Molecules on silicon: Self-consistent first-principles theory and calibration to experiments

T. Rakshit,[1] G.-C. Liang,[1] A. W. Ghosh,[1] M. C. Hersam,[2] and S. Datta[1]
[1]*School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907, USA*
[2]*Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois 60208, USA*
(Received 28 March 2005; published 6 September 2005)

In this paper based on a fully self-consistent first-principles transport calculation we show that in a silicon-molecule-STM (scanning tunneling microscope) structure, negative differential resistance (NDR) occurs for positive substrate voltage, mediated by the molecular levels on *p*-type Si(100) substrates. The positions of the NDR peaks are determined by (i) the equilibrium location of the relevant level with respect to the Si(100) Fermi energy (E_f) and (ii) how fast the level slips past the band edge. Based on (ii), we predict that by varying the STM tip-to-molecule spacing, the NDR peak location can be shifted in the current-voltage (I-V) characteristics. Recent experiments indeed show the NDR peak movement as a function of tip distance, thus strongly supporting this molecule-mediated mechanism of NDR on *p*-type substrates. Extrapolation from the NDR peak shift on the voltage axis can be used to predict the equilibrium location of the molecular level with respect to the Si(100) substrate Fermi energy. Based on (i), we conclude that to observe NDR in the negative bias direction on *n*-type substrates, much higher voltages are required. However, polarity-reversed NDR on *n*-type substrates may be observed under conditions beyond the model described here, and the relevant scenarios are discussed.

DOI: 10.1103/PhysRevB.72.125305 PACS number(s): 05.10.Gg

I. OVERVIEW

Molecular electronics traditionally has used metallic electrodes as substrates, taking advantage of the ability to grow molecular monolayers on metallic substrates. This has led to wide research activity, both theoretical (Ref. 1 and references therein) and experimental (Ref. 2 and references therein), on molecular electronics on metals. Unfortunately, metals also introduce undesirable effects, such as gap states and weak ionic bonds. In contrast molecular electronics on semiconducting substrates is a nascent field where experimental (Ref. 3 and references therein) as well as theoretical work are just emerging.[4,5] Silicon contacts provide a lot of desirable advantages over their metallic counterpart vis-à-vis molecular conduction, including a highly developed infrastructure and mature material processing connected with the I-C industry, strong covalent bonds that can tether molecules and elimination of metal-induced gap states. Furthermore, silicon substrates introduce different physics because of the presence of a band edge, which can act as an energy filter, accentuating electronic features of the molecule and making the contacts active players in the conducting process rather than passive injectors. The ability to engineer the semiconductor properties through doping provides an extra degree of control, which could conceivably be exploited in designing molecular devices.

Understanding molecular conduction through any complex substrate must invariably start with drawing a band diagram for transport. Such a band diagram helps in qualitatively understanding experiments, such as the rectifying I-V characteristics observed on silicon[6] and measurements made by a number of groups on GaAs (see Ref. 7 for references). In addition, it allows us to make predictions. Based on a band diagram (Fig. 1), we predicted a feature of a molecular conductor docked onto silicon and probed with a scanning tunneling microscope (STM) tip. Specifically, we predicted that the STM field would move the molecular levels into the band gap of silicon, which would thereafter turn off conduction, leading to negative differential resistance (NDR) peaks in the current-voltage (I-V) curve.[4] The requirement for this observation was that the substrate be heavily doped and that the measurement be taken at high bias to allow sufficient voltage for the levels to slip past the band edge. Since the conducting levels encounter the band edge only for one bias polarity, one of the key features of our prediction was that the NDRs should be observed for positive substrate bias on $p{+}{+}$ silicon substrates and negative for $n{+}{+}$ substrates if the molecular levels are close enough to the Si(100) bulk band edge.

Although band diagrams, such as Fig. 1, give us much insight, quantitative prediction poses numerous challenges that silicon substrates bring into the picture compared to their metallic counterparts.[8] Even a minimal model for silicon

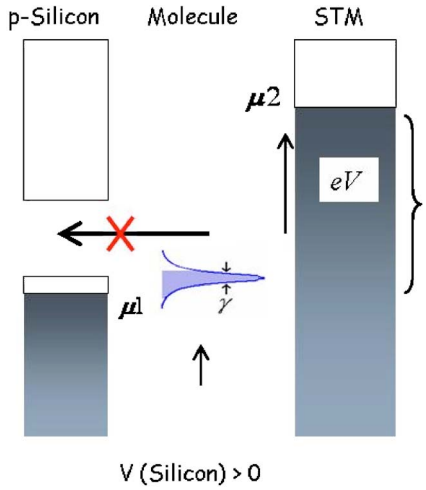

FIG. 1. (Color online) Silicon-molecule-metal band diagram.

needs to do justice to its bulk band structure, including its band gap and longitudinal and transverse effective masses. In addition, one needs to worry about band bending in a silicon substrate that is easily several atomic layers thick even for very heavily doped silicon. Furthermore, the surface chemistry of silicon needs to be considered in a way that it captures the effects of surface reconstruction and dimerization, and associated surface states inside the band gap. The band structure and surface electrostatics of silicon determine the electron injection from the contact into the molecule, whereas the surface chemistry determines the molecular orientation and coupling strengths. Typically, the former set of quantities is well studied by device theorists, whereas the latter is commonly studied by surface chemists, and needless to say, the theoretical techniques developed by these two communities are not necessarily compatible. Hence, in this paper we develop a technique that makes use of the engineers' semi-empirical treatment of bulk silicon electrostatics and band structure and the chemists' atomistic treatment of molecular bonding to the silicon surface.

In a past publication[4] we predicted that doping- and bias-dependent NDR is expected to be observed at silicon-molecule interfaces. This one-sided doping-dependent NDR was subsequently demonstrated in a series of experiments by Hersam *et al.*[7] In the earlier publication, we focused on a model geometry for a styrene molecule connected to hydrogen-passivated silicon, motivated by the availability of comparable theoretical analyses of the geometry in the literature.[9] However, the geometry studied did not correspond, in detail, to the actual experiments on NDR through isolated TEMPO (2,2′,6,6′-tetramethyl-1-piperidynyloxy) molecules on clean, rather than hydrogen-passivated *p*-type silicon. In this paper, we critically examine the experiments with a calculation that goes beyond our earlier attempts. Specifically, we address the exact experimental setup for TEMPO on silicon. We discuss the different ingredients of the heterostructure in detail, such as the bonding of TEMPO on the silicon surface, the self-consistency in the electrostatics, and the effect of the vacuum gap between the tip and the sample.

In Sec. II, we outline the solution procedure and construct an elementary model based on the same procedure to explain the essential physics behind NDRs in I-Vs in silicon-molecule-metal structures. Section III details the first-principles method based on the same procedure. Section IV presents the results, while Sec. V discusses possible scenarios that can help explain the physics of NDR on *n*-type Si substrates. Before we move on we will summarize our key results.

A. Key results

A fully self-consistent transport calculation based on density-functional theory (DFT) and nonequilibrium Green's function (NEGF) yields I-Vs that resemble the shape of the experimental I-V of TEMPO on $p++\mathrm{Si}(100)$.[3] Specifically, it shows NDR for positive bias direction and not for negative. Because the exact distance of the STM tip to the sample is an unknown, we do not focus on the absolute magnitude of current. However, we have succesfully predicted polarity-dependent relative magnitude changes in current as a function of tip distance to the molecule.[10] The reduction in current with the tip retraction is predicted to be less for the positive bias (NDR direction) than for negative (non-NDR direction). We point out that the magnitude of current at the NDR peak position is not only controlled by the absolute distance of STM tip but also the doping level of the Si(100) substrate which, in turn, gives rise to the anomalous polarity-dependent current reduction.

