

Theory of nonequilibrium single-electron dynamics in STM imaging of dangling bonds on a hydrogenated silicon surface

Lucian Livadaru,* Jason Pitters, Marco Taucer, and Robert A. Wolkow

Department of Physics, University of Alberta & National Institute for Nanotechnology, National Research Council, Edmonton, Canada, AB T6G 2M9

(Received 19 May 2011; revised manuscript received 26 July 2011; published 11 November 2011)

During fabrication and scanning tunneling microscope (STM) imaging of dangling bonds (DBs) on a hydrogenated silicon surface, we consistently observed halolike features around isolated DBs for specific imaging conditions. These surround individual or small groups of DBs, have abnormally sharp edges, and cannot be explained by conventional STM theory. Here, we investigate the nature of these features by a comprehensive three-dimensional model of elastic and inelastic charge transfers in the vicinity of a DB. Our essential finding is that nonequilibrium current through the localized electronic state of a DB determines the charging state of the DB. This localized charge distorts the electronic bands of the silicon sample, which in turn affects the STM current in that vicinity causing the halo effect. The influence of various imaging conditions and characteristics of the sample on STM images of DBs is also investigated.

DOI: [10.1103/PhysRevB.84.205416](https://doi.org/10.1103/PhysRevB.84.205416)

PACS number(s): 68.37.Ef, 72.20.Jv, 73.20.-r, 73.23.Hk

I. INTRODUCTION

Dangling bonds (DBs) on silicon crystal surfaces are becoming increasingly important candidates for building extremely miniaturized nanoelectronics devices. DBs are band-gap electronic states (or point defects) that can act as donors or acceptors, depending on their energy level with respect to the Fermi level. A surface silicon atom can host a dangling bond when only three of its valence electrons are involved in covalent bonds, the remaining valence electron being left in an unsatisfied (dangling) bond state. It turns out that a DB can be in three, energetically different states: a negative state DB^- (holding two localized electrons), a neutral state DB^0 (holding a single localized electron), and a positive state DB^+ (holding a localized hole). Traditionally, the energy eigenvalues of band-gap states are referred as transition levels, that is, the electrical levels accessible by scanning tunneling microscopy and spectroscopy. In this terminology, the transition levels are denoted as $E_{DB}(0/+)$ and $E_{DB}(-/0)$, where the slash in the middle represents the electrical level and the symbols on left/right of it represent the charge state if the Fermi level lies above/below that level.

The exact energy eigenvalues for an isolated DB on an otherwise perfectly terminated H:Si (001) 2×1 surface are not exactly known, but reasonable estimates can be obtained by various computational methods: an extended Huckel theory,¹ Poisson-Schrödinger equations,² and density functional theoretical (DFT) methods.³ The three-dimensional form of the DB wave function must, as a surface state, decay exponentially into vacuum and the crystal as well as laterally away from the center of the host Si atom. Theoretical estimates¹⁻³ and experimental measurements (by contactless capacitance-voltage method⁴) of the DB-state energy in the band gap are found to agree relatively well. In this study, we assume the energy eigenvalues for the different DB states are, with respect to the silicon valence band maximum (VBM): $E_{DB-} = 0.82$ eV,⁴ $E_{DB0} = 0.35$ eV,⁵ and $E_{DB+} = 0.0$ eV.² Because the latter state is degenerate with the silicon valence band, in this study, it is assumed not to play a significant role.

Studying constant-current empty-state scanning tunneling microscope (STM) images for different sample bias voltage, V_{bias} , and current setpoints reveals in some cases, the existence of a dark halo surrounding a bright spot of atomic size at the location of the DB. The appearance of such a halo has been also observed when imaging single dopant atoms located near the surface for other types of semiconductors.^{6,7} The simple and intuitive explanation is that this dark halo is caused by upward bending of the energy bands at the surface, leading to a reduction of the STM current (in constant-height imaging mode) or equivalently a decrease in the STM tip height (in constant-current imaging mode). In order to distinguish such a spatially limited band bending from one of a large scale, e.g., for a great density of DBs, it is important to understand to what extent band bending due to the charging of a collective of such surface states occurs *uniformly* along the surface and where the delimiting behavior occurs. Arguably, a rule of thumb is that band bending becomes uniform (and the Fermi level becomes pinned at the surface) when the average distance between two dangling bonds becomes equal to the screening length in the semiconductor. However, for the cases approached in this study, due to the great extent of H termination of our silicon surfaces, this condition is not fulfilled. Therefore the Fermi level is not pinned at such surfaces. This fact implies that each isolated DB (unless very close to others), has its own individual screening field due to mobile charge carriers in the crystal. On the other hand, isolated *groups* of closely spaced DBs (<2 nm) are electronically coupled,^{1,3,8} and have more or less a common screening field around the whole DB group. The coupling of such groups in the presence of an STM tip was recently studied experimentally and theoretically within an extended Hubbard model⁸ to yield time-average occupations of individual DBs without explicitly accounting for the dynamics of electrons in/out the DB group.

The properties of various types of single defect levels in semiconductors have been previously studied in the context of their STM imaging appearance.⁹ For DBs, the carrier capture properties^{10,11} and the on-site Coulomb interaction between two localized electrons can be extracted from spectroscopic

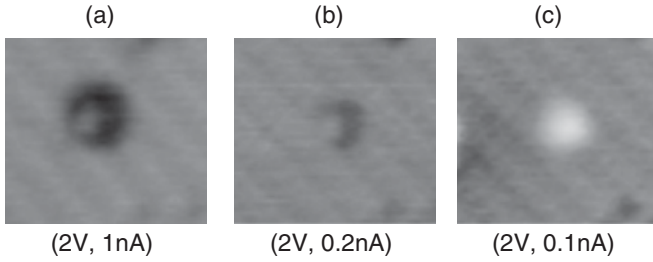


FIG. 1. STM images of a 4.0×3.6 nm region of the 2×1 H-Si (001) surface with an isolated DB in the unoccupied imaging mode (sample biased positively) at imaging setpoints (sample bias and current) indicated below each image.

measurements¹² on isolated DBs on a B-doped Si(111)-($\sqrt{3} \times \sqrt{3}$) $R30^\circ$. However, in these latter studies, electron dynamics is treated in a simplified manner, particularly with respect to the effects of tip-induced band bending at the surface. Here, we study single electron dynamics at a DB during STM imaging, accounting for all the important charge transfer mechanisms arising when band bending is present at the surface. Ultimately, we gain important insights into the charging state of a DB and its consequences in the appearance of STM images.

II. DBS IN THE POTENTIAL LANDSCAPE OF THE IMAGING PROBE

In Fig. 1, we show unoccupied-state STM images of a single isolated dangling bond on H-terminated Si (001) with a 2×1 surface reconstruction. In Fig. 1(a), the DB can be identified close to the center of the image as a bright spot surrounded by a darker ringlike or “halo” feature. As the setpoint current is decreased to 0.2 nA in (b), the halo diminishes and a further current decrease to 0.1 nA in (c) causes the halo to disappear and be replaced by a larger and brighter protrusion. This appearance is consistent with many experimental observations taken at such imaging conditions, and similar findings have been previously reported in the literature.¹³ Also, such imaging features have been observed in the case of other localized charge centers, such as subsurface dopants.⁹

We define an *isolated* DB on the silicon surface as one with no equivalent nearby DBs to be significantly coupled with (nor other resonant electronic levels). This implies that at an isolated DB, an electron is strictly localized in an orbital at a single Si atom. Regardless of its charging state, we assume a simple functional form for the DB wave function, namely, a Slater-type orbital (STO) having *p*-type symmetry (two lobes):

$$\psi_{\text{DB}}(\mathbf{r}) = \psi_{\text{DB}}(r, \theta) = Nr \cos(\theta) \exp[-\zeta(\theta)r], \quad (1)$$

where r and θ are the three-dimensional spherical coordinates and N is a normalization constant. The difference from a “regular” STO is that in our case the decay rate ζ in the exponent must vary with the z coordinate in order to ensure consistency with the asymptotic decay in the specific environment of the DB as follows. Note that the function ζ (and therefore ψ_{DB}) is dependent on the charge state of the DB as described below. The decay rate is related to the ionization potential W_i of the DB electron via $\zeta = \sqrt{2mW_i}/\hbar$. This ensures the correct asymptotic behavior (from first principles) of the wave function above. As the DB orbital is located partially in vacuum and partially in silicon, the ionization potential is not a constant, but it rather depends on the “ionization path” of the DB electron, e.g., it is energetically easier to extract an electron toward the bulk (into the conduction band of silicon) than toward the vacuum. To reflect this fact, here we assume a simple form for this dependence, namely,

$$W_i(z) = W_{\text{bulk}} + \frac{1}{2}(W_{\text{vac}} - W_{\text{bulk}})[\tanh(z/w) + 1], \quad (2)$$

where W_{bulk} and W_{vac} are the ionization potentials with respect to bulk and vacuum, respectively (their exact values are different for the DB^0 and DB^- species), z is measured from the surface and negative toward vacuum, and w is a characteristic width of the transition as can be seen in Fig. 2(a). The normalization of this wave function is calculated numerically and the resulting charge density for DB^- is plotted in Fig. 2(b) as a color map for a two-dimensional axial section. A value of $w = 5$ Å was chosen for the plots shown. Note that, as a consequence of our choice for the ζ function, the two lobes of the DB orbital are quite disproportionate, with the vacuum lobe being much smaller in the spatial extent and

