Full Quantum Simulation, Design, and Analysis of Si Tunnel Diodes, MOS Leakage and Capacitance, HEMTs, and RTDs

Roger Lake

Department of Electrical Engineering, University of California, Riverside, California 92521-0204

Abstract

The Nanoelectronic Engineering Modeling software (NEMO) has been used to model the quantum electron and hole transport and charge in a wide variety of material systems and semiconductor devices. This paper provides an overview of NEMO’s current status, its applications, and its theoretical extensions.

Introduction

The NEMO program was developed during the years 1993 to 1997 at which time the source code was delivered to the U. S. Government. Raytheon owns the software and is not currently distributing it. Raytheon has been negotiating with NASA JPL to commercialize NEMO. NASA JPL owns a government license to the NEMO source code and has enhanced the 1997 version to include three dimensional calculations of the eigenstates of quantum dots. Raytheon extended the quantum transport theory to model the phonon-assisted, indirect tunnel current of low-temperature MBE grown Si tunnel diodes. The NEMO software has been used to support experimental programs working on a wide variety of materials and devices: MOS leakage and capacitance, Si / SiGe tunnel diodes, In_{0.47}Ga_{0.53}As / In_{0.52}Al_{0.48}As / InAs / AlAs high electron mobility transistors (HEMTs) and resonant tunneling diodes (RTDs), InAs / AlSb / GaSb HEMTs and RTDs, QWIPs and quantum cascade lasers. Below, we give examples of some of these applications.

Si MOS

NEMO was used extensively in the Quantum MOS program to analyze ultra-thin oxides, explore the fabrication limitations of Si / SiO_2 RTDs, and design and analyze Si / SiGe tunnel diodes. The objective of the program was to develop a Si-based tunnel diode or resonant tunnel diode that could be integrated into a CMOS fabrication process. The benefits of integration have been described by Seabaugh et al. Fig. 1 shows calculations comparing the capacitance of a MOS structure reported by Brar et al. The calculation is performed using a 4-independent band model which includes one band for the electrons in the 4 equivalent X valleys, 1 band for the electrons in the 2 equivalent X valleys, and the

Figure 1. (a) Capacitance of 3.1nm Al / SiO_2 / n-Si with n = 3 x 10^{10} / cm^2. (b) Band edges and Fermi level. (c) Electron density in accumulation layer.
light and heavy hole bands. A single-band, effective-mass Schrödinger’s equation is solved using Green function techniques for each band. The quantum charge and current is calculated for each band and the results are added. For each of the 4 calculations, the SiO₂ is assumed to have an effective mass of 0.3 m₀ where m₀ is the bare electron mass. The capacitance is found by numerically calculating \( \delta Q/\delta V \) using a 10 mV \( \delta V \). NEMO calculates the quantity \( \delta Q \) given by Eq. (1) where \( \delta q_j \) is the charge differential at each node \( j \) and \( a \) is the node spacing. In this case, \( \delta Q \) is simply the charge in the accumulation layer. This approach was adopted for more complicated devices where the positive and negative \( \delta q \) contributions alternate throughout the device. \( \delta Q^+ \) defined as in (1) with \( \delta q_j > 0 \) is also calculated as a check to ensure that \( |\delta Q^- - \delta Q^+| \leq \delta Q \).

\[
\delta Q^- = a \sum_j (\delta q_j < 0)
\]

The calculations agree well with the experimental data in the accumulation region. The quantum charge calculation for 3V applied bias is shown in Fig. 1c. The peak of the charge distribution lies approximately 0.8 nm from the oxide.

These same calculations were also used to calculate the thin oxide tunnel currents measured by Brar et al. Fig. 2 shows the comparison with the measurements for 4 thicknesses ranging from 1.65 nm to 2.86 nm. Note that the current is carried by electrons tunneling out of the accumulation layer into the Al. The boundary conditions described in references [6, 9] permit NEMO to model current injected from quasi-2D states in the same way that it models current injected from the continuum. Both Thomas-Fermi and Hartree self-consistent calculations are shown on each graph. The simulations were performed as a quick sanity check of the experimentally extracted effective mass of 0.3 m₀ for the SiO₂. With this effective mass, used in simple, single-band effective mass calculations, the calculated and experimental current track each other over 10 orders of magnitude for all of the I-Vs shown in [8].

**Si Tunnel Diodes**

Our first working design for a low temperature, MBE grown Si / SiGe tunnel diode was the direct result of fast, simple, multiple-single-band simulations performed with NEMO. Significant experimental effort had been spent on devices similar to that shown in Fig. 3a. It consisted of an 8 nm Si₁₀Ge₀₅ layer sandwiched between Si layers. Heavy doping continued throughout the Si₁₀Ge₀₅ layer with the n⁺ - p⁺ junction at the center. A few iterations with NEMO lead to the design shown in Fig. 3b. NEMO showed that the length of the tunnel region could be cut in half and that the doping

---

**Fig. 2.** Measured and calculated tunnel current through thin oxides. Thicknesses are shown in legends.

**Fig. 3.** (a) NEMO simulation of initial tunnel diode design. (b) Design after NEMO simulations. Experimental results are described by Rommel et al.
throughout the tunnel region could be eliminated with essentially no change in the conduction - valence band overlap. The delta-doping planes were sufficient to fix the band overlap and contain the built-in electric field within the tunnel region. These were significant design improvements. The tunnel current depends exponentially on the tunnel barrier length. Cutting the length of the tunnel region in half results in a large increase in the current density. The speed index, one figure of merit, is directly proportional to the current density. Removing dopants from the tunnel region removes one cause of gap states which are responsible for the excess current. Reducing the excess current improves the peak-to-valley current ratio (PVCR), a second figure of merit.