Electrostatically, we predict that because of a varying voltage drop in the intervening vacuum gap, the NDR peak location will shift in the I-V as the tip is moved, specifically the peak will show up at a higher bias as the tip is retracted and vice-versa. Experimental evidence of the peak movement[10] substantiates that it is, indeed, a molecular level that moves past the band edge to give rise to NDR. This phenomenon can also be applied to locate the equilibrium position of a molecular level that contributes to NDR, with respect to the Si(100) E_f.

For NDR to occur through unoccupied molecular levels on *n*-type substrates, we observe that the unoccupied molecular levels need to be far closer to the band edge than they are theoretically predicted.[4] We discuss possible origins of polarity-dependent NDR on *n*-type substrates[11] and suggest experiments that could help sort out these issues. We also discuss the possible origin of multiple NDR peaks in the I-Vs.

II. SOLUTION PROCEDURE

We calculate the I-V characteristics through a structure that consists of two large electrodes and a molecule sandwiched in between. Connection to the large electrodes converts the system to an open one that is driven out of equilibrium by an applied bias. Our solution procedure is based on a self-consistent field (SCF) scheme. In SCF, electron-electron interactions are incorporated through a one-electron potential U_{SCF} determined from the one-electron density matrix ρ that, in turn, depends on U_{SCF}. Because we are interested in the I-V of the molecule and the large contacts serve as injector and/or extractor of electrons, we would like to separate the molecule as the device and represent the effect of the large contacts on the molecule as self-energy functions ($\Sigma_{1,2}$) (Fig. 2). The effect of scattering is included within NEGF as described in Ref. 12. Our prototype for a molecule is TEMPO (shown later in Fig. 6) that has been experimentally characterized on both *p*- and *n*-type substrates. Once we have constructed a device Hamiltonian and figured out its boundary conditions, we self-consistently calculate ρ and U_{SCF}. Current (I) is calculated from the the converged U_{SCF} and the contact Fermi functions ($f_{1,2}$). We calculate ρ and I based on nonequilibrium Green's function (NEGF) formalism

$$G(E) = [ES - H - U - \Sigma]^{-1},$$

$$\rho = \int dE \left[\frac{-i}{2\pi} G^{<}(E) \right],$$

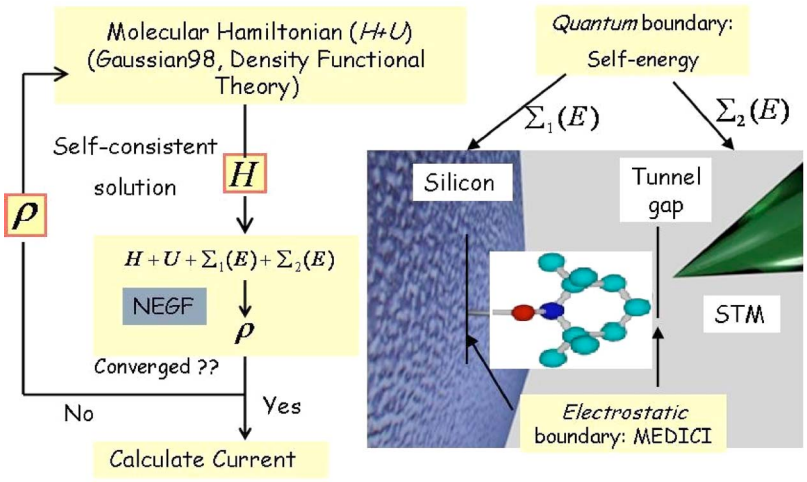

FIG. 2. (Color online) Schematic: Block diagram of detailed solution procedure based on DFT and NEGF.

$$-iG^{<}(E) = G(\Gamma_1 f_1 + \Gamma_2 f_2)G^{\dagger},$$

$$I = \frac{2e}{h}\int Tr(\Gamma_1 G \Gamma_2 G^{\dagger})(f_1 - f_2)dE. \quad (1)$$

A. Elementary example

To elucidate the solution procedure we first construct an elementary model where $H, U, \Sigma_{1,2}$ are simple matrices that are 2×2 in size. The two levels obtained from H represent one unoccupied and one occupied level of a molecule. Based on this model we explain the basic physics that is involved in observing NDR at the silicon-molecule interface. In Sec. III we get into the details of a first-principles solution that involves Hamiltonians based on 6-31$g(d)$ basis set. In general, an *ab initio* description of the molecule based on an appropriate basis set involves H that is $N\times N$, where N can be quite large. The corresponding matrices for $\Sigma_{1,2}$ are the same size as H, as described in Sec. III.

The electrostatic potential U consists of two parts, the Laplace part (U_L) and a self-consistent (U_{SCF}) part. As pointed out in Sec. I, U_L plays a very significant role in obtaining NDR at the silicon-molecule interface. It is the potential drop across the molecule that provides the slippage of molecular levels past the silicon band edge. In Sec. II B, we will introduce a factor η that provides a measure of how fast the levels move with respect to the band edge. In the elementary example a simple Hubbard-type single-electron charging is used along with the U_L to calculate the I-V self-consistently.

B. Electrostatics: Movement of molecular levels, η

U_L decides the boundary condition on the molecule. It is set by the silicon surface potential on one side and the potential drop in vacuum on the other. To accurately obtain the electrostatics we assume the molecule forms a self-assembled monolayer (SAM) providing a uniform dielectric layer between the STM+vacuum and the Si(100) substrate with a dielectric constant of 1.26.[13] Based on MEDICI[14] with two dielectric layers (one for molecular monolayer and another for vacuum) between silicon and a metal (STM), we determine the electrostatic boundary conditions. Since the total potential applied (V_{total}) from the STM tip drops across three different media: (a) the silicon surface, bending the bands at the surface with respect to bulk (V_{si}), (b) the molecule itself (V_{mol}), (c) the vacuum gap between the STM tip and the molecule (V_{vac}), we can write, $V_{\mathrm{total}} = V_{si} + V_{\mathrm{mol}} + V_{\mathrm{vac}}$. In order to observe molecule-mediated NDR we postulated[4] that the molecular levels be mobile with respect to the bulk band edge of the heavily doped Si(100) substrate. Here we define a physical parameter η that gives one an idea about how fast the levels move with respect to the bulk band edge. The potential on the molecular level is assumed to be the average of the potentials at the two edges of the molecule, in other words we do not take into account the Stark effect on the molecule for this definition.

$$\eta = \frac{V_{si} + \frac{V_{\mathrm{mol}}}{2}}{V_{\mathrm{total}}}. \quad (2)$$

Figure 3 shows the variation of η as a function of the vacuum gap. The nonlinearity in the variation of η comes from the finite nonlinear band bending at the silicon surface. The doping level of Si(100) surface was assumed to be 4×10^{19} cm^{-3} in the calculation.[3] As we will explain in the next few sections η represents the single most important experimental knob that we can tinker with to understand the physics behind the NDR peaks. As shown in Fig. 3 moving the STM tip changes η, reducing it as the tip is retracted. This has a profound effect on the I-V as explained in Sec. II C.