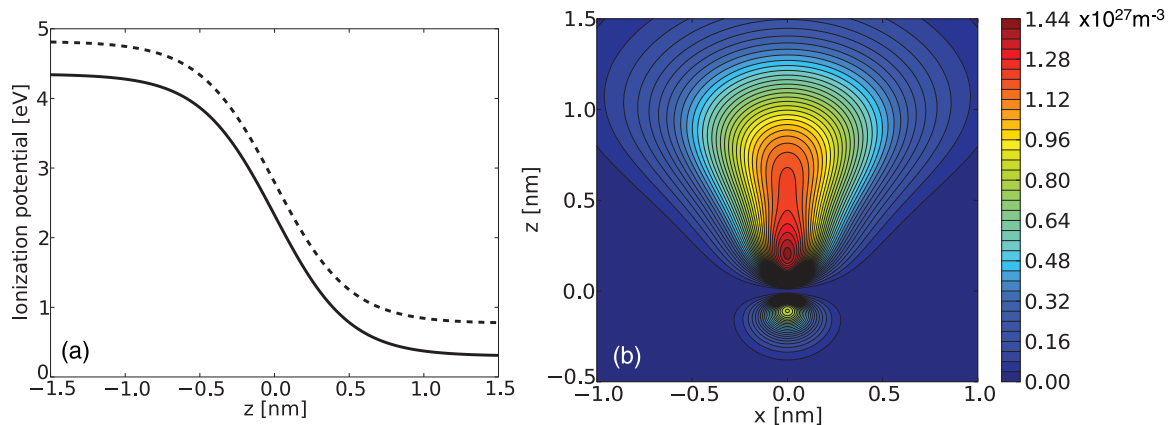


FIG. 2. (Color) Characteristics of the DB orbital. (a) The ionization potential as a function of distance from the surface (positive z in the sample) for DB^0 (dashed curve) and DB^- (solid curve) species. (b) Color map of the distribution of charge density of a DB^- in an axial plane as predicted by the modified STO function in Eq. (1). The color bar indicates the values of the charge density.

portion of charge. This choice of simplified orbital is also qualitatively consistent with DB orbital characteristics derived from *ab initio* calculations for an isolated DB at the surface of model hydrogenated silicon clusters,³ mimicking the 2×1 H-Si (001) surface. More rigorously speaking, due to the symmetry of the silicon lattice, the wave function of DB^0 exhibits sp^3 symmetry. Adding an extra electron will result in different hybridization coefficients for the DB^- wave function as compared to DB^0 .¹⁴ As the p -type component in both cases will be somewhere between 70–90%, we choose to keep only this dominant component in our model. Another simplification is that the rotational symmetry axis of the p orbital is taken to be perpendicular to the surface, while in reality, it is skewed. However, this does not affect our analysis in any essential way. Anticipating, the essential features of the DB orbital in our model are the characteristic decay tails toward the vacuum and the semiconductor, and these are appropriately captured in our model. Therefore, in the following analysis, we will make use of the simplified wave function in Eq. (1) for both DB^0 and DB^- , denoted ψ_{DB}^0 and ψ_{DB}^- , with corresponding ionization potentials W_i^0 and W_i^- , respectively, plotted in Fig. 2(a).

When the DB is negatively charged, the localized extra electron is expected to locally distort the electronic levels at the surface of the host silicon crystal, i.e., cause band bending. Band bending in a semiconductor is the phenomenon of distortion of electronic levels of the crystal in the presence of an external perturbing electric field. During STM imaging, another source of band bending is the biased STM probe itself and this is known as the tip-induced band bending. Depending on the value of the STM bias voltage, the amount of band bending can be considerable, sometimes enough to induce an inversion layer near the surface (whenever the Fermi level becomes lower than the intrinsic level of the semiconductor). The amount of the tip-induced band bending can be theoretically estimated by solving the Poisson and Schrödinger equations simultaneously, for example, by finite element methods (FEM) for an assumed probe geometry.

During STM imaging of isolated DBs, as the imaging tip is brought closer to the DB location, the band bending is turned gradually up and so is the DB level relative to the sample Fermi level. In Fig. 3, we show a typical band bending diagram calculated by FEM for the silicon surface in the presence of a scanning probe tip. The sample is n type with a resistivity of $3.5 \text{ m}\Omega \text{ cm}$, the sample bias voltage was assumed $+2 \text{ V}$, and the tip height was assumed to be 7 \AA . For these typical imaging conditions, the bands are bent upward to the extent that the sample Fermi level at the surface is 0.3 eV below the transition level ($-/0$). This value is much greater than kT , which—assuming thermodynamic equilibrium between the DB state and the bulk semiconductor—means that at room temperature, the DB should be half filled (neutral). However, the process of STM imaging also injects electrons into the DB state and an interplay of in/out transfer rates starts to form. In this study, we show how this fact has important consequences on the way surface states are imaged by STM.

An important consequence is evident from the diagram in Fig. 3; in the presence of the tip-induced field, the DB level becomes resonant with conduction band (CB) levels of the host crystal, implying that tunneling from DB into the crystal becomes possible. This situation is then similar

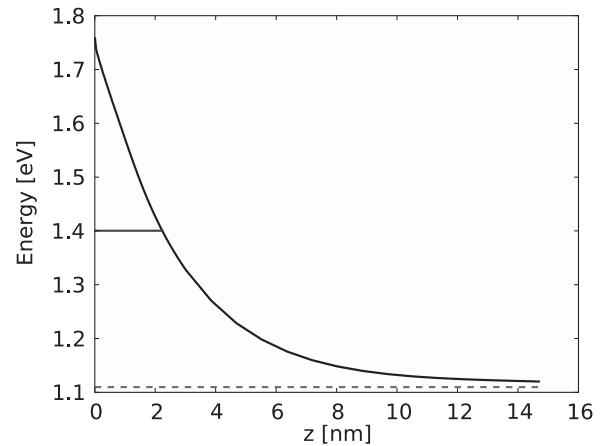


FIG. 3. Calculated band bending diagram as a function of the z coordinate. The shown curve is the conduction band minimum (CBM), the dashed horizontal line is the sample Fermi level, and the upper horizontal line shows the DB^- level with indication to tunneling into resonant levels in the silicon conduction band. The sample is n type with a resistivity of $3.5 \text{ m}\Omega \text{ cm}$, sample bias $+2 \text{ V}$, STM tip height 7 \AA .

to double barrier tunneling junctions, which were recently used to study energetic levels of single atoms,¹⁵ defects,¹⁶ and molecules^{17–19} at thin insulating films on metallic surfaces as well as metallic nanoislands on Si.²⁰ Manipulation of charge states¹⁷ and bond structure²¹ of molecules has also been demonstrated. However, the similarity of our system to these other cases is limited because in our case, there are a number of other factors at play that contribute to the charging of an individual DB, as we show below.

Careful examination of our available experimental data in conjunction with a quantitative theoretical analysis indicates that STM imaging features of dangling bonds are due to dangling bond charge caused by nonequilibrium effects during imaging, i.e., rate-limited charging/discharging of a DB during extraction/injection of electrons by the STM tip.

For routine unoccupied-state imaging conditions (sample bias $+2 \text{ V}$, current setpoint $100 \text{ pA}–1 \text{ nA}$), the tip-induced band bending renders the Fermi level of the sample to fall between the two transition levels ($0/+$) and ($-/0$) of the DB. Consequently, the transfer of an additional electron into the half-filled DB from the STM tip occurs in a non equilibrium steady state whenever the rate into the DB is higher than the rate out of the DB.

Explaining the DB imaging as purely quantum-mechanical behavior, namely, as Friedel oscillations in the local density of states, has been proposed—without consideration of band bending and nonequilibrium effects—but this explanation is inconsistent with the variation of the DB features with temperature and STM bias voltage.²²

III. NONEQUILIBRIUM CURRENT THROUGH A DB STATE

In the above discussion, we assert that a net negative charge on the DB during imaging can only be explained by a nonequilibrium phenomenon, and is in fact impossible if the DB is in electrochemical equilibrium with the silicon crystal

(due to the DB level being greatly elevated from the bulk Fermi level). In our system, we have a situation where the current flows from one reservoir (metal) at the chemical potential $\mu_{\text{tip}} = \mu_1$ to another (Si crystal) at a chemical potential $\mu_{\text{Si}} = \mu_2$ via a discrete level (DB) whose energy is somewhere between μ_{tip} and μ_{Si} . An electron can be transferred into or out of the localized DB level at two different rates, say $r_{\text{tip}} = r_1$ and $r_{\text{Si}} = r_2$, which in general include all elastic and inelastic charge transfer mechanisms. Overall, a steady state of current flow is achieved when $r_{\text{tip}} = r_{\text{Si}}$ in which the occupation of the DB level depends on the DB energy and the escape rates.

For a DB state in equilibrium with reservoir 1 only (or 2 only), its occupation equals the Fermi-Dirac distribution, $f_{\text{DB}}^{(1)} = f_{\text{FD}}(E_{\text{DB}}, \mu_1) = \{1 + \exp[(E_{\text{DB}} - \mu_1)/kT]\}^{-1}$ (or $f_{\text{DB}}^{(2)} = f_{\text{FD}}(E_{\text{DB}}, \mu_2) = \{1 + \exp[(E_{\text{DB}} - \mu_2)/kT]\}^{-1}$, respectively). However, since the DB level is in contact with both reservoirs, the DB will achieve a steady state of filling, with an occupation f_{DB}^* somewhere in between (the superscript * stands here for nonequilibrium occupation function). Then, the current through one contact equals the current through the other, $I_1 = I_2$, which is the condition for a steady-state occupation of a DB. In order to calculate the steady-state current and DB occupation, we must first determine the flow rates between the tip and DB on one hand and between sample and DB on the other. The possible mechanisms of charging and discharging a DB in the empty-state STM imaging mode are shown in Fig. 4.

Note that in our system, the first transition level (0/+) is always below the Fermi level and so the tip can not transfer electrons into that level. Therefore the single relevant occupation probability in our system is that of the DB^- level, f_{DB}^* , varying between 0 (for DB with zero time-average charge) and 1 (for DB with -1 time-average charge). In the sections below, we discuss each charge transfer mechanism shown in Fig. 4 in detail.

