Editors' Suggestion

Spectroscopy of Multielectrode Tunnel Barriers

Amir Shirkhorshidian,[1,2,*] John King Gamble,[3] Leon Maurer,[4] Stephen M. Carr,[2] Jason Dominguez,[2] Gregory A. Ten Eyck,[2] Joel R. Wendt,[2] Erik Nielsen,[4] Noah Tobias Jacobson,[4] Michael P. Lilly,[2,5] and Malcolm S. Carroll[2,†]

[1] *University of New Mexico, Albuquerque, New Mexico 87131, USA*
[2] *Sandia National Laboratories, Albuquerque, New Mexico 87185, USA*
[3] *Quantum Architecture and Computation Group, Microsoft Research, Redmond, Washington 98052, USA*
[4] *Center for Computing Research, Sandia National Laboratories, Albuquerque, New Mexico 87185, USA*
[5] *Center for Integrated Nanotechnologies, Sandia National Laboratories, Albuquerque, New Mexico 87185, USA*

(Received 8 May 2017; revised manuscript received 18 June 2018; published 1 October 2018)

Despite their ubiquity in nanoscale electronic devices, the physics of tunnel barriers has not been developed to the extent necessary for the engineering of devices in the few-electron regime. This problem is of urgent interest, as this is the specific regime into which current extreme-scale electronics fall. Here, we propose theoretically and validate experimentally a compact model for multielectrode tunnel barriers, suitable for design-rules-based engineering of tunnel junctions in quantum devices. We perform transport spectroscopy at approximately $T = 4$ K, extracting effective barrier heights and widths for a wide range of biases, using an efficient Landauer-Büttiker tunneling model to perform the analysis. We find that the barrier height shows several regimes of voltage dependence, either linear or approximately exponential. Effects on threshold, such as metal-insulator transition and lateral confinement, are included because they influence parameters that determine barrier height and width (e.g., the Fermi energy and local electric fields). We compare these results to semiclassical solutions of Poisson's equation and find them to agree qualitatively. Finally, this characterization technique is applied to an efficient lateral tunnel barrier design that does not require an electrode directly above the barrier region in order to estimate barrier heights and widths.

DOI: 10.1103/PhysRevApplied.10.044003

I. INTRODUCTION

A number of device architectures that go beyond Moore's law rely on dense packing of tunnel-barrier-based devices [e.g., tunnel field-effect transistor (FET) logic [1], quantum-dot-based quantum computing [2], quantum-dot cellular automata [3] and single-electron transistor logic [4]]. Several major challenges must be overcome to realize device architectures that use tunnel barriers as a foundational element. One of those foundational challenges is to develop rapid characterization techniques and models that can predict behavior in these densely packed layouts for which "cross talk" of voltages plays a role. For example, cross-coupling of neighboring electrodes is likely an effect that will dramatically affect tunnel-based devices much more strongly than in other previous device architectures because of potentially nonlinear or exponential dependences of currents on neighboring electrode voltages.

Although the modeling of tunnel barriers has a long history, there is little available for modeling problems of many tunnel barriers in close proximity, which (1) clarify the effects of neighboring electrodes or (2) are efficient in handling the potential complex interplays between neighboring multielectrode voltages. Success for extending designs of tunnel-barrier devices, even to the immediately relevant experimental demonstrations with small numbers of devices, requires an efficient compact design and analysis capability that does not depend on detailed microscopic modeling. A compact model foundation analogous to that used for devices in SPICE circuit modeling will be a necessary future step.

The goal of this paper is to identify an efficient root model that can be used to fit and describe the tunnel-barrier behavior over (1) a wide range of voltages on a single electrode and (2) also fits the tunnel-barrier I–V response

*ashirkho@unm.edu
†mscarro@sandia.gov

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

when multiple electrode voltages are changed. The intent is to produce a compact model that can be used to evaluate much more complex systems of multielectrode tunnel barriers, similar in spirit to simplified device models used for transistor modeling for larger circuits (e.g., SPICE). This includes significant improvement in the efficiency of the computation and clarification of the tractability of the multielectrode response of the tunnel-barrier solutions. This is well beyond the scope or intent of the traditional one-dimensional (1D) “fixed barrier height” textbook models, such as the Wentzel-Kramers-Brillouin (WKB) method or Fowler-Nordheim tunneling.

The intent of this work is also not to extract detailed microscopic information that would necessarily come from much more detailed three-dimensional (3D) Poisson-Schrödinger combined with nonequilibrium Green’s function solutions. These are important for learning and feedback loops on single barriers but are computationally very expensive and therefore impractical for device layouts with larger numbers of tunnel barriers and multidevice functionality.

Observations of linear dependence of barrier height on voltage have been conjectured for some geometries [5] but it is not clear under what conditions this holds or over how wide a voltage range. In this work, we examine simple 1D phenomenological models of a tunnel barrier that use two parameters, a barrier height and width, as a starting point to describe a multielectrode tunnel-barrier design. We use tunneling spectroscopy at 4 K to experimentally characterize the voltage dependence of a lateral electrostatically gated metal-oxide-semiconductor (MOS) tunnel barrier. The lateral electrostatically gated MOS nanostructure is itself of interest as a model system that applies abstractly to many other systems while also being directly informative to the ubiquitous MOS system and its potential application to quantum information.

We examine the models in the context of an ansatz that the barrier height and width are functions that depend independently on each electrode voltage in the lateral gate design. We identify key assumptions about previous works that are inappropriate for contemporary tunnel-based devices. We remedy these deficiencies by using the modern Landauer-Büttiker formalism (which is intrinsically quantum mechanical rather than semiclassical) and by accounting for other effects, not considered previously, that are important for cryogenically operated tunnel barriers, such as metal-insulator transition.