One point before we move on: We define η to give us a metric to estimate level movement with respect to the silicon bulk band edges. However, in our self-consistent solution a linear potential drop inside the molecule is only the initial boundary condition for a more detailed Poisson solution through GAUSSIAN98. Because of self-consistency, the potential inside the molecule develops to be highly nonlinear (as shown later in Fig. 12). The effect of this self-consistency is pronounced in the I-V characteristics.[15]

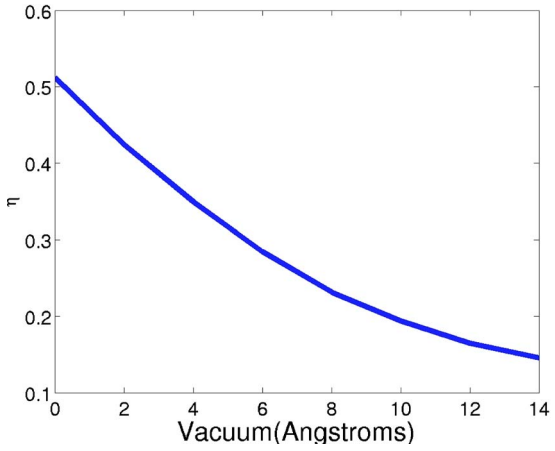

FIG. 3. (Color online) η [as defined in Eq. (2)] vs vacuum gap. As the STM tip is retracted, less applied potential drops across the molecule reducing η. This theoretical prediction is in strong agreement with experimentally extracted η.[10]

C. Effect of η on I-V

To illustrate the effect of η on the I-V, we employ the elementary model described earlier. In the elementary model, we shift the molecular levels by a fraction η of applied bias as described earlier and neglect the Stark effect (which is included in the first-principles solution). Depending on the distance of the STM tip from the molecule, from almost all to almost none of the applied bias could drop across the molecule. η is expected to be close to 0.5 for vanishingly small vacuum gap, i.e., the STM tip almost touching the molecule (magenta curve in Fig. 4). As the STM is moved farther away from the molecule, η decreases progressively as shown in Fig. 3. The effect on the I-V is twofold: (i) *electrostatic*, the NDR peak is delayed as η decreases, and (ii) *quantum*, the current level drops because of increased tunneling distance. However, as we will point out in Sec. II D, the peaks move only if they are molecular in nature. If the NDR is a contact-related effect ($\eta \approx 1$) the peak positions are expected to be rather insensitive to STM movement.

D. Equilibrium band lineup: Silicon-molecule-metal

In this section, we explain the equilibrium band alignment of a silicon-molecule-metal system. To draw the equilibrium band diagram (Fig. 5), we first place all the relevant levels (Si band edges, highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) levels) with respect to a common vacuum level. However, a work-function difference exists between the STM tip (tungsten, 4.6 eV) and the silicon substrate even at zero applied bias (work function of the silicon substrate depends on the type and degree of doping). In order to satisfy the requirement of a single Fermi energy at equilibrium, the vacuum levels bend and the molecular levels and the silicon surface potential move up and down by an amount determined by the potential drop in the intervening media, as explained in Fig. 3. For an n-type substrate, the vacuum level (E_{vac}) bends up at the silicon surface, the molecule, and the tunnel gap, whereas for p-type silicon, E_{vac} bends down. The molecular levels move up and down by an amount $\eta \times (E_{f,\text{STM}} - E_{f,si})$, as explained previously.

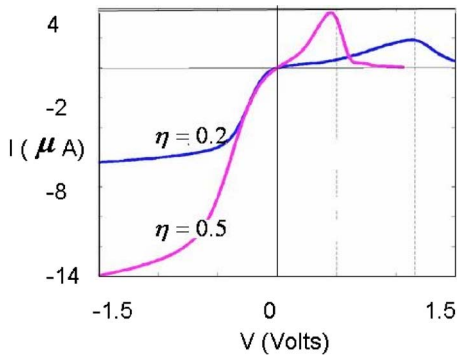

FIG. 4. (Color online) Effect of STM retraction on the I-V. The NDR peak shifts and the current level reduces. The molecule is represented using only two levels (one occupied and one unoccupied).

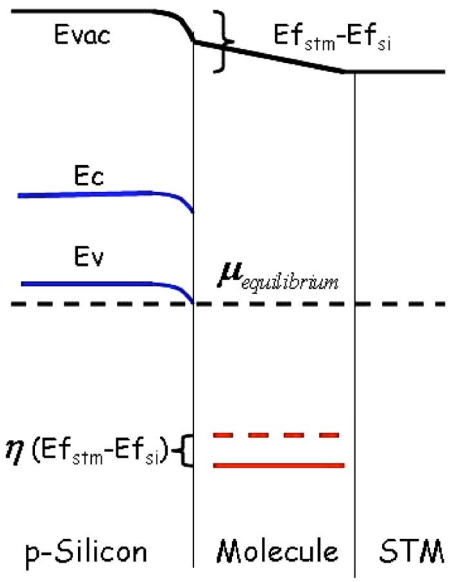

FIG. 5. (Color online) Equilibrium band lineup. The work-function difference between the tip and the Si surface gives rise to an electric field at equilibrium that bends the absolute vacuum levels as shown. As a result the molecular levels move in energy.

III. DETAILED METHOD

A. Molecular Hamiltonian: *Ab initio*

In this section we move on from a simple 2×2 model Hamiltonian to an appropriate first-principles Hamiltonian based on DFT with a 6-31$g(d)$ basis set. To describe the molecule from first principles, there exists a plethora of SCF schemes to choose from. The different schemes differ in the way they treat the one-electron potential U_{SCF}. From a transport point of view, we believe that local density approximation (LDA) within the density-functional theory (DFT) helps provide an appropriate SCF description.[16] Hence, we model TEMPO based on LDA and 6-31$g(d)$ basis set. In SCF calculations, the single-particle eigenenergies are interpreted as the energy required to extract an electron (HOMO) and the energy required to inject an electron (LUMO). HOMO-

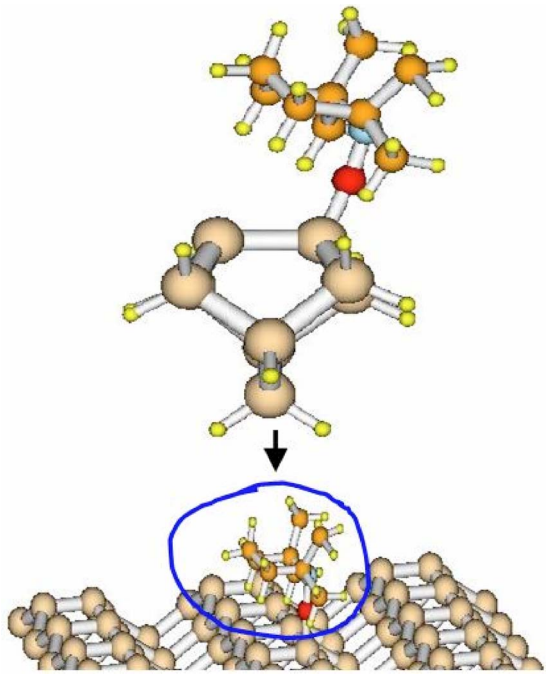

FIG. 6. (Color online) TEMPO sitting on the bare 2×1 reconstructed Si(100) surface.

LUMO gaps obtained from LDA match the numbers obtained from electrochemical measurements[17] quite well.

The only adjustable parameter that we use in our first-principles calculation is the equilibrium position of the molecular levels with respect to the equilibrium E_f. The precise location of E_f can depend very sensitively on many factors, such as surface conditions, geometrical factors, and presence of impurities.[18] In the absence of detailed information of all these factors in a particular experimental setup, it is necessary to treat the position of equilibrium E_f as adjustable.[19] The E_f is adjusted by surrounding the molecule with point charges with appropriate polarity and magnitude.