A. Rate of electron transfer between STM tip and an isolated DB

During an STM experiment, a possible mechanism for electron transfer from the STM tip to a DB (and vice versa) is by elastic tunneling across the vacuum gap. Another possible mechanism is a two-step capture process in which an electron from the tip first tunnels into the Si crystal in the vicinity of the DB and then immediately gets captured by the DB before it has the chance to become an extended-state electron. Henceforth, we call this mechanism ‘‘STM vicinity capture of tunneling electrons by a DB.’’ (1) The rate of elastic transfer between the STM tip and an isolated DB can be estimated (see Appendix A for more details):

$$I_{\text{tip-DB}}(d) = \frac{4\pi e}{\hbar} \frac{1}{\Omega_{\text{DB}}} \rho_{\text{tip}}(E_{\text{DB}}) \times [f_{\text{tip}}(E_{\text{DB}}) - f_{\text{DB}}^*] |\tilde{M}_{\text{tip-DB}}(E_{\text{DB}}, d)|^2, \quad (3)$$

where we denote the Fermi-Dirac distribution simply by f with a subscript indicating the respective electrode (‘‘tip’’ for tip, or ‘‘Si’’ for silicon crystal), Ω_{DB} is the volume of the DB (defined as the region in space where the DB charge density is greater than 1% of its maximum), and $|\tilde{M}(E_{\text{DB}}, d)|^2$ is a

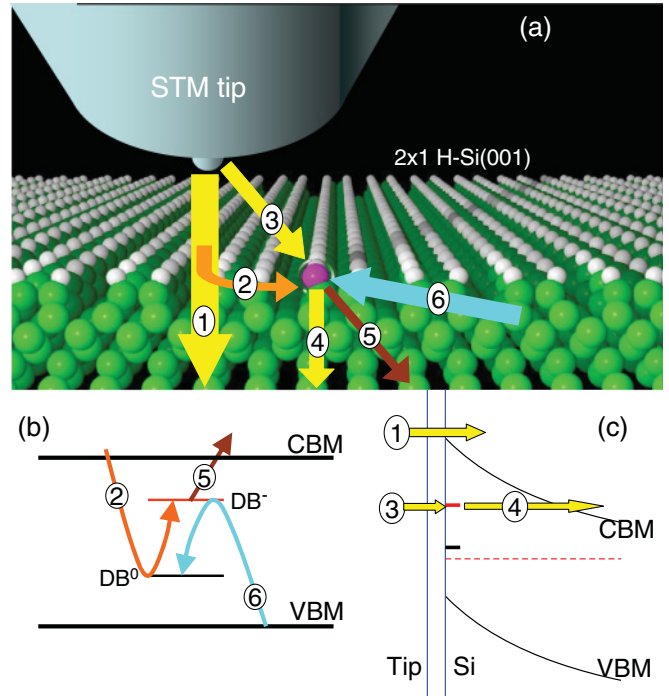


FIG. 4. (Color) Three-dimensional rendering of an isolated DB on the H-Si(001) surface in the presence of an STM tip in the unoccupied imaging mode. (a) Various channels for elastic and inelastic charge transfers between the STM tip and the silicon sample in the vicinity of the DB are indicated by colored arrows and labeled for further discussion. Si atoms are green, H atoms are white, and the DB is shown schematically as purple sphere. (b) Diagram of the silicon band gap, which further explains the various channels shown in (a): (2) inelastic capture by the DB^0 level (black horizontal line) of an STM electron after which DB becomes negative, (5) thermal electron emission from the DB^- level (red horizontal line) into the CB, and (6) inelastic capture of a hole by the negative DB, followed by recombination and the DB becoming neutral. (c) Tip-induced band bending diagram in the direction normal to the surface. The elastic charge transfers shown in (a) are detailed here: (1) direct tip-silicon tunneling, (3) tip-DB tunneling across the vacuum gap, and (4) tunneling from the DB^- state into the CB across the potential barrier created by band bending. The Fermi level in silicon is indicated by a dashed red line.

properly normalized transfer matrix element for separation d (between DB and tip centers) defined as

$$|\tilde{M}_{\text{tip-DB}}(E_{\text{DB}}, d)|^2 = \left(\frac{\hbar^2}{2m}\right)^2 \left| \int_{S_0} \mathbf{dS} \cdot [\varphi_{\text{tip}} \nabla(\varphi_{\text{DB}}^0)^* - (\varphi_{\text{DB}}^0)^* \nabla \varphi_{\text{tip}}] \right|^2. \quad (4)$$

Note that in Eq. (3) above, we also replaced the Fermi-Dirac distribution function for the DB with the nonequilibrium occupation function. As explained in Appendix A, the functions φ above represent dimensionless wave functions of the respective states under consideration defined as

$$\varphi(\mathbf{r}) = \sqrt{\Omega} \psi(\mathbf{r}), \quad (5)$$

where ψ is the normalized wave function and Ω is the volume of the electrode or the localized state, respectively. Also note

that the wave function of the neutral DB state enters the transfer matrix element above. For the purpose of estimating the tunneling currents, the wave function of the STM tip is taken in the conventional form of the Tersoff-Hamann approach:

$$\psi_E^{\text{tip}}(\mathbf{r}) = c_{\text{tip}}(R_0) \frac{\exp[-\zeta_{\text{tip}}(E)|\mathbf{r} - \mathbf{r}_{\text{tip}}|]}{\zeta_{\text{tip}}(E)|\mathbf{r} - \mathbf{r}_{\text{tip}}|} \quad (6)$$

for $|\mathbf{r} - \mathbf{r}_{\text{tip}}| > R_0$, where \mathbf{r}_{tip} is the position of the tip center, R_0 is a characteristic tip radius, $c_{\text{tip}}(R_0) = c_{0t}\zeta_{\text{tip}}R_0 \exp(\zeta_{\text{tip}}R_0)/\sqrt{\Omega_{\text{tip}}}$, c_{0t} is a dimensionless constant of the order of unity, $\zeta_{\text{tip}}(E) = \sqrt{2m(E_{\text{vac}} - E)/\hbar}$, and E_{vac} the vacuum level.

Generally speaking, one would expect sequential tunneling events from the STM tip, into both (0/+) and (-/0) transition levels. However, the DB can never be in charge state DB⁺ (holding a localized hole) because (i) the Fermi level of Si is always above the first transition level and (ii) the (0/+) is resonant with the silicon valence band (VB). Therefore tunneling into the first transition level can be neglected.

(2) The STM vicinity capture of tunneling electrons by a DB involves tunneling from the STM tip into the Si crystal for a range of electron energies and momentum orientations and subsequent capture of those hot electrons by a neutral DB. The corresponding current into the DB is denoted by C_n^{STM} and the details of its calculation are described in Appendix C.

During an STM experiment, in addition to the tunneling current from the tip to the DB, other transfer mechanisms contribute to the overall balance of charge on a DB orbital. They are discussed in detail below.

B. Rate of electron transfer between an isolated DB and the silicon crystal

Once an incoming electron becomes localized at a DB state, there are three main mechanisms of transfer into the silicon crystal.

(1) The localized DB electron escapes by emission into the conduction band via thermal excitation. This thermal escape rate only depends on the barrier height and temperature; the escape rate is given by

$$e_n = \sigma_n v_n N_c \exp[-(E_{\text{CBM}} - E_{\text{DB}})/k_B T], \quad (7)$$

where the prefactor is an attempt frequency and N_c is the effective density of states at the bottom of the conduction band:

$$N_c = \frac{1}{\sqrt{2}} \left(\frac{m_{n,\text{eff}} k_B T}{\pi \hbar^2} \right)^{3/2} \quad (8)$$

with $m_{n,\text{eff}}$ being the effective mass of electrons in Si.

(2) The localized DB electron tunnels elastically into the bulk CB level (at the classical turning point level where the CBM intersects the DB⁻ level). Tunneling from DB occurs across the space charge layer and the tunneling distance depends, via band bending, on the doping level of the sample, STM tip height, and bias voltage. The tunneling current can be estimated by using the formula (see Appendix A)

$$I_{\text{DB-Si}} = \frac{4\pi e}{\hbar} \frac{1}{\Omega_{\text{DB}}} g_{\text{Si}}(E_{\text{DB}}) [f_{\text{DB}}^* - f_{\text{Si}}(E_{\text{DB}})] |\tilde{M}_{\text{DB-Si}}(E_{\text{DB}})|^2, \quad (9)$$

where $\tilde{M}_{\text{DB-Si}}$ in this case also depends on the tunneling barrier for an electron with energy E_{DB} , which is given by the band-bending dependence on z in the presence of a negative DB. Also note that $\tilde{M}_{\text{DB-Si}}$ is dependent on the wave function of the negative DB. Upon tunneling discharge into the bulk, the DB resumes a neutral state.

To complete the picture, for the purpose of calculating transfer coefficient between the tip and the silicon sample, we assume sample wave functions derived from the ‘‘jellium model’’²³ with modifications to account for the corrugation of the surface:

$$\psi_{E,\mathbf{K}}^{\text{Si}}(\mathbf{r}) = -c_{\text{Si}} \frac{k_z \exp\left\{-\sqrt{\zeta_{\text{Si}}^2 + K^2}[z - H(\mathbf{R})]\right\}}{\sqrt{\zeta_{\text{Si}}^2 + K^2 - ik_z}} \times \exp(i\mathbf{K} \cdot \mathbf{R}), \quad (10)$$

where E is the eigenenergy of the state, $\mathbf{K} = (k_x, k_y)$ is the surface parallel wave vector, k_z is the surface normal wave vector [the last three being related via $E = \hbar^2(K^2 + k_z^2)/2m$], $\mathbf{r} = (x, y, z)$, $\mathbf{R} = (x, y)$, Ω_{Si} is the sample volume, $\zeta_{\text{Si}}^2 = 2m(E_{\text{vac}} - E)/\hbar^2$, and $c_{\text{Si}} = i c_{0s}/\sqrt{\Omega_{\text{Si}}}$ with c_{0s} a dimensionless constant of the order of unity. The corrugation function of the sample $H(\mathbf{R})$ is, in general, two-dimensional, but in our model we treat the x and y directions as independent. We assume the corrugation component along dimer rows (for which we calculate STM current traces) has the form $H(x) = A \cos^2(\pi x/L_x)$, with A the corrugation amplitude and L_x its period.