We show in this work (1) that the 1D model can describe this multielectrode tunnel barrier and might therefore act as the foundation for extending to more complex networks of tunnel barriers and (2) that a regime exists where the barrier height scales exponentially with applied voltages when an electron channel forms below a gate, in contrast to the usually assumed linear dependence.

This paper is organized as follows. In Sec. II we describe the device fabrication and transport measurements of a relatively disorder-free MOS multielectrode tunnel barrier. This tunnel-barrier layout is one that has been used successfully for MOS quantum-bit (qubit) demonstrations [6]. The barrier is therefore a representative case for which a compact model would be desirable (e.g., this is an example qubit layout that requires the simultaneous tuning of multiple cross-coupled barriers).

We next consider that the tunneling transmission is sensitive to the wave number of the electrons at the Fermi level. Thus in Sec. III we present a calculation of the Fermi energy. This includes a discussion of the threshold, the metal-insulator transition, and the lateral quantum confinement to obtain more accurate estimates of the Fermi energy and electrostatics of the nanostructure.

Section IV describes the 1D barrier model. A numeric approach is presented that offers both an efficient and accurate solution (i.e., it avoids unjustified assumptions that are present in common workhorse models such as the WKB and Fowler-Nordheim). A second simple analytic model is described and compared to the numeric model. The analytic model can provide more intuitive insight and is easily implementable (e.g., back of the envelope). Good agreement is observed in the comparison, which provides a measure of confidence and guidance about when this simpler model can be applied.

Then, in Sec. V, we discuss the results of fitting the 1D barrier model to the data, including extraction of barrier heights and widths. The fits produce good agreement with the measured currents over a wide bias range and for multiple electrode voltages, achieving a proof-of-principle demonstration that the 1D compact model could be used to describe this multielectrode barrier.

Finally, in Sec. VI, we further compare these results to electrostatic simulations and lithographic dimensions to better understand whether the quantitative barrier heights and widths are reasonable. We find that the extracted values are qualitatively in agreement and semiquantitatively accurate.

II. THE EXPERIMENTAL APPROACH

The tunnel barriers are formed in a metal-oxide-semiconductor (MOS) field-effect-transistor (FET) inversion channel at a Si/SiO_2 interface. The starting material is a p-type Si substrate produced by a float-zone process and with a doping concentration of roughly 10^{14} cm^{-3} boron. The structures are made by first forming heavily doped $n+$ ohmics using ion implantation of As. Next, a 35-nm SiO_2 gate dielectric is grown by thermal oxidation at 900°C. Degenerately doped n-type poly-Si is formed over the SiO_2 to complete the MOS gate stack. The gate structures are patterned using electron-beam lithography and a dry etch, using a fabrication process described in more