B. Structural optimization: Molecule on silicon surface

The first step in carrying out a detailed calculation of current flow through a molecule adsorbed onto silicon is to understand how the molecule bonds to the silicon surface. The nature of bonding can affect the device Hamiltonian, which, in turn, may affect the final I-V characteristics. The molecular bonding at the surface is particularly complicated for silicon because the surface can come in different flavors that depend on the nature of surface terminaton. In this paper, we are going to focus only on the Si(100) surface. Although the Si(100) surface has been found to exhibit a large number of reconstructions, it is generally believed to assume a 2×1 reconstruction at room temperature, with the Si—Si dimer bond is 2.2 Å long and adjacent dimers are 3.8 Å apart.[20] We consider a TEMPO molecule that binds with one of the silicon atoms of the dimer as shown in Fig. 6. It is the terminal oxygen atom of TEMPO that forms a covalent bond with one of the silicon surface atoms. There are other molecules, for example, cyclopentene (C_5H_8), which bond to silicon via a strong silicon-carbon tether.

The main purpose of a detailed optimization is to model the oxygen and carbon to silicon bonding at the surface and the resulting rearrangement, if any, of the molecule and the adjacent silicon surface. We start with a silicon cluster that has one surface dimer and extends three layers into the bulk. The bulk Si atoms are terminated with hydrogen atoms that have approximately the same electronegativity as Si atoms. To optimize the geometry, the hydrogen atoms are usually kept fixed to mimic the correct behavior of bulk silicon, otherwise the cluster shrinks too much. First, this bare silicon cluster is optimized using the B3LYP method and 6-31$g(d)$ basis set. The optimization yields an asymmetric dimer at the surface with a dimer bond length of 2.2 Å that compares favorably with published results.[21] Next, the TEMPO molecule is placed on top of this silicon cluster such that the terminating oxygen on TEMPO forms a bond with one of the surface dimer atoms. This structure is optimized using B3LYP and 6-31$g(d)$. The resultant configuration is used to construct the extended device Hamiltonian, which consists of TEMPO, and one silicon surface atom that forms a bond with TEMPO.

C. Self-energy-Σ_1: Silicon(100)

Connecting a molecule seamlessly to a silicon substrate is a rather challenging theoretical task. An isolated molecule is quite accurately described by *ab initio* schemes[22] based on nonorthogonal basis sets. On the other hand, silicon bulk is usually modeled employing a different class of basis sets and schemes, from empirical tight-binding[23] to *ab initio*.[24] However, *ab initio* schemes and basis sets that describe an isolated molecule adequately may not describe bulk properties of silicon as well. Because the physics of NDR lies in molecular levels moving past the band edge, it is essential that our model estimates bulk silicon properties, such as the correct band gap, and describes the Si(100) surface around the band gap effectively. Once we have obtained a reasonable surface Green's function for silicon, we need to convert its effect onto an *ab initio* basis set that describes the molecule. This is of paramount importance to us since the charge transfer at the interface depends sensitively on what the coupling is at the silicon-molecule junction. Therefore, our effort to model silicon has two goals: (i) obtain the correct band-structure and surface properties of silicon and (ii) couple this silicon to a molecule described from first principles.

We take two approaches to describe and couple silicon to the molecule. The first one is an elementary two-band model describing the conduction and valence bands at a parabolic density-of-states (DOS) effective mass level based on a finite element grid. This description, although admittedly simplistic, yields a rather straightforward coupling to an *ab initio* basis set that describes the molecule. To couple silicon to the molecule, we directly calculate the overlap between a P_1 finite element shape function and the 6-31$g(d)$ basis set that describes the molecule (Fig. 7). Following,[25] we write the jellium Green's function as

$$G(x,y;x,y') = -2\sum_{k_m} \frac{\sin k_m x}{\hbar v_m} e^{ik_m x} \chi_m(y) \chi_m^*(y'), \qquad (3)$$

where $k_m = \sqrt{2mE/\hbar^2 - k_y^2}$, $v_m = \hbar k_m / m$, and $\chi_m(y) = \exp(ik_m y)/\sqrt{L}$. The self-energy ($\Sigma$) expression is

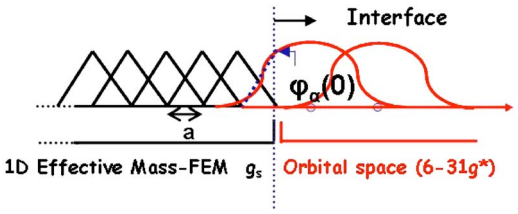

FIG. 7. (Color online) Coupling finite element method (FEM) basis to *ab initio* 6-31$g(d)$ basis set.

$$\Sigma_{\alpha\beta} = \left(\frac{\hbar^2}{2m}\right)^2 \int dy dy' \phi_\alpha(0,y)\phi_\beta^*(0,y')\frac{G(b,y;b,y')}{b^2}. \tag{4}$$

When discretized based on a 1D uniform finite element grid the above reduces to

$$G = -2\frac{\sin kx}{\hbar^2 k_m/m}e^{ikx}, \tag{5}$$

so that

$$\Sigma = -\frac{\hbar^2}{2mb} - i\frac{\hbar^2 k}{2m}. \tag{6}$$

Evidently, we recover the well-known one-dimensional (1D) tight-binding expression for Σ.

The second approach is more rigorous. We simulate semi-infinite Si bulk properties using extend Huckel theory (EHT) with optimized parameters, which provide good Si bulk band structure.[26] Then we connect the silicon surface (that includes four relaxed layers) according to the optimized geometry as mentioned in Ref. 27. The surface density of states is calculated using Green's function and is in good agreement with STM spectra as shown in Fig. 8. Then we couple the physicists' silicon to a chemists' molecule as explained in detail in Ref. 28. Although we recognize that far more accurate band-structure calculations could be there for silicon with more desirable attributes, we think at this time it is crucial to adopt a minimal, transferrable model for the contacts that can interface with a molecule and, nevertheless, does justice to the geometry, bulk band structure, and surface states. For the results shown in Sec. IV, we have used the former approach to connect TEMPO to silicon.

Silicon not only acts as a bulk electrode, the neighboring surface states of silicon [π and π^* for a Si(100)] interact with the molecule by exchanging charge as observed in experiments.[13] In our model we include this effect based on an energy-independent Buttiker probe, which does not draw any current from the molecular device.

D. Self-energy-Σ_2: STM and vacuum gap

One added complication in any STM measurement is the presence of a vacuum gap between the sample and the STM tip. This affects the I-V characteristics in two ways: (i) *quantum*, rate of electron tunneling through the barrier posed by the gap and (ii) *electrostatic*, part of the applied bias drops across the gap, affecting the movement of levels through the Laplace part (as explained in Sec. II).

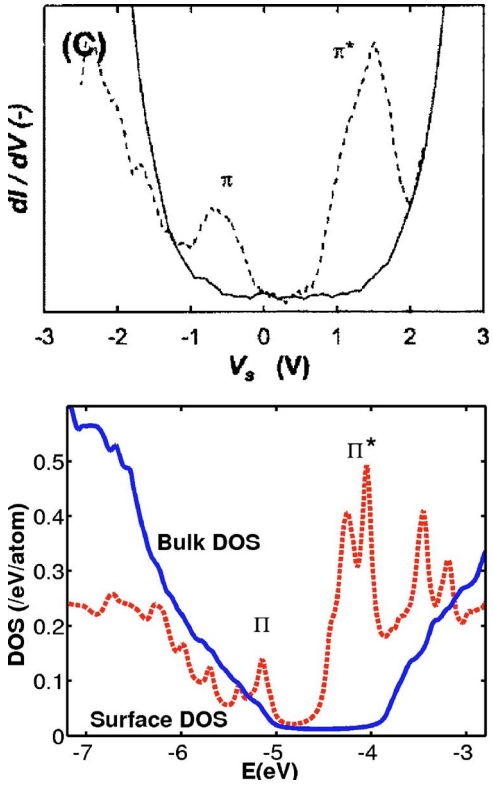

FIG. 8. (Color online) Si surface states Si(100) and bulk Si: experiment (top) and theory (bottom). Top: Dashed curve shows STM image of Si(100) surface without a molecule showing prominent π and π^* states. Solid curve depicts the image when a cyclopentene molecule is bonded at the surface, releasing the band edge (Ref. 34) (reproduced with permission from the American Physical Society). Bottom: Dashed curve shows the calculated Si(100) surface DOS, and solid one shows total bulk DOS.