(3) A hole striking the surface is captured by the occupied DB level (defect mediated recombination). The associated recombination current is calculated in a customary approximation by assuming that the rate of recombination at a DB site is proportional to concentration of holes, p , at the vicinity of the DB, the thermal velocity of those holes, v_p , and a capture cross section of a hole by a negatively charged DB, σ_p :

$$r_{\text{rec}} = \sigma_p p v_p. \quad (11)$$

Their numerical values are $v_p = 1.87 \times 10^7$ cm/s, $\sigma_p = 1.1 \text{ \AA}^2$, while p depends on the band bending at the surface and can be anywhere from 70 cm^{-3} in the bulk of highly doped n -type Si (resistivity 0.01 \Omega cm) to 10^{17} cm^{-3} in an inversion layer. The recombination current at the DB is given by

$$I_{\text{rec}} = C_p = e \sigma_p p v_p f_{\text{DB}}^*. \quad (12)$$

An additional note of the capture cross section; defects in semiconductors, both in bulk and at surface/interface, have been known to act as very effective recombination centers. In fact, carrier recombination in a semiconductor is actually dominated by trapping processes at defects, such as DB states. An important fact is that defects with energies close to band edges (as opposed to the middle of the band gap) are more efficient at trapping carriers and causing recombination.²⁴ Also, the above rough values of recombination rates rely on the semiclassical estimates for the hole concentration at the surface and include no charge quantization effects due to band bending. This quantum effect can significantly alter the carrier concentration at the surface and therefore can be important for the recombination rates.

Generally speaking, the capture cross section has a thermally activated behavior:

$$\sigma_p = \sigma_{0p} \exp(E_a/k_B T), \quad (13)$$

where σ_{0p} is a constant and E_a is an activation energy. For our purpose, we note that the capture cross section of a DB is more or less independent of temperature between 100–400 K,²⁵ for both electron and hole capture. It follows, that in the absence of tunneling, the flow of electrons in/out of a DB is given by

$$R_n = C_n - E_n = e\sigma_n v_n n(1 - f_{DB}^*) - ee_n f_{DB}^* \quad (14)$$

with capture competing with emission and the unknown occupation, f_{DB}^* , dictated by the nonequilibrium conditions.

IV. SOLVING FOR THE NONEQUILIBRIUM OCCUPATION OF A DB

Combining all the above mechanisms, we obtain the total current flow between DB and bulk:

$$I_{DB}^{in} = I_{DB}^{out}, \quad (15)$$

yields

$$I_{tip-DB} + C_n^{STM} + C_n = I_{DB-Si} + E_n + C_p, \quad (16)$$

where the additional current components, in ampere, are given by

$$C_n = e\sigma_n v_n n(1 - f_{DB}^*), \quad (17)$$

$$E_n = ee_n f_{DB}^*, \quad (18)$$

$$C_n^{STM} = e\sigma_n I_{tip-Si}^{DB0} (1 - f_{DB}^*) \int_{\Omega_{DB}} d\mathbf{r} |\psi_{DB}^0(\mathbf{r})|^2 \frac{D[\theta(\mathbf{r})]}{2\pi r^2}, \quad (19)$$

$$C_p = e\sigma_p v_p p f_{DB}^*, \quad (20)$$

where v_p is the thermal velocity of holes and N_v is the effective density of states at the top of the valence band.

$$N_v = \frac{1}{\sqrt{2}} \left(\frac{m_{lp,eff} k_B T}{\pi \hbar^2} \right)^{3/2} + \frac{1}{\sqrt{2}} \left(\frac{m_{hp,eff} k_B T}{\pi \hbar^2} \right)^{3/2} \quad (21)$$

with $m_{lp,eff}$ and $m_{hp,eff}$ being the effective masses of light and heavy holes in Si, respectively.

The solution is

$$f_{DB}^* = \frac{i_{tip} f_{tip}(E_{DB}) + i_{Si} f_{Si}(E_{DB}) + c_n + c_n^{STM}}{i_{tip} + i_{Si} + c_p + e_n + c_n + c_n^{STM}}, \quad (22)$$

where the charge transfer rates in s^{-1} are given by

$$c_n = \sigma_n v_n n, \quad (23)$$

$$c_n^{STM} = \sigma_n \frac{I_{tip-Si}^{DB0}}{e} \int_{\Omega_{DB}} d\mathbf{r} |\psi_{DB}^0(\mathbf{r})|^2 \frac{D[\theta(\mathbf{r})]}{2\pi r^2}, \quad (24)$$

$$c_p = \sigma_p v_p p, \quad (25)$$

$$i_{tip} = \frac{4\pi}{\hbar} \frac{1}{\Omega_{DB}} g_{tip}(E_{DB}) |\tilde{M}_{tip-DB}(E_{DB})|^2, \quad (26)$$

$$i_{Si} = \frac{4\pi}{\hbar} \frac{1}{\Omega_{DB}} g_{Si}(E_{DB}) |\tilde{M}_{DB-Si}(E_{DB})|^2, \quad (27)$$

and where $D[\theta(\mathbf{r})]$ in c_n^{STM} is the normalized tunneling current density in the direction θ from the STM tip to the sample, which is derived in Appendix C.

The total time-average STM current is then given by

$$I_{STM} = I_{bulk}^{DB0} (1 - f_{DB}^*) + I_{bulk}^{DB-} f_{DB}^* + I_{DB}^{out}, \quad (28)$$

where I_{bulk}^{DB0} and I_{bulk}^{DB-} are the tip-to-sample currents in the presence of a neutral and a negative DB, respectively.

From *ab initio* calculations on isolated DBs at the surface of model silicon clusters,³ we know that a negatively charged DB renders its host Si atom to be slightly elevated from the plane of the surface, about 0.3 Å, which can also have an effect on the tunneling coefficient during STM imaging. This effect is also captured in our model by accordingly modifying the sample wave function, and plays a role in the appearance of the DB.

Another possible effect during DB imaging is a Stark shift of the DB level due to the tip-induced field in which the DB is found. This shift would be most pronounced when the field is stronger, i.e., when the tip apex is immediately above the DB, for a given sample bias. In principle, calculating the magnitude of this shift is possible within the frame of our current theory by solving the Schrödinger equation in the presence of the external field, e.g., by a perturbative approach or finite element method. This would produce another parameter in our model, namely, the magnitude of the Stark shift as a function of the tip-DB distance and tip height. However, for the sake of simplicity, we decide not to include this additional parameter in our model, but rather to comment on its possible consequences on the model results.

Note that if the “equilibrium would-be” populations are equal, the net current through the DB is zero (this only happens if the chemical potentials on the two sides are equal). Also note from above formulas that, if one of the flow rates is much smaller than the other, then the net current will be reduced to just that rate. As another general remark, an implicit assumption above is that the total STM current (through DB and bulk states) is small enough that it does not significantly change the Fermi level of the reservoirs in the vicinity of the two contacts.

V. RESULTS OF THE IMAGING MODEL

In this section, we present the numerical results of the theory presented above for the imaging of DBs. We illustrate our model with experimental and theoretical results on a highly doped *n*-type silicon sample, with a H-terminated 2×1 (001) surface reconstruction. We only focus on the unoccupied-state imaging mode of STM, where most of our experimental data were obtained. In this case, upward band bending occurs at the surface and an excess concentration of holes is present there. The STM tip injects electrons into the sample and into isolated DBs as well. Localized electrons at DBs discharge into the sample by the different channels described above, including recombination with the excess holes at the surface. Other sample doping type and levels are also amenable to this model, keeping in mind that the system parameters will be different and require making appropriate adjustments: Fermi level, band bending, local electron, and hole concentrations at

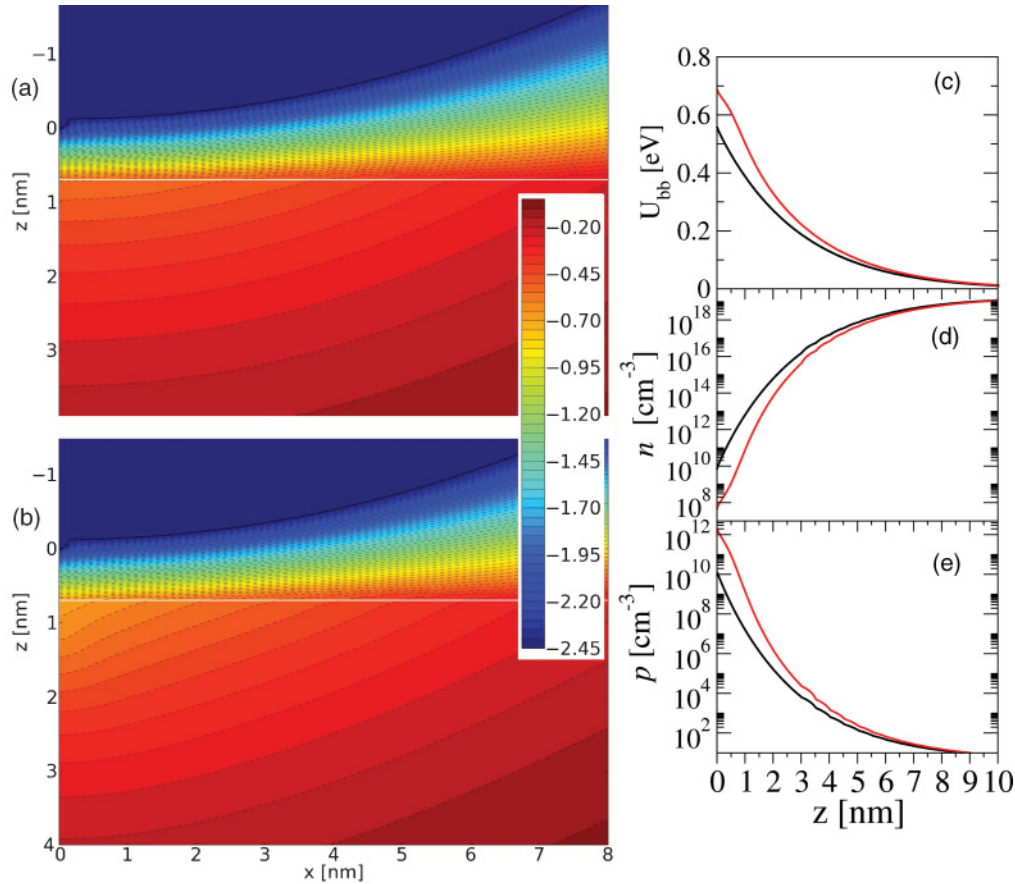


FIG. 5. (Color) (a) Two-dimensional section of the electrostatic potential landscape caused by a biased STM probe in the vicinity of a silicon surface for unoccupied state imaging and for axial symmetry (tip central axis is at $x = 0$). The potential penetrates the surface and vanishes deep in the crystal. The Si crystal is n type with a resistivity of $3.5 \text{ m}\Omega \text{ cm}$, voltage bias on the STM tip is -2 V , and the tip height was assumed to be 0.7 nm (tip apex is at $z = 0$). The tip contour is drawn as a black line and the surface as a white line. (b) The electrostatic potential for the same biasing conditions as in (a) and in the presence of a negative DB directly under the tip apex. The color bar indicates the values of the electrostatic potential in V, and is identical for both plots for easy comparison. (c)–(e) Subsurface band bending and carrier concentration along the central axis of the tip plotted as a function of sample depth, z , for biased tip in the presence of a DB⁰ (black curves), and in the presence of a DB⁻ (red curves) placed directly under tip apex.

the surface, etc. For example, in the occupied-state imaging, electrons are extracted from the sample into the STM tip, and downward band bending occurs at the surface, which brings about an excess of electrons, as opposed to an excess of holes. This has consequences on the filling of isolated DBs prior to their discharge to the STM tip.