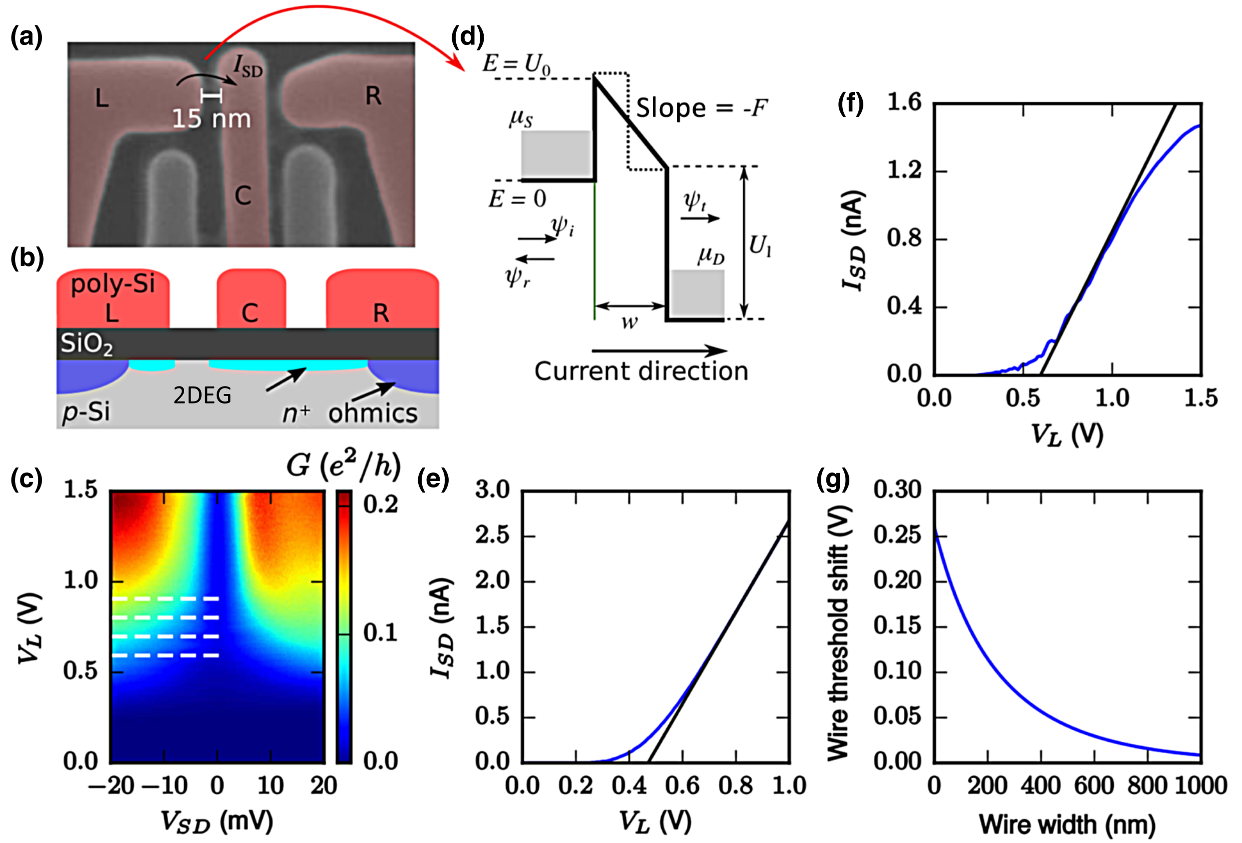

FIG. 1. The device geometry and threshold measurements. (a) A false-color scanning-electron-microscope (SEM) image of the measured device. Only the red-colored gates (L, C, and R) are used; all other gates are grounded. (b) A cross-section schematic of the device (not to scale). The device stack includes a p-Si substrate (light gray), 35 nm thermally grown SiO_2 (dark gray), and 200 nm poly-Si gates (red). The L and R gates spread out to overlap n^+ ohmics (purple). The channel or two-dimensional electron gas (2DEG) is formed at the interface (light blue). (c) The differential conductance for $V_R = 2.13$ V and $V_C = 1.60$ V. The dashed lines correspond to the line cuts shown in Fig. 3(a). (d) The energy diagram as described in the text, showing the conduction-band minimum energy vs distance along the direction of current flow. (e) The measured current in the field of the device, showing a field threshold of 0.48 V. (f) The measured current through the nanostructure for $V_R = 2.13$ V and $V_C = 5.00$ V, giving a threshold of 0.60 V. The V_C value is chosen because it is representative of voltages above approximately 2.00 V. The threshold decreases slightly for $V_C < 2.00$ V. (g) The calculated threshold shift as a function of the wire width for a field threshold of 0.48 V. Lateral quantum confinement is predicted to produce a threshold shift. Details of the calculation are found in Appendix B.

detail elsewhere [7]. A scanning-electron- microscope (SEM) image of the nanostructure gates and a cross-section schematic view of the device are shown in Figs. 1(a) and 1(b). As seen in Fig. 1(a), the relevant gates are separated from each other by approximately 15 nm. The inversion channel is formed under the poly-Si enhancement gates and the barriers are formed at gaps between the poly-Si gates.

We measure the differential conductance $G = dI/dV_{SD}$ using a lock-in amplifier with an ac excitation voltage of 100 μV at a modulation frequency of 149 Hz in combination with a dc source-drain bias V_{SD}. We also monitor the dc source-drain current I_{SD} using a digital multimeter. All measurements are performed at approximately 4 K by immersing the sample in liquid He at an atmospheric pressure of approximately 625 torr. A plot of the differential conductance as a function of V_L and V_{SD} for $V_R = 2.13$ V and $V_C = 1.60$ V is shown in Fig. 1(c). Disorder such as charge defects in the gate oxide [8], impurities [9], or strain from the gate stack [10] can cause subthreshold resonances in these plots. The absence of resonances suggests a single tunnel barrier between the source and drain leads. We also note that the combination of the contact resistance and sheet resistance of the channel leads to approximately 10 kΩ, which limits the maximum conductance.

III. THE CALCULATION OF THE FERMI ENERGY

To calculate the tunneling current, we need to estimate the voltage dependence of the Fermi energy in the leads. The conduction-band energy at low temperature is

shown in Fig. 1(d). A high-density electron channel forms when gates L, C, and R are biased above threshold. If we decrease the bias on one of the gates below threshold, electrons are locally depleted and the conduction-band edge rises above the Fermi level, forming a tunnel barrier with source and drain leads to the left and right of the barrier. The extent to which the Fermi level in the leads is above the conduction-band edge is proportional to the electron density. Assuming a two-dimensional (2D) density of states (DOS), the difference between the Fermi energy E_f and the conduction-band edge E_C is

$$E_f - E_C = \frac{\pi \hbar^2 n}{m_t g_v}, \tag{1}$$

where n is the 2D electron density, g_v is the valley degeneracy, and m_t is the electron transverse effective mass. Spin degeneracy, $g_s = 2$, is already assumed. We use a valley degeneracy $g_v = 2$ and an effective mass of $m_t = 0.19 m_0$, appropriate for 2D electron inversion layers in Si(100) [11], where m_0 is the free-electron mass.

We first estimate the electron-density dependence on voltage in the field of the device. In the field, we exclude effects due to the nanoscale gates, such as fringing fields and lateral quantum confinement (i.e., we assume a bulk poly-Si gate). We identify a field threshold of 0.48 V, Fig. 1(e). At cryogenic temperatures, there is a metal-insulator transition for which there is a critical density of electrons necessary before appreciable conduction starts. We measure a percolation-density threshold of $n_p = 3.1 \times 10^{11}$ cm^{-2} at approximately 4 K in a Hall bar with a nominally similar oxide-silicon interface as used in this tunnel-barrier experiment. We estimate the critical density with the percolation density [12]. Assuming that $n = n_{\text{crit}}$ at $V_{\text{gate}} = V_{\text{th}}$, where V_{th} is the threshold voltage, then the electron density as a function of the gate voltage is

$$n\left(V_{\text{gate}}\right) = \frac{C_{\text{ox}}}{e}(V_{\text{gate}} - V_{\text{th}}) + n_{\text{crit}}, \tag{2}$$

where C_{ox} is the oxide capacitance per unit area and e is the electron charge. The electron density and Fermi energy can be calculated in the field using Eqs. (2) and (1). The accounting of the metal-insulator transition accommodates fitting over the full range of voltages, including the subthreshold region, using a simple 1D capacitance model to approximate the density in the leads.

The electron density in the nanostructure will differ from that in the field due to several effects: fringing fields, lateral quantum confinement, and possibly 2D to 1D weak localization or disorder effects in the confined region. The electrostatic effects of the narrowed lead are directly calculated. We calculate electrostatic modifications of threshold due to fringing fields numerically. For these dimensions, the effects are small, less than 10%. We neglect them in the compact model. The effect of quantum confinement is, however, more significant [13].