When it comes to tip-sample separation, we observe that there can be two limits: (i) if the tip is reasonably far away from the molecule, then the vacuum in between the molecule and the tip acts as an energy filter to the traveling electrons. Electrons traveling from the conduction band encounter a lower barrier compared to the electrons traveling from the valence band. This can be approximated using a Wentzel-Kramers-Brillouin (WKB) type function with an appropriate barrier height. (ii) When the tip is very close to the sample, in the range of a bondlength between the molecule and the tip, a physical collapse of the tunnel barrier accompanies the transition from tunneling to direct-contact regime. In this limit, we model tip-sample interaction by evaluating the direct coupling matrix element between the end carbon atom of TEMPO, and a single tungsten atom assumed to be sitting at the apex of the STM tip.[29] Because of a lack of experimental knowledge of precise STM tip structure, we ignore any structural effect of the tip.

IV. RESULTS

A. NDR on *p*-type Si(100) substrate

Now that we have explained how we treat each of the different entities in the solution block diagram (Fig. 2), we

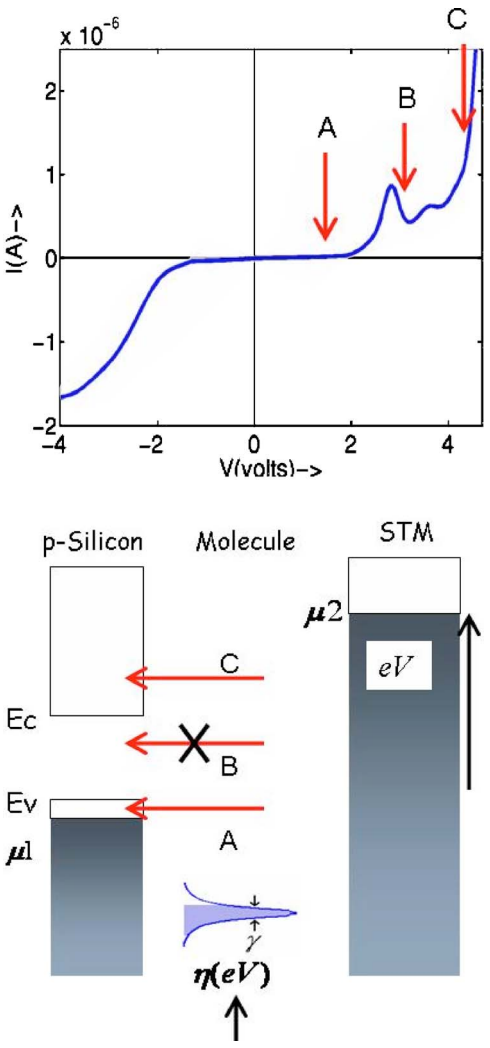

FIG. 9. (Color online) Fully self-consistent I-V based on LDA/6–31$g(d)$ description of TEMPO on p++type Si(100) substrate (top). The I-V shows clear polarity reversal, showing NDR in the positive bias direction and not in the negative. The I-V in the NDR direction shows three distinct regions, labeled A, B, and C. The corresponding locations of the HOMO level with respect to the bulk band edge are depicted in the schematic (bottom).

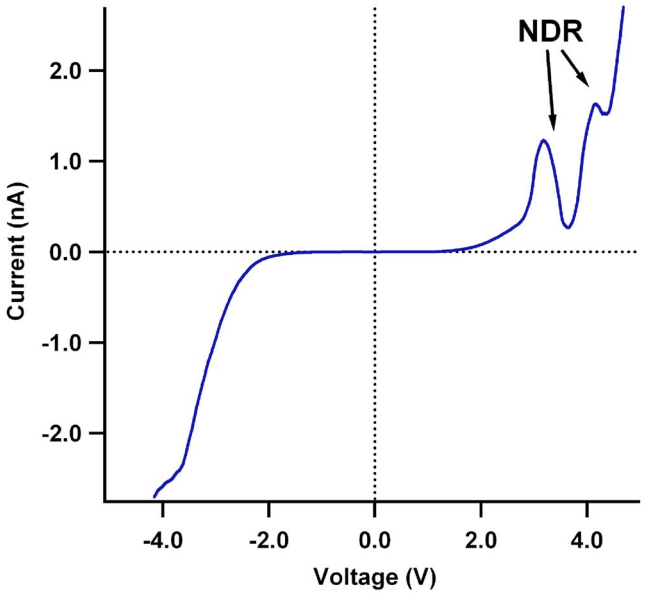

FIG. 10. (Color online) Experimental I-V showing one-sided NDR through TEMPO on p++ Si(100).

are ready to discuss in detail the results obtained. We couple the nonequilibrium Green's function (NEGF) formulation of transport[25] with an LDA molecular Hamiltonian of TEMPO calculated based on Gaussian98[22] using a 6-31$g(d)$ basis set as outlined in Sec. III. Current (I) through the semiconductor-molecule-metal junction is calculated from Eqs. (1). The current thereby calculated includes contributions both from thermionic emission over the semiconductor surface barrier, as well as quantum-mechanical tunneling through it.

I-V obtained from a detailed *ab initio* calculation [Fig. 9 (top)] for TEMPO on p-type substrate shows clear polarity reversal in NDR peaks, appearing for positive substrate bias and not for negative (Fig. 10). The only fitting parameter in our calculation is the location of the molecular levels with respect to equilibrium E_f, and, in this case, the levels are moved upward (in energy) by ≈250 meV from their self-consistent equilibrium location. We obtain one clear first peak and an attenuated second one in the positive bias direction. The peaks correspond to the HOMO and HOMO-1 levels of the molecule moving past the silicon valence band edge (E_v) into the band gap. We identify three distinct regions of the I-V in the NDR bias direction and label them as A, B, and C. The corresponding position of a molecular level with respect to the silicon bulk band edge is illustrated using arrows in the schematic [Fig. 9 (bottom)]. There is very little current flow when the molecular level is not within the bias window, but approaching it. Current starts flowing as the tail of a level enters the bias window between μ_1 and μ_2 (position A). Current keeps increasing until conduction is cut off as the level moves past the bulk band edge and enters the band-gap region (position B). We observe a clear and rapid increase in current once the bias crosses the peaks (position C). This occurs because tunneling into the conduction band overwhelms the current that is obtained from the valence band.

What determines the location of the peak in the I-V? Is it (i) how far the relevant molecular level (level contributing to NDR) is from the equlibrium E_f of the metal-molecule-silicon structure or (ii) how fast the molecular level moves with respect to the silicon bulk band edge? This last part has two components to it: (a) Laplace—the Laplace solution decides how much voltage drops across the molecule as outlined in Sec. II. This, in turn, decides how fast the molecular level moves with respect to the silicon bulk band edge (Fig. 3). (b) Self-consistent charging—a plot of the movement of the molecular energy levels[15] clearly indicates that the level movement is slower than what is expected from a simple Laplace answer. This is because of the self-consistent charging effect discussed in detail in Ref. 15. In this calculation we ignore any electrostatic image corrections due to the proximity of the contacts.