Finite element analysis was used for calculating tip-induced band-bending as the self-consistent solution of Poisson and Schrödinger equations.²⁶ This solution is reliable for the regime in which quantum confinement effects (quantization of electronic levels at surface) are negligible, but also for the case of strong inversion or accumulation, when quantum effects are important.²⁷ In this latter case, many surface quantized levels are occupied, which makes the semiclassical approximation for the charge distribution valid. The semiclassical calculation fails to be reliable when the inversion or accumulation is moderate, that is, when surface quantized states exist but only a few quantum states are occupied. Some details of the FEM calculation are given in Appendix B, and the complete details of the whole calculation procedure are described elsewhere.²⁶

In Fig. 5, we show an example of a calculated potential landscape (by FEM) for chosen sample bias and tip height as [in panels (a) and (b)] a two-dimensional section of the axially symmetric system, and (c) a function of the depth inside the silicon surface. The effect of a negative DB under the tip on the band bending was also calculated and shown here. The greater differences between the two cases can be seen in the silicon region immediately under the tip apex, where the DB is located. Simultaneously, we calculate the carrier concentrations in the sample induced by the tip alone and also by the tip in the presence of a DB⁻, and plot the results in panels (d) and (e).

Unlike the case of a highly doped p -type Si sample, where the main discharging rate of a DB⁻ is that of recombination with holes,¹⁰ in our case, both types of carrier concentration are small at the surface, where the sample is rendered almost intrinsic by the tip-induced potential. As a consequence, recombination is no longer the dominant discharging rate and it competes with the other rates described above.

The calculated band bending and carrier densities are input to the mathematical formalism described above in order to

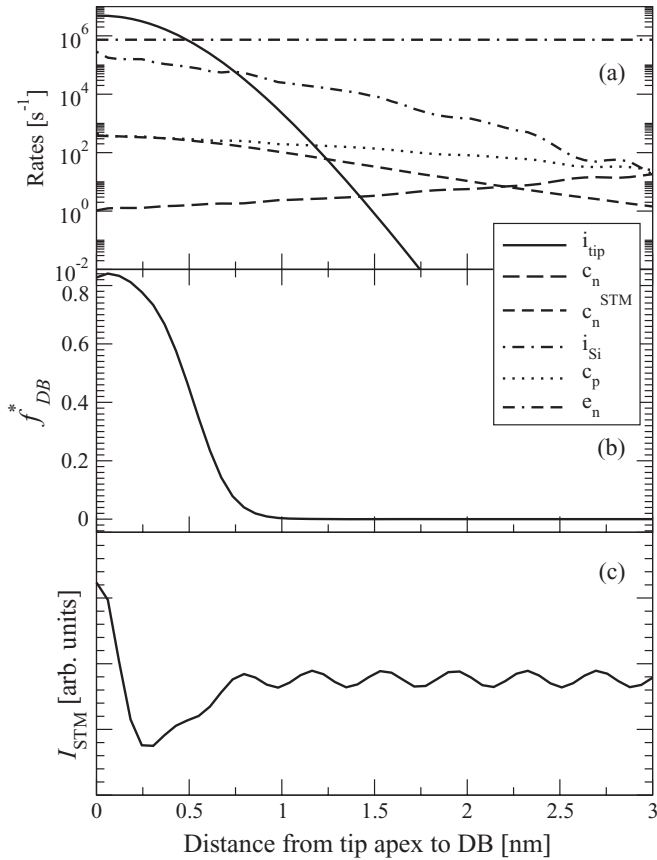


FIG. 6. Dependence of relevant physical quantities on the surface-parallel distance between the STM tip apex and the center of an isolated DB. Voltage bias on the sample is +2 V and the tip height is 7 Å from the surface. (a) The dependence of various elastic and inelastic rates of charge transfer through a DB in the presence of an STM imaging tip. (b) Nonequilibrium occupation function of the DB⁻ state versus distance. (c) Total STM current dependence on the same distance. This situation corresponds to the experimental Fig. 1(a), where a halo is formed.

compare our theory to our experimental observations. As suggested above, the key point is to determine the charging state of the DB. The main results needed to understand the nonequilibrium charging effect are shown in Figs. 6 and 7. Figure 6(a) illustrates the competitions between electron transfer rates in and out of an isolated DB in the unoccupied-state imaging mode. As the STM tip moves in from far away toward the DB, the electron injection rate from tip to DB (solid black curve) increases approximately as an exponential function, up to a critical distance r_c of about 5 Å. At that point, this injection rate becomes greater than the dominant rate of transfer out of DB, which in this case, is provided by the thermal emission of electrons into the conduction band. The second greatest discharge rate is by tunneling into resonant CB levels, process labeled i_{Si} in Fig. 6(a), and the third is by capture and recombination with an itinerant hole, c_p , which is considerably smaller. Note that STM vicinity capture of electrons, c_n^{STM} is of small magnitude, 10^1 – 10^2 s⁻¹, which is intuitively expected on the basis that it describes a Bloch state

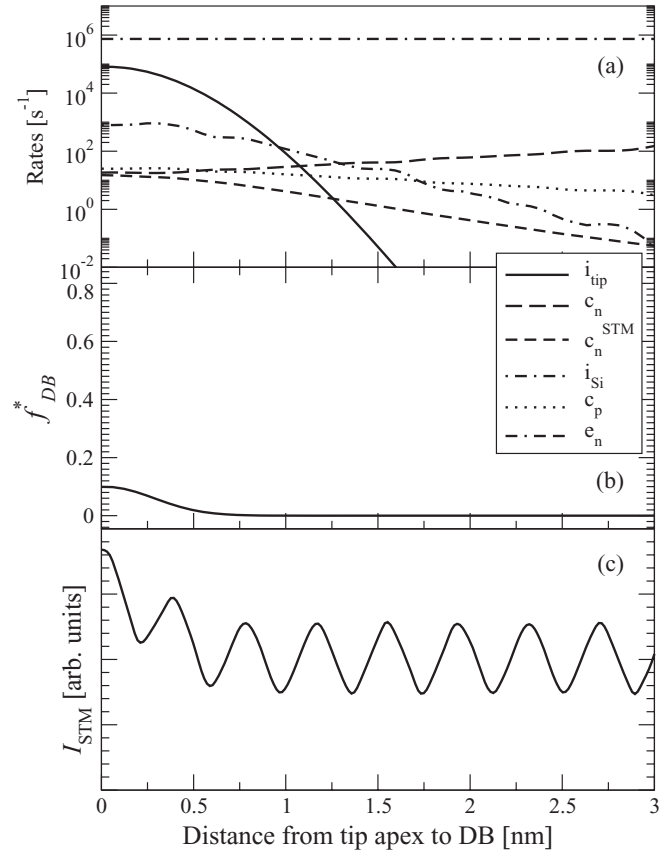


FIG. 7. The same quantities are plotted as in Fig. 6 but for a different tip height, 8.5 Å, which corresponds to the case in the experimental Fig. 1(c), where halo is not formed.

electron repelled away from the surface by the tip-induced field.

For smaller distances than r_c , the discharge rate becomes the limiting rate for current through the DB, and therefore it dictates the charging state of the DB, f_{DB}^* , plotted in Fig. 6(b). In turn, this charging state determines the total band bending (tip-induced plus DB-induced) and also the tip-sample tunneling coefficient in the vicinity of the DB and it therefore affects the total current measured in the STM scan, see Fig. 6(c). Over a relatively short distance, in the vicinity of r_c , the time-average DB occupation switches from zero to one, because the electrons localized there cannot escape at the same rate at which they are injected. With the addition of a localized negative charge at the DB center, the electrostatic landscape changes in its vicinity, i.e., upward band bending increases, and the STM current is affected, giving rise to a dark halo with the outer edge around r_c . This corresponds to the situation depicted in the experimental image in Fig. 1(a). The bright spot at the DB center is caused by the slight elevation of the host silicon atom at that location²⁸ as mentioned above, which although small around 0.3 Å, has a significant effect on the tunneling coefficient due to the exponential sensitivity of the latter with tip-sample separation.

A halo does not always appear around an isolated DB, and in our case, this happens at lower injection current, as in the situation shown in Fig. 1(c), which is consistent with a

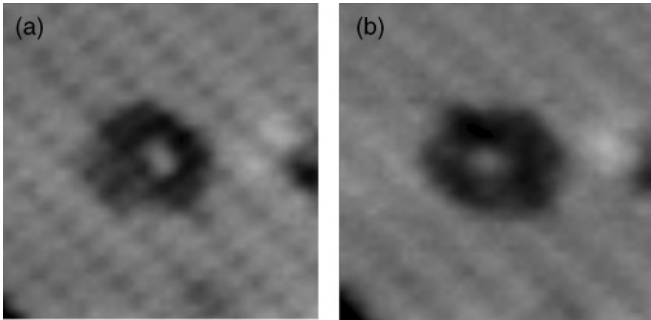


FIG. 8. Dependence of the shape of a DB halo on the particular shape of the STM tip used for imaging. The same DB was imaged in (a) and (b) at the same imaging setpoint (sample voltage 2 V, imaging current 0.2 nA), but with a tip change between (a) and (b).

greater tip height. To simulate the situation in our model, we performed the same calculations for a tip height of 8.5 Å and the results are shown in Fig. 7. We can see a decrease in this case of the STM injection rate into the DB, which becomes lower than the thermal emission rate as seen in (a). Therefore the DB is mostly neutral in this situation, as seen in (b). No halo is formed here because additional band bending by the DB does not occur.