We calculate the confinement threshold shift for this geometry over various wire widths [see Fig. 1(g) and Appendix B]. We measure a threshold of 0.60 ± 0.05 V for the nanostructure (i.e., $V_R = 2.13$ V, $V_C = 5.00$ V, and V_L is increased). This threshold is the linear extrapolation to zero current [Fig. 1(f)]. The threshold shift falls within a range that would be expected for the approximately 70 nm width of the tunnel-barrier lead. The electron density in the leads can, therefore, be estimated using the field threshold combined with an offset from quantum confinement. Disorder in 1D also can affect $n = n_{\text{crit}}$. Any increase in the local potential of a 1D wire represents a direct barrier to transport and a source of enhanced weak localization. These effects can be challenging to directly quantify in practical devices, which is one of the goals of this work (i.e., a compact model for rapid characterization of devices for future computation). Since disorder and scattering are also effects related to transmission through a 1D barrier, for this compact model approach, we combine these effects as part of the extracted barrier height and the width of a single barrier.

IV. THE THEORETICAL MODEL

We now examine how well a 1D barrier model and a voltage-dependent parametrization of the barrier height and width fits the observed tunneling for multiple bias conditions on different electrodes. At zero source-drain bias, we assume a rectangular barrier with a width w, a left barrier height U_0, and a right barrier height U_1. The two barrier heights allow for asymmetric plateaus in the leads. When a source-drain bias V_{SD} is applied, the chemical potentials or quasi-Fermi levels in the source μ_S and drain μ_D separate by an amount proportional to V_{SD}. The potential in the barrier region then varies linearly with a slope equal to F, where F is the electric field due to the source-drain bias, as shown in Fig. 1(d). Thus, for nonzero source-drain bias, the wave functions in the barrier region are Airy functions. In the leads, the wave functions are propagating plane waves with an incident wave ψ_i, a reflected wave ψ_r, and a transmitted wave ψ_t.

We use the Landauer-Büttiker formalism to model the tunneling transport. Assuming low temperature, the Fermi functions in the source and drain leads can be approximated as step functions and the current through the barrier is given by [14]

$$I = \int_{\mu_0}^{\mu_S} dE \frac{2e}{h} T(E) M(E). \tag{3}$$

Here, $T(E)$ is the transmission coefficient of an electron with energy E, $M(E)$ is the total number of available transverse modes, which depends on the DOS in the leads, and

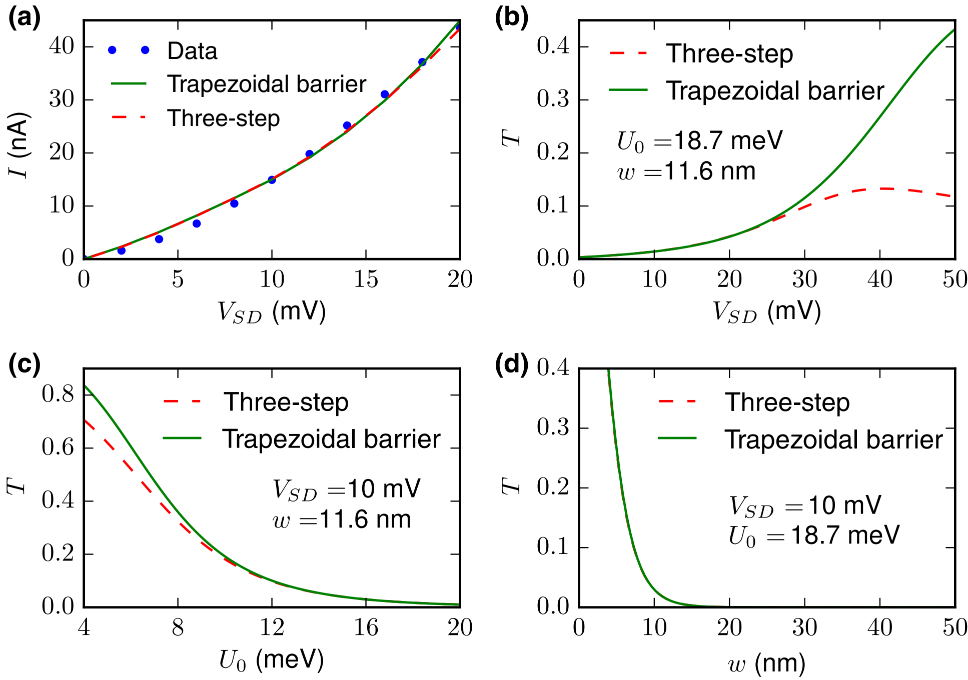

FIG. 2. A comparison between the trapezoidal and three-step barrier models. (a) Line-cut data from Fig. 1(c) for $V_L = 0.70$ V together with a fit using the trapezoidal barrier model. The extracted fit parameters are used to calculate the current using the three-step model. We then compare the resulting transmission coefficients T as a function of (b) V_{SD}, (c) U_0, and (d) w.

$\mu_0 = \max(\mu_D, 0)$. Thus, the product $T(E)M(E)$ gives the total transmission summed over the transverse modes.

We examine both a numeric and an approximate analytic solution to computing the transmission. Both are sufficiently efficient to find good fits for barrier height and width for each gate voltage. The analytic solution provides an easy reference for "back-of-the-envelope" estimates and quick insights in the laboratory, while showing very good agreement with the more accurate numeric model over a wide range of voltages. The numeric approach, is however, used primarily in this work to extract barrier heights and widths because it is more accurate. The numeric approach uses the analytic piecewise solution to the Schrödinger equation and then numerically solves the boundary-matching problem as a system of equations. This approach gives us an exceptionally fast forward solution (contrasted against fully numeric approaches) to facilitate nonlinear inference over a large experimental data set. For more details, see Appendix C.

