devices

Dejan Jovanovic and Jean-Pierre Leburton

Beckman Institute for Advanced Science and Technology,
University of Illinois,
Urbana, IL 61801

Abstract

A comprehensive self-consistent Schrödinger-Poisson simulation is presented. The model takes into account the various regions of reduced dimensionality throughout a particular structure which can consist of 0D, 1D, 2D electron gases or bulk. The feasibility of this level of simulation is made possible by an iterative extraction orthogonalization method for solving the Schrödinger equation. This method is superior to conventional eigenvalue solvers since it generates an arbitrary number of eigenstates and easily couples to the Poisson equation. High order effects such as exchange/correlation and single-electron charging are also included in the model. Transport is treated in the linear response regime and used to investigate the Coulomb blockade oscillations observed by Meirav[1].

I. Introduction

Recent advances in semiconductor fabrication techniques have resulted in a large variety of new experimental devices exhibiting quantum effects due to high degrees of confinement. In order to provide a realistic model for the physical analysis and design of these structures, we have developed a comprehensive self-consistent simulation tool that merges the statistical and quantum mechanical aspects of the problem. In this paper, we give the theoretical background of the model and its application to the analysis of the single-electron charging effects on transport through a quantum dot[1]. Previous efforts at simulating quantum structures have focused on localized regions in the device which exhibit reduced degrees of dimensionality. Although these simulations have yielded much information on the general quantum-mechanical properties in a device active region[2, 3, 4], their failure to integrate the dots (0D),leads (1D), contacts (2D), and bulk regions in a particular geometry prevents them from achieving good quantitative agreement with experimental data. This lack of coherence has been due in large part to computational time constraints in solving the Schrödinger equation over an arbitrarily large number of grid-points, N_G. Conventional eigenvalue solvers typically scale as N_G^3 and generate N_G eigenvectors thereby restricting their application to problems with small grids. In a general quantum device, however, eigenvalue problems need to be solved in all regions exhibiting reduced dimensionality and require a large N_G. In addition, only a few eigenenergies are often necessary owing to the small number of occupied eigenstates in a typical device. Clearly, a robust eigenvalue solver that addresses the above issues and allows easy self-consistent coupling to the Poisson equation is needed to bridge the gap between "localized" modeling and full scale device simulations.

II. Theoretical Background

We solve the Schrödinger equation with the iterative extraction orthogonalization method (IEOM) that propagates a set of eigenstates according to[5]

$$\left|\phi_n^i\right\rangle \;=\; e^{-\alpha\hat{H}}\left|\phi_n^{i-1}\right\rangle \quad , \quad \mathrm{i^{th}\ iteration} \tag{1}$$

where $\hat{H}$ is the Hamiltonian and n indicates a particular eigenstate. The parameter α is selected so as to minimize the error in the Taylor expansion of the exponential operator $exp(-\alpha\hat{H})$. In general, N_i iterations are required to maintain the accuracy of the Taylor expansion and eliminate all the error projections of the basis states $|m\rangle$ of $\hat{H}$ onto the initial guess state $|\phi_n^o\rangle$. Gram-Schmidt orthonormalization is performed over the entire set of states after each iteration to prevent any excited states from collapsing to the ground state. This procedure also eliminates all projections of states with $m < n$ such that after N_i iterations,

$$\left|\phi_n^{N_i}\right\rangle \;\simeq\; |n\rangle + \sum_{m=n+1} |m\rangle\, e^{-\alpha N_i(E_m-E_n)} \frac{\langle m \mid \phi_n^o\rangle}{\langle n \mid \phi_n^o\rangle}. \tag{2}$$