Regardless of tip height, the DB is negative when the tip is laterally very far from the DB. There are indeed two “far regimes:” (i) asymptotically far, when the tip does not affect the potential landscape around the isolated DB and (ii) too far for tunneling into the DB (a few nm to tens of nm), but where band bending is still significant. However, the regime in (i) is not captured in our Figs. 6 and 7 because the lateral range would need to be extended to distances much greater than radius of the STM tip. Also in the asymptotic regime (i) the corresponding rates in these two figures become independent on tip height.

Over the course of many experiments, we also observed other related imaging trends of DBs: (i) there is a tip dependence in the exact shape of a DB halo, see Fig. 8, and (ii) halolike features around groups of two or more tunnel-coupled DBs. The dependence in (i) can be qualitatively described by the fact that different STM tips have atomic size protrusions at different locations in the vicinity of the apex. During scanning, as the tip approaches a DB, there are different onset points for complete charging ($f_{DB}^* = 1$) as the smallest distance between tip and DB (and the corresponding tunneling rate) depends on the location of the protrusions for each tip.

For (ii), the same imaging mechanism is at work for groups of coupled DBs, except the localized charge is shared among more than one DBs, and the charge per atom is less.³ Consequently, the DB-induced band bending tends to be lower. Also, due to the shape of the STM tip, the tip-induced band bending amounts to different values at different DBs, and therefore their levels are pulled upward with respect to the bulk VBM by different amounts. The localized charge tends to reside at the state whose energy level is lower (farthest from the tip).

Other possible effects were also taken into account in our analysis. The calculation of the charge carrier densities at the surface, accounting for charge quantization in the

surface-perpendicular direction, was also performed in a one-dimensional model (z dependence only) by solving the Schrödinger equation for the quantum well formed by the band bending in the z direction. However, for the situation depicted in Fig. 5, band bending is not strong enough to induce significant quantization of the space charge layer, and the semiclassical treatment is reliable.

As we mentioned above, during an STM scan when the tip apex is directly above the DB, the electric field under the apex can be very strong (around 2 V/nm in the vacuum gap), and this can induce a Stark shift of the DB level and a deformation of the DB orbital. This shift is of the order of a few tens of meV (from our preliminary estimates) and is not included as a parameter in our model. However, this shift has possible consequences on the appearance of DB, especially in the region close to the DB center. An upward shift in E_{DB} would decrease the band bending barrier and thus increase the i_{Si} rate, and also the thermal excitation rate e_n (these rates are very sensitive to the exact DB level). This means that more current would pass through the DB state itself, contributing to a brighter peak at the DB center. We expect this effect to become significant only for images taken at lower tip heights (higher setpoint currents).

Band bending in a semiconductor is usually explained by appealing to semiclassical arguments. However, for the semiclassical theory to be accurate, one must involve the approximation that the perturbation in the crystal varies slowly on the spatial scale of the crystal lattice constant. In order to test this approximation for our case, here we also undertake a more general approach that can be used beyond the semiclassical approximation, which can also cast insight into the long-range behavior of screening of a negative DB. For this purpose, the calculation of the electrostatic potential landscape was done by separating the whole problem into two components, considered to be approximately independent, according to the source of electrostatic perturbation: (i) the tip-induced band bending and (ii) band bending by an isolated charged DB alone. Solving these two problems requires significantly different approaches caused by the different spatial scales of the problems. A perturbation approach^{29–31} was used for the problem of DB-induced band bending. This approach includes the semiclassical approximation and the random phase approximation (RPA) as particular limits, and allows us to analyze other quantum mechanical effects, such as Friedel oscillations.

In order to make the computational task feasible for long spatial range (100 nm), the angular dependence of the DB orbital was integrated out in this case (with only radial dependence remaining). The DB is considered to be placed in a background charge density determined by the tip-induced band bending evaluated by FEM calculations. In Fig. 9, we show the solution of the (separate) electrostatic problem of screening of a DB by the mobile carriers in the semiconductor. The maximum band bending due to the DB⁻ is around 0.17 eV, which is slightly larger than the all-semiclassical results presented above. Thermally damped Friedel oscillations are indeed present in our results [see inset in Fig. 9(a)], but they cannot be associated with the appearance of dangling bonds because: (i) they are extremely small in magnitude compared to the other effects discussed in our model (by more than

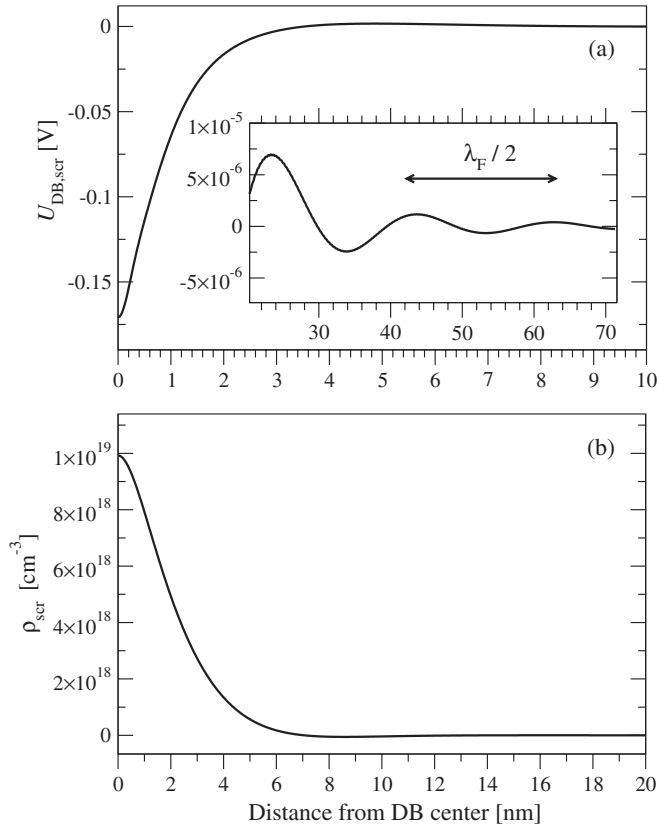


FIG. 9. (a) The screened electrostatic potential solution for a negatively charged single DB as a function of distance from the DB center calculated using the perturbation approach.^{29–31} Inset: the long-range behavior of the potential displaying thermally damped Friedel oscillations, albeit with much lower amplitudes than the potential at close range (the horizontal axis is the distance in nm for all plots). (b) Charge density of the mobile carriers (holes) brought about to screen the DB charge.

four orders of magnitude) and (ii) they are manifest at greater distances (>20 nm) than the length scales associated with imaging halos.

Over the course of many experiments, we noticed a significant effect of the sample preparation methods on the way DBs image. In particular, there is a strong correlation between the value of the temperature used for flash heating the sample and the formation of (or lack of) dark halos around DBs. This effect is caused by the fact that at high flash heating temperatures (>1250 °C), there is a substantial migration of dopant atoms away from the surface region. This dopant depletion happens up to a depth ranging from hundreds of nm to a few micrometers and alters the local Fermi level at the surface. In turn, this has consequences on the charge transfer rates discussed above and ultimately on the charging state of the DB under imaging conditions. These facts also suggest possible control mechanisms of the charging state of a DB, by adjusting the local Fermi level at the surface, e.g., by applying an external potential via a local electrode. This was also illustrated and discussed in a recent study on groups of closely coupled DBs on hydrogen-terminated silicon surfaces.⁸

VI. SUMMARY AND OUTLOOK

In this paper, we propose a mechanism for understanding the STM imaging of isolated DBs on silicon surface with quantitative focus on the n -type surfaces and unoccupied-state imaging mode. A quantitative analysis of the electrostatics of our system shows that the tip-induced band bending plays an important role in determining the charging state of the isolated DB by effectively pulling the DB state out of equilibrium with its host crystal. Under these conditions, the equilibrium Fermi-Dirac occupation of a DB is replaced by an occupation function dictated by the steady-state flow of current through the DB, which is very sensitive to the lateral DB-tip distance. Therefore the charging state of the DB can vary quickly upon lateral approach of the scanning tip, thus causing halo-like features with sharp edges around the DB. The sharpness of the edge is related to the slope of the variation of the DB occupation function f_{DB}^* with the lateral distance, see Fig. 6.

We also found that the all-semiclassical approach to band bending by an isolated DB^- compares reasonably well, in terms of excess band bending by a DB^- , to more general approaches, specifically the RPA, and that Friedel oscillations do not account for imaging features present in this system.

As the dominant DB discharging mechanism in our specific case study is the thermal emission of electrons into the silicon CB, lowering the system temperature should have a drastic effect on the DB image, and the halo should be increased. The similarity of our system to double barrier tunneling junction can also be exploited in certain regimes, e.g., by lowering the temperature and decreasing the thermal emission, thus rendering the DB-Si tunneling as the dominant discharging rate. These will be the subject of a further experimental investigation.

Our model calculates the total STM current in a time-average view of the system. If the time resolution of the STM was high enough, it should be possible to discern a “telegraph signal” of the current inside the halo region of an isolated silicon DB, as the band bending fluctuates with the charging/discharging of the DB. Our model does not reproduce the STM tunneling currents in a strictly quantitative fashion. However, it does capture all the relevant mechanisms at play in image formation, and the main qualitative trends that were observed experimentally, with the variation of system parameters such as doping level, sample bias, and current setpoint. Because the charging and discharging rates have exponential sensitivities to band bending values and distances between electrodes and localized states, their particular balance can be greatly affected by small changes in these parameters. Therefore imaging of DBs will be different for different semiconductor materials (Ge, GaAs), different surface reconstructions [e.g., Si(111)], for samples doped at lower or higher levels, or p type, and for different tips. The latter is the cause of daily variations in the appearance of DBs in the laboratory. However, the general formalism presented here can be adapted to suit those similar systems.

ACKNOWLEDGMENTS

This work is supported by iCORE, the National Research Council of Canada, NSERC, and Cifar. We are grateful to

Professor Randall Feenstra and Dr. Gino DiLabio for fruitful discussions.