As noted above, reasonably good quantitative agreement can also be provided through solving the transmission coefficient by approximating the linearly varying potential in the barrier region with two potential steps [see Fig. 1(d)]. One step has a fixed height of U_0 while the other has a variable step height that depends on V_{SD}. The transmission coefficient is then solved using transfer matrices and yields a simple analytic formula for transmission that is a good approximation to the trapezoidal barrier problem over a wide parameter range. More details are provided in Appendix D. We find that both 1D models can match the measured current dependences on voltage [Fig. 2(a)]. We further find that the three-step model agrees well with the full numeric solution for lower V_{SD} relative to the barrier U_0 [Figs. 2(b) and 2(c)]. For the rest of this paper, we will use the numeric solution to the trapezoidal model for the more accurate quantitative analysis.

The three-step model and, in particular, the trapezoidal barrier model are useful alternatives to traditional approximate models of tunneling phenomena, such as the WKB approximation [15] and Fowler-Nordheim tunneling [16,17]. In particular, this approach avoids unnecessary and often unjustified assumptions by exactly solving the Schrödinger equation and more accurately accounting for the lead geometry. Over the course of a numeric fit, the assumptions of these approximations can easily become violated, invalidating the parameter extraction. Fowler-Nordheim, for example, is only appropriate for large electric fields relative to the barrier but will not be satisfactory over the range of voltages fitted in this work. Both our numeric solution to the trapezoidal barrier and the three-step model are simple enough to allow rapid estimates of the barrier height and width from transport measurements using a modern computer. We also note that this analysis applies to other semiconducting systems (e.g., holes and Si/SiGe heterostructures) although the parameters (e.g., the effective mass) and the energy-band diagram may require modification.

V. RESULTS AND DISCUSSION

Figure 3(a) shows line cuts at $V_L = 0.90$, 0.80, 0.70, and 0.60 V from the data in Fig. 1(c). We treat the barrier height U_0 and width w as free parameters and fit the trapezoidal barrier model to the data by minimizing a

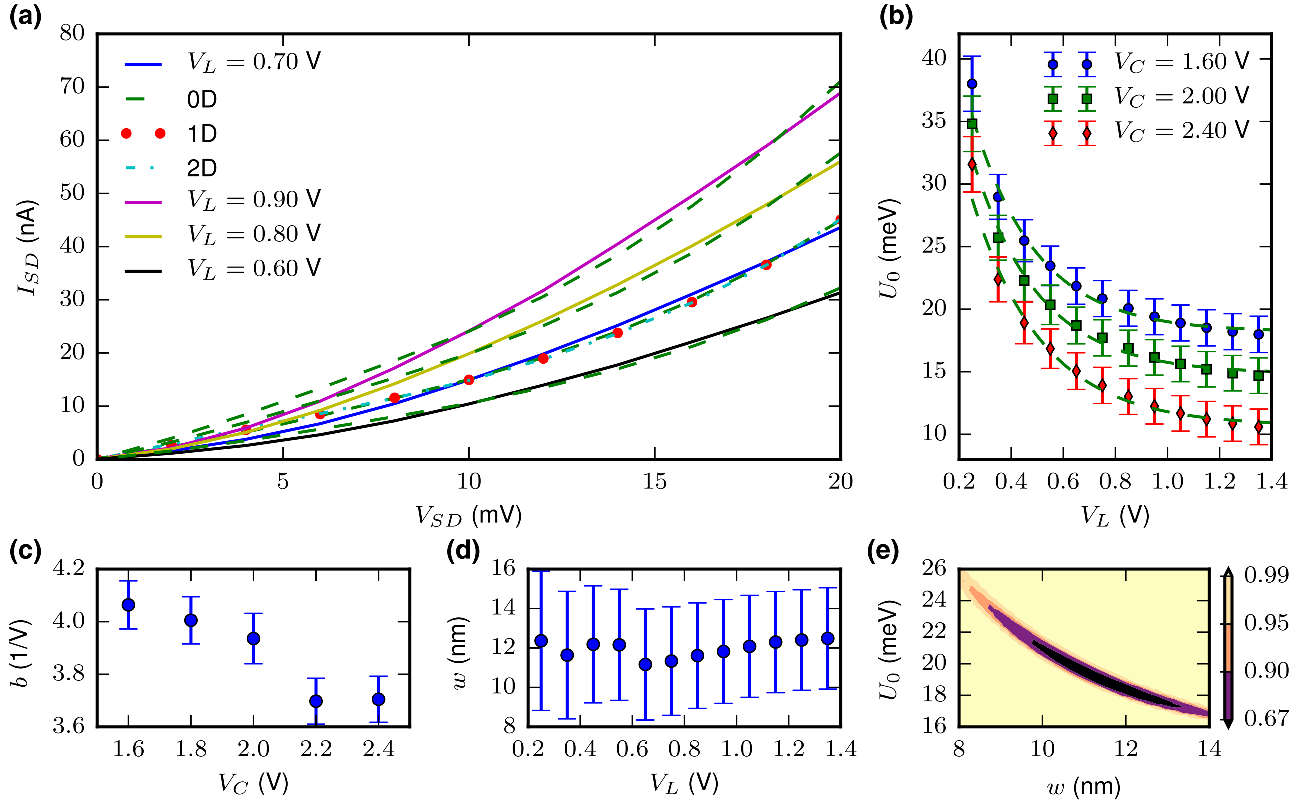

FIG. 3. Parameter extraction from the experimental data. (a) Line-cut data from Fig. 1(c) together with fits using the semianalytic model. For $V_L = 0.70$ V, we compare zero-dimensional (0D), 1D, and 2D DOS in the leads. Note that the data are plotted as the absolute value of the dc source-drain current I_{SD} versus the absolute value of V_{SD}. (b) The extracted barrier height U_0 as a function of the gate V_L for a fixed width $w = 12$ nm. The dashed lines are fits to an exponential function. The curves are artificially offset vertically to allow each case to be viewed independently. For $V_C = 1.60$ V, the offset is +3 meV and for $V_C = 2.40$ V the offset is −3 meV. There is a high degree of orthogonal control of the barrier, depending strongly on L and weakly on C. (c) The extracted exponent from the fits shown in (b) as a function of V_C. The average exponent is 3.9 V^{-1}. (d) The extracted barrier width w as a function of the gate V_L for a nonfixed U_0. (e) A plot of the confidence region as a function of w and U_0 for $V_L = 0.70$ V. The contours correspond to 67%, 90%, 95%, and 99% confidence levels.

chi-squared test statistic, resulting in maximum-likelihood estimates (MLEs) of the parameters under the assumption of independent and identically distributed Gaussian errors. These fits are shown in Fig. 3(a). The MLEs and expected errors for the parameters are shown in Table I.