Now that we have described what determines the location of the I-V peaks, let us focus our attention on the magnitude

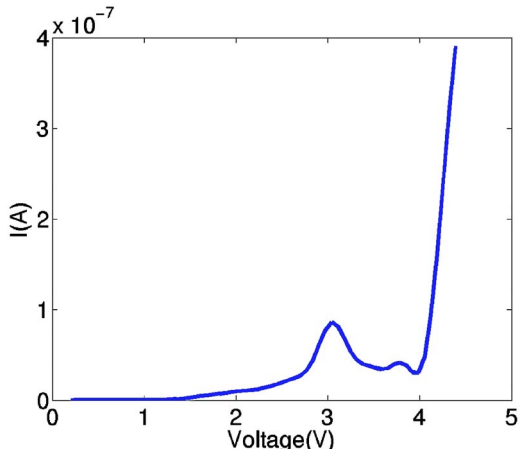

FIG. 11. (Color online) I-V with lower doping, which produces a gap of 10 meV between the work function and valence band edge of Si(100) compared to 50 meV in the previous *ab initio* I-V (Fig. 9). The STM was kept fixed at the same distance as in the case of the previous I-V plot.

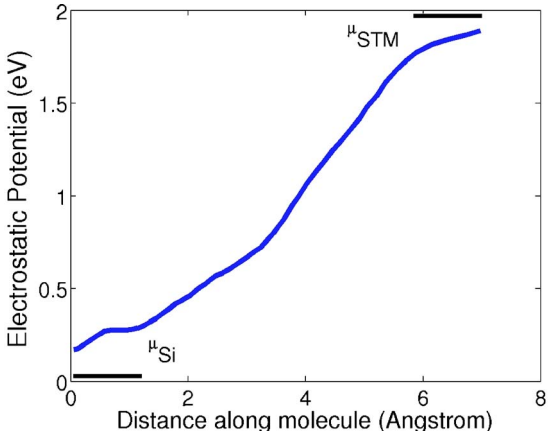

FIG. 12. (Color online) Self-consistent potential profile that develops inside the molecule. Note that some voltage is dropped at the molecule-electrode interface; hence, less than the applied 2 V drops inside the molecule.

of current. Since the exact distance of the STM tip from the molecule is an unknown,[30] in our theoretical calculations we assumed that the STM is almost touching the molecule. This configuration provides the best possible slippage of the molecular levels (Fig. 3) accentuating the NDR peak. However, moving the STM tip outward will move the location of the NDR peak. In that case, we argue that the equilibrium location of the molecular levels with respect to the Si(100) E_f is closer than what is deduced from equilibrium band lineup and charge transfer. As pointed out earlier, this shift is not unexpected given that the exact experimental conditions of the interface and the surface is not clearly known.

One additional conclusion from our calculations is that the magnitude of current in this structure is *not* solely determined by the STM tip, as is the case for a metallic substrate and as described in classical STM literature.[31] At the NDR point, current magnitude depends on *both* how far the STM tip is from the molecule and the doping level of the substrate. The effect of substrate doping is to move the silicon E_f with respect to the band edges. This, in turn, limits the number of electrons that are injected from the silicon substrate contributing to NDR. For example, we predict that for the same STM distance an increase (decrease) in doping would increase (decrease) the current level significantly at the NDR point (Fig. 11). However, for negative bias direction the magnitude of current is expected to be determined mostly by the tip-molecule distance, since one is away from the Si band edges in that case. Initial experiments seem to agree reasonably well with the above conclusions.[10]

In summary, the *magnitude* of NDR peak current depends on (i) the doping level of the Si(100) substrate and (ii) how far the STM tip is from the molecule. The peak position depends on (i) the equilibrium location of the relevant level with respect to the Si(100) E_f and (ii) how fast the level slips past the band edge. Hence, given experimental tip-molecule distances and thus estimates of voltage drop on the molecule (Fig. 3), one can extrapolate the equilibrium position of the HOMO of a molecule with respect to the Si(100) E_f. This will facilitate electronic characterization of molecular HOMO levels of different molecules with respect to the technologically important Si(100) surface.

B. Absence of NDR on bare Si(100) substrate

Until now we have explained both from a simple band diagram and a detailed first-principles solution that a molecular level slipping past the silicon band edge can give rise to NDR on *p*-type silicon. One good question to ask is how do we make sure that it is indeed the interaction of the molecular level and the silicon band edge that produces NDR? The answer to that would be the absence of NDR on a silicon surface without a molecule. And that is indeed the case when we calculate current on a bare Si(100) surface both for *n*- and *p* types. In experiments, too, there is no evidence of NDR on bare Si(100) substrates.[3]

The current calculated on the bare *n*-type Si surface is shown in Fig. 13. We represent vacuum between the tip and surface based on a finite difference grid. The total current from the surface to the tip varies with (i) distance between the STM and the surface (ii) the effective area of the STM tip, and (iii) the barrier height in vacuum. Since there is no

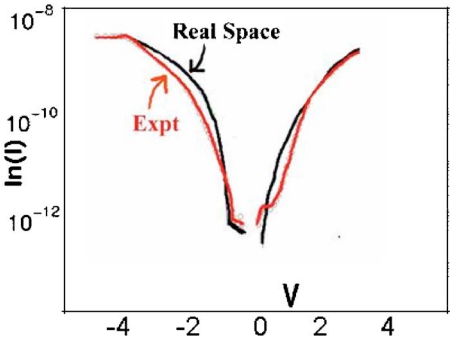

FIG. 13. (Color online) I-V: Bare Si(100) that shows no evidence of NDR. Current was calculated through vacuum based on a real-space finite difference grid.

charge in vacuum we assume a linear potential drop inside it. The initial barrier height posed by vacuum is ≈ the work function of the STM tip (≈4.6 eV). We obtain a good match with experimental current levels for an area ≈1 nm^2 and distance=1 nm for a 1D problem.

V. DISCUSSION: CONNECTION TO EXPERIMENTS AND POSSIBLE SCENARIOS OF NDR ON *n*-TYPE

In Sec. IV we showed that we can reproduce the experimental location and polarity of NDR peaks on *p*-type substrates starting from first principles. However, from a simple band diagram, it is straightforward to see that it would require enormously high biases to observe NDR through molecules on *n*-type Si(100) since the equilibrium position of the *unoccupied* levels are far removed from the conduction band edge (E_c). Experimentally though,[11] NDR is observed through cyclopentene docked onto n++Si(100) substrates at 300 K. For NDR to occur due to the cyclopentene LUMO, it needs to be ≈0.9 V away from E_c based on a realistic level movement scenario.[32] However, theoretically the cyclopentene LUMO (based on LDA) is located 3.0 V away from E_c. Hence, theoretically NDR peaks are expected to occur at 8.5 V. The numbers are far worse in the case of TEMPO, where peaks are observed in experiments[3] at sufficiently low biases as well. In this section we suggest some physics that may help explain the early occurence of NDR on *n*-type substrates. Furthermore, we propose experiments that may help ascertain which of these effects is really playing a role to produce NDR on *n*-type silicon at room temperature.

A. Early appearance of peaks: STM adsorbate effects

In this section we analyze the effect of adsorbates sticking to the STM tip on the I-V characteristics. The adsorbate can be viewed as another molecule, the identity of which is as yet unknown. Nonetheless, this adsorbate can be expected to have discrete energy levels broadened by the tip density of states. Once a bias is applied from the tip, the adsorbate levels move with an electrostatic speed $\eta \approx 1$. This, in turn, ensures that an energy level (not from the molecule but from the adsorbate) comes into the bias window early enough to cause current conduction. As these level(s) move past the band edge and into the silicon band gap, conduction is cut off giving rise to NDR. One good question to ask at this point is why should we not see any NDR on bare Si(100) since the adsorbate level is there even if the molecule is not? The answer lies in the fact that bare Si(100) surface has surface states that swamp the band edges.[33] However, when a molecule bonds chemically to the Si(100) surface, the bond locally quenches the dangling bonds of silicon surface releasing the band edge.[34] This allows conduction cutoff for an adsorbate or molecular level but not for a bare surface.