APPENDIX A: FORMULAS FOR CALCULATING TUNNELING CURRENTS

Throughout this study we use a generalized formula³² for tunneling currents between two electrodes (1 and 2) of volumes Ω_1 and Ω_2 in the presence of a externally applied voltage, V_{bias} . They are equal to the difference between the chemical potentials on the two sides $\mu_1 - \mu_2$:

$$I_{1 \rightarrow 2} = \frac{4\pi e}{\hbar} \int dE [f(E, \mu_1) - f(E, \mu_2)] g_1(E) g_2(E) |\tilde{M}(E)|^2, \quad (\text{A1})$$

Here, $g_1(E)$ and $g_2(E)$ are appropriately generalized local density of states (LDOS) per unit energy and unit volume (in units of $\text{J}^{-1}\text{m}^{-3}$) of the two electrodes, and $|\tilde{M}(E)|^2$ is an appropriately defined (see below) transition matrix element between states in each electrode having energy E . We assume that our system fulfills the conditions in which the above formula is equivalent to the golden-rule expression for the tunneling current³³

$$I_{1 \rightarrow 2} = \frac{2\pi e}{\hbar} \sum_{m,n} [f(E_m, \mu_1) - f(E_n, \mu_2)] \delta(E_m - E_n) |M_{mn}|^2, \quad (\text{A2})$$

where E_m and E_n are the energy levels of eigenstates m and n on each side with the (spin excluding) wave functions ψ_m and ψ_n , respectively. In a first-order perturbation theory, the transition matrix element was shown to be³⁴

$$|M_{mn}|^2 = \left(\frac{\hbar^2}{2m} \right)^2 \left| \int_{S_0} dS \cdot [\psi_m \nabla \psi_n^* - \psi_n^* \nabla \psi_m] \right|^2, \quad (\text{A3})$$

where the surface integral is carried out over a median surface S_0 .

If we define dimensionless wave functions φ , instead of normalized ones ψ , by

$$\varphi_m(\mathbf{r}) = \sqrt{\Omega_1} \psi_m(\mathbf{r}), \quad (\text{A4})$$

then we can further write

$$|M_{mn}|^2 = \frac{1}{\Omega_1 \Omega_2} \left(\frac{\hbar^2}{2m} \right)^2 \left| \int_{S_0} d\mathbf{S} \cdot [\varphi_m \nabla \varphi_n^* - \varphi_n^* \nabla \varphi_m] \right|^2, \quad (\text{A5})$$

and define a volume-independent transfer matrix element as

$$\begin{aligned} |\tilde{M}_{mn}|^2 &= \Omega_1 \Omega_2 |M_{mn}|^2 \\ &= \left(\frac{\hbar^2}{2m} \right)^2 \left| \int_{S_0} d\mathbf{S} \cdot [\varphi_m \nabla \varphi_n^* - \varphi_n^* \nabla \varphi_m] \right|^2, \end{aligned} \quad (\text{A6})$$

which appears in Eq. (A1) and fulfills the correct dimensionality of the integral.

We apply the above formalism for calculating the tunneling current for the three tunneling cases in our system: (i) between the STM tip and the silicon sample, (ii) between the STM tip and an isolated DB, and (iii) between an isolated DB and the

silicon sample. Below we adapt the general formula above to each of these cases in order to derive the simpler forms given in the text.

For the first case, we assume that the tip wave functions are given as in the Tersoff-Hamann approach by Eq. (6) and its LDOS is constant in energy. Also, we assume that the sample LDOS is constant in space and has the customary square root dependence on energy for the bands in semiconductors:

$$g_{\text{Si}}(E) = \frac{8\pi\sqrt{2}}{\hbar^3} m_{\text{n,eff}}^{3/2} \sqrt{E - E_{\text{CBM}}} \quad (\text{A7})$$

for the conduction band levels (above CBM) and

$$g_{\text{Si}}(E) = \frac{8\pi\sqrt{2}}{\hbar^3} m_{\text{p,eff}}^{3/2} \sqrt{E_{\text{VBM}} - E} \quad (\text{A8})$$

for the valence band levels (below VBM), with $g_{\text{Si}}(E)$ being zero in the band gap ($m_{\text{n,eff}}$ and $m_{\text{p,eff}}$ are the effective masses for electron and hole in silicon, respectively).

For the tunneling between a DB and an electrode (tip or sample), we take advantage of the fact that the LDOS of the DB is a δ function in energy, namely, $g_{\text{DB}}(E) = \delta(E - E_{\text{DB}})/\Omega_{\text{DB}}$, and therefore the integral over energy in Eq. (A1) is reduced to a single-energy term for E_{DB} . This yields a simplified formula

$$\begin{aligned} I_{\text{tip-DB}} &= \frac{4\pi e}{\hbar} \frac{1}{\Omega_{\text{DB}}} \rho_{\text{tip}}(E_{\text{DB}}) \\ &\times [f_{\text{tip}}(E_{\text{DB}}) - f_{\text{DB}}^*] |\tilde{M}_{mn}(E_{\text{DB}})|^2, \end{aligned} \quad (\text{A9})$$

where Ω_{DB} is the volume of the DB, defined in our case as the region in space where the DB charge density is greater than 1% of its maximum. A completely similar equation can be written for the tunneling between DB and Si just by replacing the subscript ‘‘tip’’ above with ‘‘Si’’. In this latter case, \tilde{M}_{mn} depends on the tunneling barrier between the DB and CB states lying at the same energy, which is given by the form of the band bending in the Si region surrounding the DB. Using the semiclassical approximation, we assume the tails of the wave functions for states in the conduction band in the barrier region scale as $\varphi_{\text{CB}}(z) = \exp(-\zeta_{\text{CB}}z)$, with $\zeta_{\text{CB}} = \sqrt{2m\bar{U}_{\text{bb}}/\hbar}$ and \bar{U}_{bb} the average tunneling barrier height between DB and CBM.

For all cases described above, the tunneling matrix elements were calculated numerically by assuming the corresponding wave functions given in the text.

APPENDIX B: CALCULATING BAND BENDING IN THE VICINITY OF A DB DURING STM IMAGING

Tip-induced band bending was calculated within the semiclassical approximation using finite element methods (FEM). We assume that the STM tip has an axially symmetric shape ending in a hemispherical surface of radius $R_{\text{tip}} = 20$ nm. On top of this surface, we assume the existence of an atomic-sized protrusion of radius R_0 , which coincides with the spherical tip radius assumed in the Tersoff-Hamann model for the tip wave function.

For an STM system, the total potential drop between the tip and the sample is

$$(V_{\text{tip}} - V_{\text{sam}}) = V_{\text{bias}} + \Phi_{\text{tip}} - \chi - (E_{\text{CBM}} - E_{\text{F}}), \quad (\text{B1})$$

where Φ_{tip} and χ are the tip work function and electron affinity of Si, respectively. Another component of U_{DB} can arise in the vicinity of a charged DB from the field created by the localized electron.

The calculation involves solving a system of nonlinear equations consisting of the Poisson and Schrödinger (PS) equations for the electrostatic potential U and the charge densities (n, p) of electrons and holes, respectively, in the semiconductor. Poisson equation reads

$$\nabla^2 U(\mathbf{r}) = -\frac{e}{\epsilon_{\text{Si}}} [N_D^+ + p(\mathbf{r}) - N_A^- - n(\mathbf{r})], \quad (\text{B2})$$

where e is the elementary charge, ϵ_{Si} is the electric permittivity of the semiconductor, N_D^+ the ionized donor concentration, and N_A^- the ionized acceptor concentration. The same U that satisfies the above equation, must also satisfy the Schrödinger equation

$$-\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}) + U(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad (\text{B3})$$

where, in our case, ψ are extended states of the crystal belonging to the CB or VB bands, and their corresponding charge distribution determines n and p .

In the semiclassical approximation, the extended states ψ entering the above equation are affected by a local potential just by a rigid shift in their energies, which equals to $-eU(\mathbf{r})$. As a consequence, the local charge densities are shown to be proportional to the (1/2)-Fermi-Dirac integrals, $F_{1/2}$:

$$\begin{aligned} n(\mathbf{r}) &= N_c F_{1/2} \left\{ [E_F - E_{\text{CBM}}^{\text{bulk}} + eU(\mathbf{r})] / k_B T \right\}, \\ p(\mathbf{r}) &= N_v F_{1/2} \left\{ [E_{\text{VBM}}^{\text{bulk}} - eU(\mathbf{r}) - E_F] / k_B T \right\}, \end{aligned} \quad (\text{B4})$$

where E_F is the Fermi level (chemical potential) of the crystal. The above PS equations were solved self-consistently for the tip-induced band bending (no charge present on a DB) using an iterative FEM scheme in which the solution to Eq. (B4) was input as a source term in the RHS of Poisson Eq. (B2). To ensure convergence of the iterations, the Anderson mixing scheme was used.

In order to calculate the distortions in the bands caused by the field of a DB^- , we performed a second FEM calculation in the presence of the charged DB. For practical reasons, we did not employ FEM on a three-dimensional geometry (for the case when the DB is laterally displaced from the central axis of the tip). Such a calculation is not practically feasible for the following reasons: (i) the size of the finite element domain must extend very deep into the semiconductor bulk in order to ensure that boundary condition there (zero electrostatic potential) does not introduce nonphysical artifacts in the results, (ii) a very fine mesh is needed for the tip vicinity where the field is very strong in contrast to the bulk semiconductor where the field is relatively weak, and (iii) many PS iterations are required to reach convergence, depending on the biasing conditions and doping level in the semiconductor.