The expected errors are computed by constructing the profile likelihood function (separately, for each parameter) and numerically determining a confidence level of 95%. To do this, we make the standard assumption that the difference in the profile likelihood function from the MLE value should be $\chi^2(1)$ distributed. Unless otherwise noted, we assume throughout this paper that the experimental standard error in the observed currents is $\delta I = 0.1 \cdot \max(I_{SD})$, where $\max(I_{SD})$ is the maximum source-drain current for a given gate voltage. From Table I, we see that the barrier width seems to be fixed at about 12 nm while the barrier height and its uncertainty increase as V_L decreases.

TABLE I. The extracted barrier parameters U_0 and w for the fits shown in Fig. 3(a).

V_L (V)	DOS of leads	U_0 (meV)	w (nm)
0.90	0D	17 ± 3	12 ± 3
0.80	0D	18 ± 4	11 ± 3
0.70	0D	19 ± 5	12 ± 3
0.70	1D	27 ± 8	12 ± 3
0.70	2D	29 ± 9	12 ± 3
0.60	0D	20 ± 5	12 ± 3

We repeat the fitting procedure for different V_L values. Figure 3(b) shows a plot of the extracted barrier height as a function of V_L for a fixed width of 12 nm and three values of V_C. In this case, we assume that $\delta I = 0.3 \cdot \max(I_{SD})$ because the errors are found to be smaller for a one-parameter fit. We find that at high V_L the barrier height varies linearly with the gate voltage. However, for low V_L the barrier height increases nonlinearly.

Overall, we observe that the dependence fits an exponential function of the form $f(x) = ae^{-bx} + c$. In Fig. 3(c),

we compare the exponent b as a function of V_C, finding that it varies between 4.1 V^{-1} and 3.7 V^{-1} as V_C changes from 1.60 to 2.40 V. The average exponent is 3.9 V^{-1}. This nonlinear increase in barrier height approximately coincides with the regime for which the charge density is changing exponentially as the electron channel is forming under the gate.

The width of the barrier is plotted as a function of V_L for $V_C = 1.60$ V in Fig. 3(d). In this case, we show a two-parameter fit where the barrier height is not fixed. We see that the width appears to be constant over the range of V_L values considered. The average value is about 12 nm. We note that the extracted widths correspond well with our gate pattern.

In the context of compact models, the overall barrier height and width dependences on V_L and V_C have relatively simple forms that describe the behavior over very wide bias ranges.

In addition to profile-likelihood error bars, we also compute full 2D confidence regions for the parameters U_0 and w. Similar to the 1D case, we assume that the difference between the MLE point and the likelihood function is distributed as $\chi^2(2)$. Figure 3(e) shows the result for $V_L = 0.70$ V, where the contours represent different confidence levels.

Besides performing the above parameter extractions, we consider cases where the DOS in the leads is 0D, 1D, or 2D to examine if there are any distinguishing signatures in the calculated dependences. The difference in DOS affects the total number of available modes $M(E)$ in Eq. (3). All three cases fit the data well when the barrier heights and widths are allowed to adjust to compensate for the change in the DOS. We note that the estimated widths of the leads would be consistent with transversal quantization and that the energy-level splitting is too large to warrant 1D DOS for the transmission modes (i.e., multiple contributing 1D sub-bands). We therefore use the 0D DOS (i.e., a single 1D mode or sub-band) for the extracted barrier heights and widths. More generally, it is unclear at this time what factors in the tunnel-barrier geometry express clearer signatures due to the DOS dimensionality. This is a topic for future examination.

VI. ELECTROSTATIC SIMULATIONS

Lastly, we compare the results of the 1D barrier model with electrostatic simulations of the electron density and electric potential (i.e., the conduction-band edge, E_C). The electrostatics of the device are modeled using the COMSOL Multiphysics software. We use the Thomas-Fermi approximation to model the electric potential and the electron density at the oxide-semiconductor interface with a 2D density of states for the experimental gate voltages. The electron density is calibrated at a gate voltage of 0.48 V by applying an offset voltage to all electrodes while ramping gate L. The offset voltage is used to ensure that the electron density in the leads agrees with the definition of threshold given by Eq. (2).