These design simulations were again multiple single band calculations in which Schroedinger’s equation was solved independently for each band and then the results for the quantum charge were added to perform a self-consistent calculation of the electrostatic potential. One novel twist from the previous calculations is that the strain breaks the degeneracy of the 6 conduction band valleys and two valence bands. The actual 4 bands resulting from the calculation are shown in Fig. 3. Such an approach allowed fast turn around for design, but, since it was based on single band equations, was unsuitable for calculations of the indirect, interband, tunnel current. The Green function implementation of the multiple sequential scattering approach algorithm described in reference [6] would have worked, however, we felt that it would be too numerically demanding. Therefore, we derived an expression that is essentially Fermi’s Golden Rule written in Green function form which calculates the overlap of the wavefunction envelopes, one injected towards the tunnel junction from the left emanating from a conduction band X valley and the other injected towards the tunnel junction from the right emanating from a valence band near Γ. The results of the current calculations are described in reference [4].

\[ \text{In}_{0.47}\text{Ga}_{0.53}\text{As} / \text{In}_{0.52}\text{Al}_{0.48}\text{As HEMTs} \]

For the development of high-speed analog-to-digital converter (ADC) and tunneling static RAM (TSRAM) circuits at Raytheon, NEMO was used extensively to understand the gate leakage current and the non-alloyed contact resistance of In\textsubscript{0.47}Ga\textsubscript{0.53}As / In\textsubscript{0.52}Al\textsubscript{0.48}As HEMTs. The gate leakage current of a HEMT shows a strong temperature dependence (see Fig. 4a) which contradicts the simple analytical expression for the tunneling current in an insulator. The temperature dependence is not the result of thermionic emission over the barrier, but the result of the strong energy dependence of the tunneling transmission coefficient shown in Figure 4b. From the Fermi level of the metal to the top of the barrier, the transmission increases by 14 orders of magnitude. This results in a strong temperature dependence of the integral of the current distribution J(E) which is the product of T(E) and the Fermi distribution F(E-μ). A similar temperature dependence is observed in Si MOS diodes and it has been proposed to use the effect for an integrated on-chip temperature gauge.

\[ \text{In}_{0.52}\text{Al}_{0.48}\text{As} \]

The non-alloyed contact process places the conflicting requirements on the In\textsubscript{0.52}Al\textsubscript{0.48}As gate layer to act as a good insulator under the gate and to act as a low resistance layer under the source and drain contacts. Figure 5a shows a device cross section, and Figures 5b and 5c show the bands under the gate and source/drain, respectively. A parametric study was performed to determine the contact resistance and gate leakage current for various epi stacks. The calculated contact resistance for this particular structure was 0.6 μΩ cm\(^2\). Calculations of the contact resistance were performed in two ways. For each case, the 20 nm In\textsubscript{0.47}Ga\textsubscript{0.53}As quantum well was treated as a contact in equilibrium with the substrate. In the first approach, electrons were injected from the metal non-alloyed contact and the current was calculated coherently through the whole epi stack. In the second approach, the epi stack from the metal to the n+ InP was considered to be in equilibrium with the metal. The electrons were injected from the n+ InP layer into the In\textsubscript{0.47}Ga\textsubscript{0.53}As well. The results from the 2 calculations were essentially identical. This indicates
that the main contribution to the contact resistance arises from transport over the In_{0.52}Al_{0.48}As gate barrier. For this fixed gate-barrier thickness, we found the epi structure shown in Fig. 5c to be the optimum one for reducing the non-alloyed contact resistance. We could only reduce the contact resistance by reducing the thickness of the In_{0.52}Al_{0.48}As gate barrier. Such a reduction then affects the threshold voltage and the gate leakage current.

In_{0.47}Ga_{0.53}As / In_{0.52}Al_{0.48}As / InAs / AlAs and InAs / AlAs RTDs for THz Applications

An approach that is being pursued to build THz receivers requires RTD based local oscillators (LOs) that operate at low THz frequencies. Our design approach for the RTDs has been to extract the circuit parameters for the small-signal model described by Brown et al.\textsuperscript{14} from NEMO DC calculations of the I-V and C-V curves. The use of the DC curves should be a good approximation as long as the inverse of the quantum state lifetime is larger than the frequency of operation. The physics that we have found particularly interesting is the quantum capacitance that peaks precisely at the maximum magnitude of the negative differential resistance (NDR) which is the desired operating point. The peak in the quantum capacitance results from the discharge over a small voltage range of the charge in the quantum well. Based on other calculations, we believe that a Hartree calculation overstates the peak.\textsuperscript{15} However, it does exist and has been measured experimentally.\textsuperscript{16} Fig. 6 illustrates the calculations and predictions for the resistive cutoff frequency and self-resonant frequency for an InAs / AlAs RTD.

Acknowledgements: This work was supported in part by Texas Instruments, Raytheon, and AFOSR/DARPA.