We found a way to approximate the full solution of the PS system in the presence of a charged DB by taking advantage of the fact that the STM tip has a radius much greater than the size of the DB, and therefore it can be seen as locally flat in the vicinity of the DB. This implies that the potential due to the DB^- alone, $U_{\text{DB}^-}(\mathbf{r})$, will be approximately the same (as

measured from the center of the DB) whether the DB is directly under the tip apex or slightly displaced laterally by a few nm. This allows us to extract $U_{\text{DB}^-}(\mathbf{r})$ as the difference between two different FEM solutions for two axially symmetric systems: (i) STM tip alone and (ii) STM tip with a DB^- directly under the apex. Then, for the case when the DB is not on the central axis of the tip but close to it, the solution is well approximated by adding the potential $U_{\text{DB}^-}(\mathbf{r} - \mathbf{r}_{\text{DB}})$ to the tip-induced potential. More details of the finite element calculation will be published in an upcoming paper.²⁶

The calculation of the charge carrier densities at the surface including charge quantization in the surface-perpendicular direction was also performed. We numerically solved the PS equations self-consistently for a planar metal-vacuum-semiconductor system in a one-dimensional case, assuming translational symmetry in xy plane in which case the PS equations need only to be solved in the z direction. This approximation is justified by the fact that the surface-parallel length-scale of the sample in which tip-induced band bending is manifest is on the order of hundreds of nanometers, and much greater than the surface-normal length scale over which the hole density is significant (tens of nanometers). In other words, the confinement of the holes in the potential well created by the biased STM probe in the surface-parallel direction is much weaker than the confinement in the surface-normal direction. Charge quantization was also estimated using a 1D Poisson solution³⁵ corresponding to the limit of a very large tip radius (or flat tip); those results were consistent with ours.

APPENDIX C: STM VICINITY CAPTURE OF TUNNELING ELECTRONS BY A DB

The rate of capture for a deep-level state is proportional to the velocity of the electron being captured,³⁶ $c_n = \sigma_n v_n n$. In equilibrium, at finite temperature, electrons sit at the bottom of the conduction band, and their energy distribution is well described by Maxwell-Boltzmann statistics, as long as $E_{\text{CBM}} - \mu_{\text{sam}} \gg k_B T$. This allows the capture rate to be described by the familiar equation $c_n = \sigma_n v_{\text{th}} n$, expressed in terms of the thermal velocity. The thermal velocity is the average group velocity for all electrons in the conduction band. Assuming equilibrium statistics and a parabolic conduction band minimum, the thermal velocity is given by $v_{\text{th}} = \sqrt{8k_B T / \pi m^*} \approx 10^7$ cm/s.

However, in calculating the excess capture rate, c_n^{STM} , due to electrons injected from the tip into the conduction band, we no longer have recourse to Maxwell-Boltzmann statistics. We therefore use the group velocities directly. For simplicity, we retain the assumption of a parabolic band, so that $E_{\mathbf{k}} = \hbar^2 k^2 / 2m$. In addition to simplifying calculations, this ensures that group velocities are parallel to their associated wave vectors.

The injected electrons can tunnel into eigenstates in the conduction band with energies below μ_{tip} , however, we expect that tunneling will occur most readily into states with high k_z (surface normal) and low \mathbf{K} (surface-parallel) values. Thus their velocities are not evenly distributed across the range of polar angles, θ . We solve for the angular distribution of tip-injected electrons by considering the matrix element,

$|M_{mn}(\mathbf{k})|^2$, for tunneling from tip states to sample states with wave vector, \mathbf{k} , described by Eq. (A6). The angular distribution of wave vectors, and hence also of velocities, of the injected electrons is then given by

$$D(\theta) = \mathcal{N} \int_0^{\sqrt{2m(\mu_{\text{tip}} - E_{\text{CBM}})/\hbar}} k^2 |M_{mn}(\mathbf{k})|^2 dk, \quad (\text{C1})$$

where

$$\mathcal{N} = \frac{2\pi}{\int_0^{2\pi} \int_0^{\pi/2} \int_0^{\sqrt{2m(\mu_{\text{tip}} - E_{\text{CBM}})/\hbar}} k^2 |M_{mn}|^2 \sin\theta dk d\theta d\phi}. \quad (\text{C2})$$

Ultrafast pump-probe reflectivity measurements of the Si(001) surface place the momentum relaxation time of free carriers at 32 fs.³⁷ Electrons traveling with speed v_{th} (much slower than the average velocity for injected electrons) will travel a distance of roughly 6 nm in this time. We therefore make the approximation that injected electrons retain their initial group velocities over the distance scales relevant to the

present problem. The local excess electron density at each point in the silicon sample, due to the injected current from the STM tip when the DB is neutral, $I_{\text{tip-Si}}^{\text{DB}^0}$, is then given by

$$n^{\text{STM}}(\mathbf{r}) = D(\theta) \frac{I_{\text{tip-Si}}^{\text{DB}^0}}{2\pi r^2 v}, \quad (\text{C3})$$

where $\mathbf{r} = (r, \theta, \phi)$ being the distance vector from the STM tip apex with the polar angle θ being measured from the tip axis pointing toward the silicon crystal.

Finally, we account for the occupation of the DB with one electron, $1 - f_{\text{DB}}^*$, and write the capture current by the DB level as

$$C_n^{\text{STM}} = (1 - f_{\text{DB}}^*) \int_{\Omega_{\text{DB}}} d\mathbf{r} |\psi_{\text{DB}}^0(\mathbf{r})|^2 D(\theta) I_{\text{tip-Si}}^{\text{DB}^0} \frac{\sigma_n}{2\pi r^2}, \quad (\text{C4})$$

where we assume that the capture cross section for an infinitesimal volume is $|\psi_{\text{DB}}^0|^2 \sigma_n d\mathbf{r}$. The capture rate c_n^{STM} in s^{-1} is obtained by dividing C_n^{STM} by the elementary charge and assuming a neutral DB, i.e., $f_{\text{DB}}^* = 0$.

*lucian@ualberta.ca

¹H. Raza, *Phys. Rev. B* **76**, 45308 (2007).

²T. Blomquist and G. Kirzenow, *Nano Lett.* **6**, 61 (2006).

³M. B. Haider, J. L. Pitters, G. A. DiLabio, L. Livadaru, J. Y. Mutus, and R. A. Wolkow, *Phys. Rev. Lett.* **102**, 46805 (2009).

⁴T. Yoshida and H. Hasegawa, *Jpn. J. Appl. Phys.* **39**, 4504 (2000).

⁵M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian 03, Revision C.02. Gaussian, Inc., Wallingford, CT, 2004.

⁶K. Teichmann, M. Wenderoth, S. Loth, R. G. Ulbrich, J. K. Garleff, A. P. Wijnheijmer, and P. M. Koenraad, *Phys. Rev. Lett.* **101**, 76103 (2008).

⁷D. H. Lee and J. A. Gupta, *Science* **330**, 1807 (2010).

⁸J. L. Pitters, L. Livadaru, M. B. Haider, and R. A. Wolkow, *J. Chem. Phys.* **134**, 4712 (2011).

⁹Ph. Ebert, *Surf. Sci. Rep.* **33**, 121 (1999).

¹⁰M. Berthe, R. Stiuflu, B. Grandidier, D. Deresmes, C. Delerue, and D. Stievenard, *Science* **319**, 436 (2008).

¹¹B. Grandidier, X. de La Broise, D. Stievenard, C. Delerue, M. Lannoo, M. Stellmacher, and J. Bourgoin, *Appl. Phys. Lett.* **76**, 3142 (2009).

¹²T. H. Nguyen, G. Mahieu, M. Berthe, B. Grandidier, C. Delerue, D. Stievenard, and Ph. Ebert, *Phys. Rev. Lett.* **105**, 226404 (2010).

¹³L. Liu, J. Yu, and J. W. Lyding, *Mater. Res. Soc. Symp. Proc.* **705**, Y6.6.1 (2002).

¹⁴W. A. Harrison, *Surf. Sci.* **55**, 1 (1976).

¹⁵J. Repp, G. Meyer, F. E. Olsson, and M. Persson, *Science* **305**, 493 (2004).

¹⁶J. Repp, G. Meyer, S. Paavilainen, F. E. Olsson, and M. Persson, *Phys. Rev. Lett.* **95**, 225503 (2005).

¹⁷J. Repp and G. Meyer, *Appl. Phys. A* **85**, 399 (2006).

¹⁸S. W. Wu, G. V. Nazin, X. Chen, X. H. Qiu, and W. Ho, *Phys. Rev. Lett.* **93**, 236802 (2004).

¹⁹G. Mikaelian, N. Ogawa, X. W. Tu, and W. Ho, *J. Chem. Phys.* **124**, 131101 (2006).

²⁰J. Oh, V. Meunier, H. Ham, and R. J. Nemanich, *J. Appl. Phys.* **92**, 3332 (2002).

²¹F. Mohn, J. Repp, L. Gross, G. Meyer, M. S. Dyer, and M. Persson, *Phys. Rev. Lett.* **105**, 266102 (2010).

²²M. Wenderoth, M. A. Rosentreter, K. J. Engel, A. J. Heinrich, M. A. Schneider, and R. G. Ulbrich, *Europhys. Lett.* **45**, 579 (1999).

²³C. R. Leavens and G. C. Aers, *Phys. Rev. B* **38**, 7357 (1988).

²⁴J. M. Moison and M. Bensoussan, *Rev. Phys. Appl.* **22**, 293 (1987).

²⁵D. Goguenheim and M. Lannoo, *Phys. Rev. B* **44**, 1724 (1991).

²⁶L. Livadaru and R. A. Wolkow (unpublished).

²⁷R. M. Feenstra, *J. Vac. Sci. Technol. B* **21**, 2080 (2003).

²⁸L. Livadaru, P. Xue, Z. Shaterzadeh-Yazdi, G. A. DiLabio, J. Mutus, J. L. Pitters, B. C. Sanders, and R. A. Wolkow, *New J. Phys.* **12**, 083018 (2010).

²⁹N. H. March and A. M. Murray, *Proc. R. Soc. A* **261**, 119 (1961).

³⁰N. H. March and A. M. Murray, *Proc. R. Soc. A* **266**, 559 (1962).

³¹N. H. March and J. C. Stoddart, *Rep. Prog. Phys.* **31**, 533 (1968).

- ³²T. E. Feuchtwang, P. H. Cutler, and N. M. Miskovsky, *Phys. Lett. A* **99**, 167 (1983).
- ³³J. A. Appelbaum and W. F. Brinkman, *Phys. Rev.* **186**, 464 (1969).
- ³⁴J. Bardeen, *Phys. Rev. Lett.* **6**, 57 (1961).
- ³⁵I. Tan, G. L. Snider, L. D. Chang, and E. L. Hu, *J. Appl. Phys.* **68**, 4071 (2009).
- ³⁶W. Shockley and W. T. Read Jr., *Phys. Rev.* **87**, 835 (1952).
- ³⁷A. J. Sabbah and D. M. Riffe, *Phys. Rev. B* **66**, 165217 (2002).