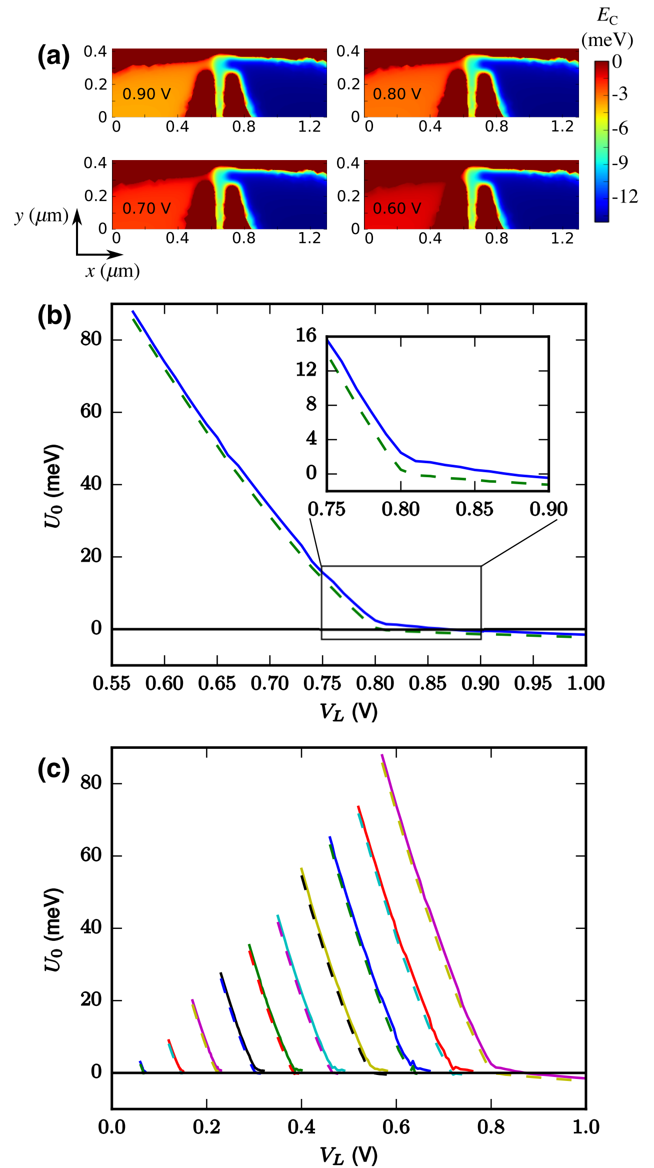

FIG. 4. A 3D electrostatics simulation of the tunnel barrier. (a) Conduction-band energy plots for different values of V_L. The other gates are $V_C = 1.60$ V and $V_R = 2.13$ V. The Fermi level is defined at zero in these plots. (b) The simulated barrier height as a function of V_L for $V_C = 1.60$ V, $V_R = 2.13$ V, and an offset voltage of 0.50 V. The solid line shows barrier height values where the effect of quantum confinement has been accounted for, while the dashed line does not take this effect into account. Inset: an enlargement of the boxed region to clarify the effect of lateral confinement on the barrier height. (c) The barrier height as a function of V_L for a range of offset voltages. The offset voltage varies from 0.05 V (leftmost curve) to 0.50 V (rightmost curve) in steps of 0.05 V. The solid and dashed lines have the same meaning as in (b).

Plots of the conduction-band energy as a function of V_L and position are shown in Fig. 4(a). As V_L is decreased, a single barrier forms between gates L and C. The barrier height is determined by calculating the saddle point in the potential, which is found by calculating the gradient of the potential. We note that the shape of the barrier will generally not be trapezoidal for a 1D cut of the conduction-band energy along a curve through the saddle point. The barrier height as a function of V_L for $V_C = 1.60$ V, $V_R = 2.13$ V, and an offset voltage of 0.50 V is shown in Figs. 4(b) and 4(c).

In the constriction where the barrier forms, we expect quantum confinement to have an effect. These transverse confinement energies are computed for this device geometry and become appreciable at the lower densities where the channel is narrow (see Appendix B). The effect of lateral confinement is shown more clearly in the inset of Fig. 4(b). The confinement increases the barrier and rounds the dependence on voltage at the intersection of two relatively linearly dependent regions. The barrier height is relatively linearly dependent on voltage well before and well after the electron channel is completely formed. The barrier goes to approximately zero once the tunnel-barrier gap is completely saturated with electron density. A qualitatively similar rounded region between the intersection of two linearly voltage-dependent regions is seen in the extracted barrier heights in Fig. 3(b), although extended over a wider voltage range. Better quantitative agreement is beyond the scope or goal of this paper (i.e., a simple compact model for device characterization). Improved agreement would certainly require a more accurate 3D simulation of the barrier, including self-consistent Schrödinger-Poisson calculations and 2D tunneling models (e.g., nonequilibrium Green's function approaches) [18–20].

We examine different offset voltages to further investigate the magnitude of the quantum-confinement effect on increasing the barrier and softening the intersection of the two linear regions (i.e., full depletion and channel-formation regions). Figure 4(c) shows the barrier height as a function of V_L for a variety of offset voltages (i.e., threshold offsets for the COMSOL calculation). Each curve shows the range from no accumulation (large barrier) to a nonexistent barrier (the barrier height is zero or lower). We do see some increase in the rounding of the regions of intersection for the highest V_L. The larger rounding observed in experiment might, therefore, be indicative of stronger quantum confinement or disorder effects in the lead than are being predicted by the simple confinement model used in this work.

VII. CONCLUSION

In conclusion, we measure and model the transport spectroscopy of a silicon MOS surface electrode-defined tunnel barrier at 4 K. Multiple electrodes are used to form the electrostatically defined barriers, making for a large parameter space on which the tunnel barrier is dependent.

Using a voltage-dependent barrier height and width, we examine a 1D barrier model that is found to fit the current-voltage dependence well over a large parameter range. The model is quasianalytic, providing both accurate 1D transmission values for the barrier model while also being a fast extraction method to more readily enable its application to multielectrode cases. This includes providing confidence that this formulation can be used reliably to efficiently simulate multielectrode tunnel barriers without the need of supporting computationally expensive microscopic simulation of the tunnel barrier. Cryogenic and quantum-confinement effects are included in the model to account for threshold and barrier height shifts.

The barrier height dependence on the gate voltage is shown to be relatively linear in the high- and low-voltage regimes, with an intermediate nonlinear regime that is nearly exponential. The exponential dependence correlates with the depletion or enhancement of the electron channel and simulations suggest that the barrier height can be sensitive to 1D quantum-confinement effects.

To clarify how quantitatively accurate the 1D model might be, we also compare the results of the 1D model to semiclassical solutions of Poisson's equation using COMSOL and find that the simulations qualitatively agree well with the three regimes predicted by the model.