We suggest one test to determine experimentally whether, indeed, NDR occurs because of the adsorbate. We postulate that as one retracts the STM tip, the NDR should show up at the same or very close to the original bias. This is expected since the adsorbate-induced level is tied to the STM tip and moves at $\eta \approx 1$ with respect to silicon. If the position of the NDR peak moves significantly as one retracts the STM tip, the level contributing to NDR is not adsorbate or contact related.

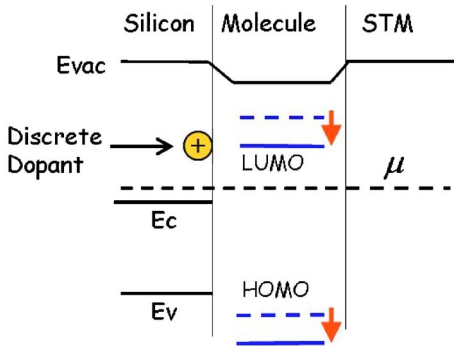

FIG. 14. (Color online) Dopant schematic: The presence of uncovered charge at the Si surface can create a potential on the molecule that may help lower (*n*-type) or raise (*p*-type) molecular levels with respect to the bulk band edge, thus expediting observation of NDR.

We would also like to point out that the aforementioned mechanism retains the polarity dependence of the observed NDR. In this regard it is different from the mechanism pointed out by Lyo and Avouris[35] when NDR shows up in no specific bias direction, depending on the type of doping.

B. Earlier appearance of peaks: Surface dopant effect

It is conceivable that due to dopant out-segregation the concentration of dopants at the Si(100) surface may be higher than in the bulk. This may have a profound electrostatic effect on the molecular energy levels. Imagine a donor dopant ion sitting at the surface. It creates a negative potential thereby lowering the molecular energy levels with respect to the silicon band edge, as shown schematically in Fig. 14. The question is whether the levels can be pulled down by as much as 5–6 V or not. In the early days of semiconductors, though, threshold voltage shifts of the order of 20 V, due to sodium ion impurities were not unusual.[18] Hence in the early days of silicon-based molecular electronics it may not be terribly far-fetched to imagine a shift of the order of several volts for molecular levels due to random dopant ions. There can be two different ways that the dopants can affect the band lineup at the molecular interface: (i) globally altering the lineup or (ii) locally changing the interface lineup right around a single molecule. Evidence for (i) would be the advent of Fowler-Nordheim tunneling[36] for the I-V on bare silicon, when the vacuum level between the surface and the STM tip can be pulled down globally by several volts thereby dramatically increasing the rate of tunneling. One point before we move on, the electrostatics of surface dopant ions helps lower levels for *n*-type (donor) and raise levels for *p*-type (acceptor) substrates, both of which would expedite observation of NDR.

C. Multiple peaks: Inelastic phonons

In experimental I-V curves, particularly on the *n*-type substrates, multiple NDR peaks are observed. This may be

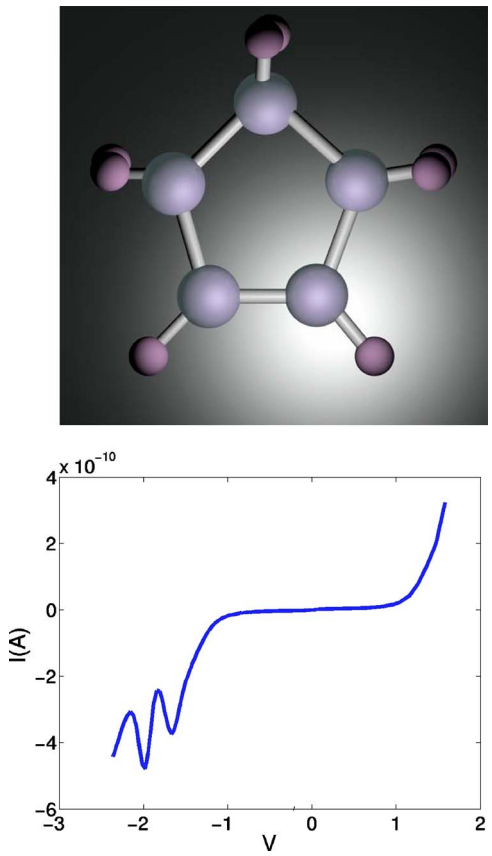

FIG. 15. (Color online) Cyclopentene molecule (top) and IV through cyclopentene, including inelastic scattering. Cyclopentene is represented with one orbital per site, and scattering is included within the first-Born approximation based on NEGF.

caused by multiple electronic levels crossing the band edge or one electronic level and its phonon sidebands crossing the band edge. The difficulty in adding one or more electrons (due to single-electron charging and structural deformation) into unoccupied level(s) prompts us to explore the second option.

To calculate an I-V that is affected by phonons starting from NEGF equations we need (i) a Hamiltonian, (ii) the dominant phonon frequencies, and (iii) the deformation potential developed due to the relevant phonon modes. We use cyclopentene as a prototype molecule in our calculations because the phonon modes of cyclopentene on silicon surfaces are well documented. For the Hamiltonian, we construct a simple explanatory one where cyclopentene is represented at a one-orbital per carbon atom site orthogonal basis set. The eigenenergy levels are parametrized so that the NDR peaks show up at the locations as seen in experiments.[11]

The phonon modes of cyclopentene bonded onto Si(100) surface are well known from high-resolution electron-energy-loss spectroscopy (HREELS) experiments.[37] The dominant one is 96 meV. The deformation potential due to the phonon modes on the molecular levels is estimated as follows: We displace the molecular coordinates by a small amount and calculate the resulting change in the electronic levels based on Gaussian98.[38] Then we calculate the ratio of the phonon mode and the resultant shift in energy level and linearly extrapolate the deformation potential that corresponds to the particular phonon mode. Next we calculate current through the molecule within the first Born approximation based on NEGF.[12] We evaluate an inscattering matrix $\Gamma_s^{\text{in}}(E)$ and outscattering matrix $\Gamma_s^{\text{out}}(E)$ from

$$\Gamma_s^{\text{in}}(E) = \int (d(\hbar\omega))D^{ph}(\hbar\omega)n(E+\hbar\omega),$$
$$\Gamma_s^{\text{out}}(E) = \int (d(\hbar\omega))D^{ph}(\hbar\omega)p(E-\hbar\omega), \tag{7}$$

where the phonon spectral function can be written as the sum of an emission term and an absorption term

$$D^{ph}(\hbar\omega) = \Sigma_i D_i[(N_i+1)\delta(\hbar\omega-\hbar\omega_i) + N_i\delta(\hbar\omega+\hbar\omega_i)], \tag{8}$$

with N_i representing the number of phonons of frequency $\hbar\omega_i$ and D_i its coupling. We assume N_i to be given by the Bose-Einstein factor, but it is conceivable that the phonons could be driven off equilibrium, requiring N_i to be evaluated from a transport equation for phonons. Low-frequency phonons with $\hbar\omega_i$ much smaller than other relevant energy scales can be treated as elastic scatterers with $\hbar\omega_i \approx 0$, $D_i(N_i+1) \approx D_i N_i \equiv D_0^{ph}$. From Eqs. (8) then, we can write