The repeated exponential scalings of the error projections and Gram-Schmidt orthonormalizations therefore eventually convert $|\phi_n^o\rangle$ into a pure basis state $|n\rangle$. The convergence criterion for the algorithm is determined by the expression

$$\alpha N_i \;=\; max\left[\frac{1}{E_m - E_n}\; log\left(\frac{\langle m \mid \phi_n^o\rangle}{\epsilon\,\langle n \mid \phi_n^o\rangle}\right)\right] \tag{3}$$

where ϵ is an imposed error tolerance. Clearly, α should be chosen to be as large as possible to reduce the number of iterations and still allow a high degree of accuracy in the Taylor expansion of $exp(-\alpha\hat{H})$. Eq. 3 points out an overall limitation of this method in that the number of iterations required to achieve convergence scales inversely with the energy separation between eigenstates. This problem can be alleviated somewhat by selecting initial states with the appropriate symmetry such that the error projections vanish for eigenstates with opposite parity. Following previous time dependent treatments[6, 7, 8, 9], the exponential operator in Eq. 1 is typically cast into a split form which separates its potential and kinetic energy components. The execution time of each iteration depends on how efficiently the discretized kinetic energy portion can be solved. For the rectilinear geometries often encountered in quantum devices, the kinetic energy operator can be further separated into (x, y, z) components which can be treated independently and solved by rapid elliptical solvers. In general, each application of the propagator scales with N_G and the Gram-Schmidt algorithm scales with $N_G N_E^2$ where N_E is the number of eigenenergies required in a particular simulation. If N_E is relatively small, as is often the case in nanostructures, $N_G N_E^2 \ll N_G^3$ and the IEOM shows a significant improvement over conventional eigenvalue solvers. The chief advantage offered by the rapid convergence of this method is the ability to accurately treat the dimensionality of each region in a quantum device out to an appropriate set of boundaries. We solve the 3D Schrödinger equation in the quantum dot (0D) regions, the 2D Schrödinger equation in the lead (1D) regions, and 1D solutions are obtained in the contact (2D) regions. Semiclassical solutions are

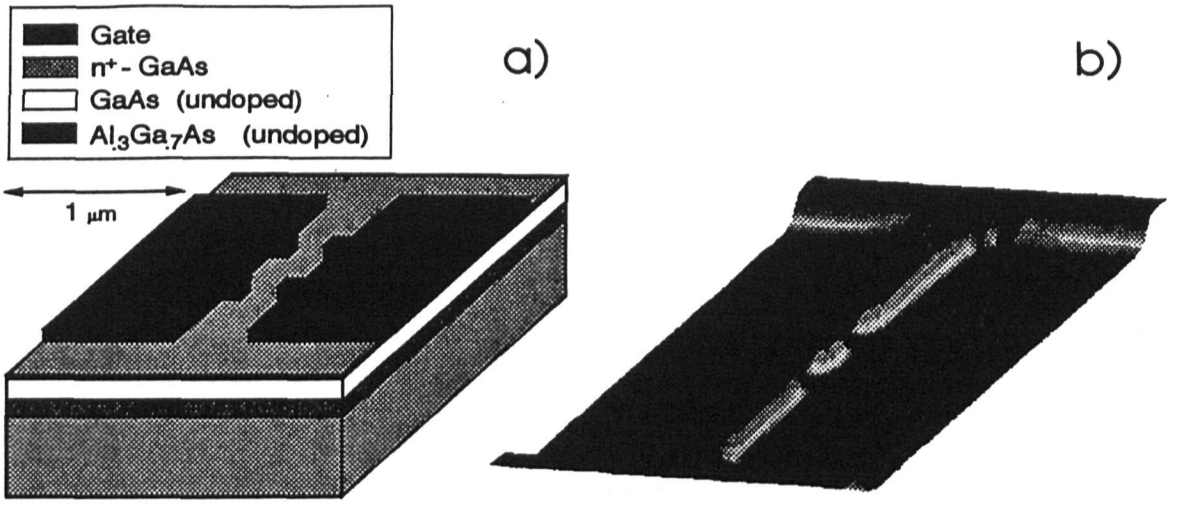

Figure 1: (a) Experimental device geometry used for investigating single electron charging effects. A negative bias on the gate confines charge in the lateral direction to form 1D leads and a quantum dot between the restrictions. Modulation of the bottom gate bias adds electrons to the quantum dot one at a time.(b) Theoretical electron density for a slice taken at the device active region. The formation of a quantum dot is visible as an island of charge between the two 1D leads.