$$\Gamma_s^{\text{in}}(E) = D_0^{ph} n(E),$$
$$\Gamma_s^{\text{out}}(E) = D_0^{ph} p(E), \tag{9}$$

so that $\Gamma_s = \Gamma_s^{\text{in}} + \Gamma_s^{\text{out}} = D_0^{ph} D(E)$. This Γ_s can now be used to evaluate the correlation function G^n that decides what $n(E)$ and $p(E)$ are. From a knowledge of $n(E)$ and $p(E)$, we can recalculate what Γ_s is and we keep on doing this until we reach convergence. Once convergence is achieved, I-V is calculated. The resultant I-V shows multiple peaks, the first one coming from the original unoccupied electronic level and the following ones originating due to the phonon sidebands. We conclude that phonon sidebands from phonon emission can give rise to multiple peaks in the near vicinity of the main electronic level. However, it does not help shift the unoccupied levels downward. Hence there ought to be a negative shift in the molecular energy levels from a different source, namely, possibilities (i) and (ii). The calculation was done assuming phonon equilibrium;[12] however, there may be phonon absorption peaks that will produce NDRs that would precede a peak due to an electronic level.[39–43]

VI. CONCLUSION

In this paper we present a detailed first-principles solution procedure based on DFT and NEGF to analyze I-V through molecules on silicon. We obtain polarity-dependent NDR for I-Vs through TEMPO on *p*-type silicon substrates. With more experimental data on NDR-peak position as a function of STM tip distance, it should be possible to map out the location of the HOMO level with respect to the *p*-type Si(100) band edge. Based on theoretical estimates, we predict that high biases would be required to obtain similar

polarity-dependent I-Vs on *n*-type silicon. However, in experiments on *n*-silicon, the peaks show up earlier than expected. Hence it seems likely that other physics help produce this expedited NDRs on *n*-silicon. We point to two such possibilities and further suggest experiments that would help ascertain the cause of such NDRs.

ACKNOWLEDGMENTS

We acknowledge N. P. Guisinger, K. Bevan, and L. Siddique for many useful discussions. This research was funded by ARO-DURINT, NSF, NASA Institute for Nanoelectronics and Computing, and MARCO Focus Center.

[1] M. DiVentra, S. T. Pantelides, and N. D. Lang, Phys. Rev. Lett. **84**, 979 (2000).
[2] J. Chen and M. A. Reed, Chem. Phys. **281**, 187 (2002).
[3] N. P. Guisinger, M. E. Greene, and R. Basu, A. S. Baluch, and M. C. Hersam, Nano Lett. **4**, 55 (2004).
[4] T. Rakshit, G.-C. Liang, A. W. Ghosh, and S. Datta, Nano Lett. **4**(10), 1803 (2004).
[5] K. Walzer *et al.*, Surf. Sci. **595**, 532 (2003).
[6] S. Lenfant, C. Krzeminski, C. Delerue, G. Allan, and D. Vuillaume, Nano Lett. **741**, 3 (2003).
[7] M. C. Hersam and R. Reifenberger, MRS Bull. **29**, 385 (2004).
[8] P. S. Damle, A. W. Ghosh, and S. Datta, Phys. Rev. B **64**, 201403 (2001).
[9] R. A. Wolkow, Jpn. J. Appl. Phys., Part 1 **40**, 4378 (2001).
[10] N. P. Guisinger, N. L. Yoder, and M. C. Hersam, Proc. Natl. Acad. Sci. U.S.A. **102**, 8838 (2005).
[11] N. P. Guisinger, R. Basu, M. E. Greene, A. S. Baluch, and M. C. Hersam, Nanotechnology **15**, S452 (2004).
[12] S. Datta, Nanotechnology **15**, S-433 (2004).
[13] K. Hata, Y. Shibata, and H. Shingekawa, Phys. Rev. B **64**, 235310 (2001).
[14] Technology Modeling Associates, Inc., Sunnyvale, CA, *TMA Medici, Two-Dimensional Device Simulation Program, version 4.0 User's Manual* (1997).
[15] T. Rakshit, G.-C. Liang, A. W. Ghosh, and S. Datta, J. Comput. Electron. (to be published).
[16] T. Rakshit and S. Datta (unpublished).
[17] H. Lund *et al.*, *Organic Electrochemistry* (M. Dekker Inc., 1990).
[18] R. Pierret, *Semiconductor Device Fundamentals* (Addison-Wesley, Reading, MA, 1991).
[19] In fact, in the early days of semiconductor industry a threshold shift of the order of 20 V in the capacitance-voltage (C-V) measurements was not unusual.
[20] G. P. Srivastava, *Theoretical Modeling of Semiconductor Surfaces* (World Scientific, Singapore, 1998).
[21] Q. Liu, J. Am. Chem. Soc. **4082**, 117 (1985).
[22] M. J. Frisch *et al.*, *Gaussian '98, Revision A.7* (Gaussian, Inc., Pittsburgh, 1998).
[23] T. Boykin, G. Klimeck, and F. Oyafuso, Phys. Rev. B **69**, 115201 (2004).
[24] M. Stadele, M. Moukara, J. A. Majewski, P. Vogl, and A. Gorling, Phys. Rev. B **59**, 10031 (1999).
[25] S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, England, 1995).
[26] J. Cerda and F. Soria, Phys. Rev. B **61**, 7965 (2000).
[27] A. Ramstad, G. Brocks, and P. J. Kelly, Phys. Rev. B **51**, 14504 (1995).
[28] A. W. Ghosh (unpublished).
[29] J. M. Blanco, C. Gonzalez, P. Jelinek, J. Ortega, F. Flores, and R. Perez, Phys. Rev. B **70**, 085405 (2004).
[30] S. Datta, W. D. Tian, S. H. Hong, R. Reifenberger, J. I. Henderson, and C. P. Kubiak, Phys. Rev. Lett. **79**, 2530 (1997).
[31] C. J. Chen, *Introduction to Scanning Tunneling Microscopy*, Oxford Series in Optical and Imaging Sciences Vol. 1 (Oxford, London, 1993).
[32] Based on experimental observation we estimate that the STM tip is roughly 4–6 Å. That corresponds to an $\eta \approx 0.35$.
[33] M. McEllistrem, G. Haase, D. Chen, and R. J. Hamers, Phys. Rev. Lett. **70**, 2471 (1993).
[34] R. Akiyama, T. Matsumoto, and T. Kawai, Phys. Rev. B **62**, 2034 (2000).
[35] I.-W. Lyo and P. Avouris, Science **245**, 1369 (1989).
[36] Y. Taur, *Semiconductor Device Fundamentals* (Addison-Wesley, Reading, MA, 1991).
[37] S. Machida, K. Hamaguchi, M. Nagao, F. Yasui, K. Mukai, Y. Yamashita, J. Yoshinobu, H. S. Kato, H. Okuyama, and M. Kawai, J. Phys. Chem. B **1691**, 106 (2002).
[38] K. Bevan (private communication).
[39] G. P. Lopinski, D. D. M. Wayner, and R. A. Wolkow, Science **406**, 48 (2000).
[40] M. Brandbyge, J. L. Mozos, P. Ordejon, J. Taylor, and K. Stokbro, Phys. Rev. B **65**, 165401 (2002).
[41] S. Datta, *Quantum Transport: Atom to Transistor* (Cambridge University Press, England, 2005).
[42] G. C. Liang and A. W. Ghosh (unpublished).
[43] J. S. Hovis, H. Liu, and R. J. Hamers, Appl. Phys. A: Mater. Sci. Process. **S553**, 66 (1998).