used for carriers that do not exhibit confinement (holes) and regions that are not governed by quantum mechanics. Charge densities are constructed by scaling the amplitude squared of each wavefunction with its appropriate statistical weight[10]. The distribution functions used in the scaling of the 0D localized states are derived from the grand canonical ensemble with the constraint that only an integer number N^* of electrons occupy the quantum dot. N^* is determined for a given gate bias by free energy minimization[11]. The potential $\phi(\vec{r})$ in each region is inherently coupled by the Poisson equation which maintains continuity of $\phi(\vec{r})$ and its first derivative. The Schrödinger and Poisson equations are solved self-consistently by a modified Newton method that incorporates a line-search step to allow convergence at low temperatures. In addition, exchange and correlation effects are self-consistently treated with the Kohn-Sham approach[12] using Perdew-Zunger parameterization for the correlation potential[13]. Presently, the simulation is carried out in equilibrium although in principle, nonequilibrium solutions are possible for accurate evaluations of the quantum-mechanical current. We therefore compute the quantum transport properties of devices in the linear response regime which is typically the case for Coulomb blockade measurements.

III. Results

We have applied the simulation to the analysis of single electron charging effects in the inverted semiconductor-insulator-semiconductor structure reported by Meirav et al[1]. Fig. 1 shows the specific device geometry used in our investigation along with a charge density surface taken at the device active region. A negative bias on the top gate confines charge

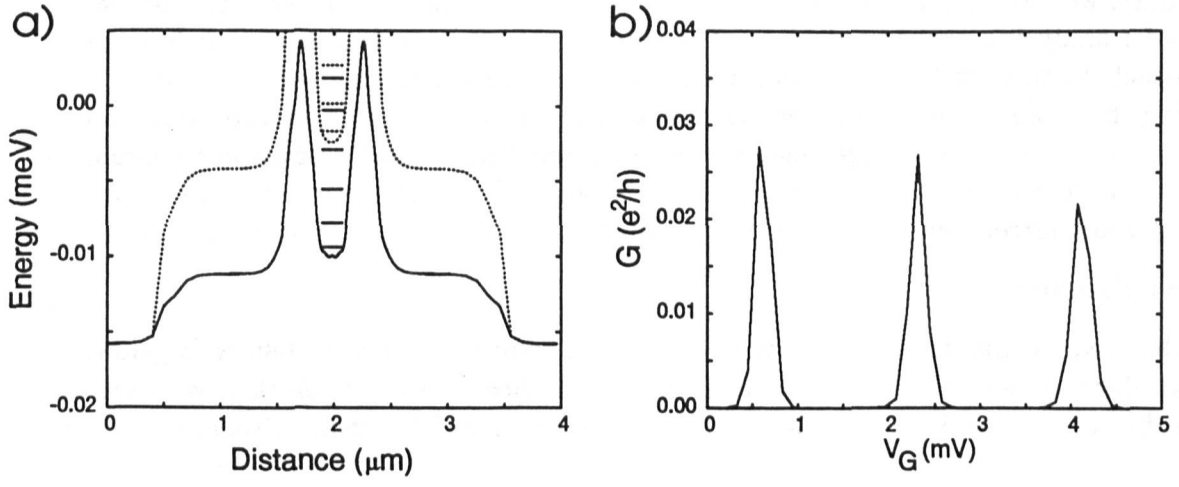

Figure 2: (a) The quasi-1D eigenenergies and their respective localized 0D eigenenergies. The eigenenergies associated with the first (solid) and second (dotted) quasi-1D states are the only ones occupied ($E_F = 0meV$). The formation of sharp barriers leads to tunneling into and out of the quantum dot. (b) The transport characteristics of the device exhibit sharp thermally broadened ($T = 50mK$, $L_o = .8\mu m$) peaks that correspond to the addition of a single electron to the dot and the subsequent lifting of the Coulomb blockade.

in the lateral direction to form 1D leads and a quantum dot between the restrictions. The quantum dot is visible in Fig. 1b as an island of charge between the two 1D leads. In addition, the presence of charge in the leads and contacts is also apparent. As the bias on the backgate is increased, electrons are added to the dot in single increments. This results from the electrostatic Coulomb repulsion between localized electrons which modifies the occupation probability of each level and allows only an integer number of electrons to occupy the dot. The 0D region is delineated by sharp barriers separated by a distance L_o which bring the dot into weak contact with the leads via tunneling. Fig. 2a shows the localized eigenenergies in the dot superimposed over their respective quasi-1D adiabatic eigenenergies obtained by extending the simulation of the leads into the quantum dot region. Although two quasi-1D modes have access to the dot, the upper channel is essentially closed owing to its relatively wide tunnel barriers and low statistical weight. The device should therefore exhibit single mode characteristics. Transport in the device is calculated using the appropriate expression derived for the Coulomb blockade regime[14, 15]. Transmission probabilities are evaluated using a transfer matrix calculation based on the quasi-1D eigenenergy as a function of distance. The charge imbalance caused by single electron occupation is modeled by minimizing the free energy with respect to the number of electrons. The difference between successive free energies then gives the charging energy required to add a single electron to the dot for that gate bias[14]. Transport characteristics of the device (Fig. 2) exhibit sharp thermally broadened peaks that correspond to the addition of a single electron to the dot and the subsequent lifting of the Coulomb blockade. The peak amplitude and periodicity show good agreement with the experimental data[1].

IV. Conclusion

In summary, we have demonstrated the viability of a comprehensive device-scale simulation tool for analyzing quantum devices. The lengthy eigenvalue calculation, which is the chief obstacle to this level of simulation via conventional methods, has been overcome by employing the rapidly converging iterative extraction orthogonalization method that can treat an arbitrary number of eigenenergies and wavefunctions. Finally, transport characteristics for a device exhibiting single-electron charging effects were obtained and shown to exhibit good agreement with experimental findings.

Acknowledgements

The authors would like to thank Tetsuo Kawamura, Paul von Allmen, Jeffery Bigelow, Massimo Macucci, and Karl Hess for their helpful advice. Support for this work was provided by NSF Grant ECS 91-20641 and the Joint Services Electronics Program.

References

[1] U. Meirav, M. A. Kastner, and S. J. Wind, Phys. Rev. Lett. **65**, 771 (1990).
U. Meirav, P.L. McEuen, M.A. Kastner, E.B. Foxman, A. Kumar, and S.J. Wind, Z. Phys. B **85**, 357 (1991).

[2] S. E. Laux and F. Stern, Appl. Phys. Lett. **49**, 91 (1986).

[3] A. Kumar, S. E. Laux, and F. Stern, Phys. Rev. B **42**, 5166 (1990).

[4] D. Jovanovic and J. P. Leburton, IEEE Electron Device Lett. **14**, 7 (1993).

[5] R. Kosloff and H. Tal-Ezer, Chem. Phys. Lett. **127**, 223 (1986).

[6] M. D. Feit, J. A. Fleck, Jr., and A. Steiger, J. Comp. Phys. **47**, 412 (1982).

[7] R. Kosloff, J. Phys. Chem. **92**, 2087 (1988).

[8] M. H. Degani, Appl. Phys. Lett. **59**, 57 (1991).

[9] J. Bigelow, J. P. Leburton, and M. Degani, *Proc. Int. Workshop Computational Electron.* Urbana, IL., 309 (1992).

[10] D. Jovanovic and J. P. Leburton, to be published

[11] H. Van Houten, C.W.J. Beenakker, and A. A. M. Staring, in *Single Charge Tunneling, NATO Advanced Study Institute, Series B: Physics*, edited by H. Grabert and M.H. Devoret (Plenum, New York, 1991).

[12] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[13] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[14] C. W. J. Beenakker, Phys. Rev. B **44**, 1646 (1991).

[15] Y. Meir, N. S. Wingreen, and P. A. Lee, Phys. Rev. Lett. **66**, 3048 (1991).