Transport spectroscopy is a fast way to characterize tunnel barriers relative to more time-intensive and complex pulsing approaches used for these kinds of barriers [21–23]. The combination of this method with transport spectroscopy offers a relatively rapid way to build a compact model of a tunnel barrier for multiple-gate electrodes and reasonably large bias ranges. This work should provide useful insight into the details of electrostatic barriers and how to characterize them.

ACKNOWLEDGMENTS

We gratefully recognize conversations with D.R. Ward and P.A. Sharma about this work and the manuscript. The work was supported by the Sandia National Laboratories Directed Research and Development Program. This work was performed, in part, at the Center for Integrated Nanotechnologies, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under Contract No. DE-NA-0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be

expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

APPENDIX A: CALCULATION OF THE CRYOGENIC THRESHOLD VOLTAGE

We calculate the threshold voltage for a MOSFET at low temperatures by considering (1) the metal-insulator transition and (2) the consequences of low temperature on the parameters used to calculate the threshold, such as the metal-semiconductor work function, ϕ_{ms}. First, we consider the metal-insulator transition.

One definition of the threshold voltage for MOSFETs at high temperatures (such as room temperature) is that it is the gate voltage for which the mobile electron charge $Q_n = 0$. As noted in the text, at cryogenic temperatures there is a metal-insulator transition where we must achieve a critical electron density n_{crit} before appreciable conduction begins. Thus at low temperatures the threshold voltage is defined as the gate voltage for which $Q_n = -en_{\text{crit}}$. The expression for the threshold voltage is then given by

$$V_{\text{th}} = V_0 + V_{\text{crit}}. \tag{A1}$$

In this expression, V_0 is the standard MOSFET threshold voltage (discussed below) and $V_{\text{crit}} = en_{\text{crit}}/C_{\text{ox}}$, where C_{ox} is the oxide capacitance per unit area. The critical density can be extracted from Hall measurements [12].

We next examine V_0. The usual formula for the threshold voltage, neglecting the body effect (i.e., a zero voltage difference between the source and bulk) and assuming a p-type substrate, is given by the following equation [24]:

$$V_0 = V_{\text{FB}} + 2|\phi_p| + \frac{1}{C_{\text{ox}}}\sqrt{4\epsilon_s e N_a|\phi_p|}, \tag{A2}$$

where V_{FB} is the flat-band voltage, ϕ_p is the bulk potential (difference between the intrinsic and quasi-Fermi levels in the bulk of the semiconductor), and the third term is the voltage across the oxide due to the depletion layer charge. The flat-band voltage is given by $V_{\text{FB}} = \phi_{ms} - Q_f/C_{\text{ox}}$. The electrical permittivity of the semiconductor is ϵ_s and the acceptor concentration is N_a.

The parameters ϕ_{ms} and ϕ_p are most affected by changes in temperature. At cryogenic temperatures, the quasi-Fermi level in the semiconductor is pinned halfway between the acceptor energy levels and the top of the valence band due to freeze-out, as shown in Fig. 5. For boron, the acceptor energy levels are about 0.045 eV above E_V. It is well known that the band-gap energy E_g changes with temperature; for Si, the band-gap energy is 1.17 eV at low temperatures. Thus we find $\phi_p = E_g/2 - 0.045/2 = 0.563$ V.

The parameter ϕ_p also affects ϕ_{ms} through the semiconductor work function ϕ_s. Unlike E_g, the effect of low temperatures on the metal work function ϕ_m and the electron affinity χ is not clear. We use $\phi_m = 4.05$ V for n^+ polysilicon and assume $\chi = 4.05$ V for Si. This gives $\phi_s = \chi + E_g/2 + \phi_p = 5.20$ V and a metal-semiconductor work function of $\phi_{ms} = \phi_m - \phi_s = -1.15$ V.

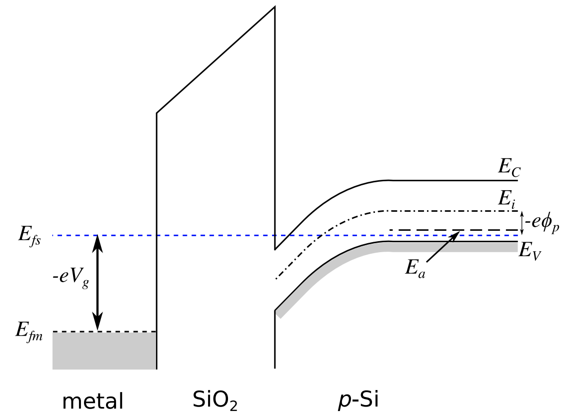

FIG. 5. The low-temperature energy-band diagram for a MOS system with a p-type Si substrate biased into inversion. The quasi-Fermi level in the semiconductor E_{fs} is pinned between the acceptor level E_a and the top of the valence band E_V due to freeze-out.

It is important to establish the defect densities below which "clean" MOS tunnel barriers begin to be observed and we can use the modified threshold to estimate fixed charge for this device. We use the observed tunnel-barrier threshold voltage and critical density at cryogenic temperatures to extract an effective fixed charge density, which we find to be $Q_f = 3.9 \times 10^{10}$ cm^{-2}, following the adjusted MOS threshold calculation above. Notably, the fixed charge would be below zero if the critical density were not considered.

Fixed charge densities are associated with the MOS interface and the density is often estimated from room-temperature MOSFET behavior. In this work, we are not able to establish a good measure of the room-temperature threshold because the devices are fabricated on ultra-high-resistivity substrates. The poor junction isolation leads to substantial conduction through the substrate from the contacts and regions where the channel forms. Ultra-high-resistivity substrates are used to minimize background electron spins, which is important for the intended longer-term quantum-information application. An additional reason to characterize the effective defect charge at low temperature instead of relying on a room-temperature fixed charge characterization is because charge-trap densities might be higher at lower temperature. That is, small fluctuations in potential might not confine charge at room temperature but would trap charge at lower temperature.

As an additional reference, we also note other measures of the disorder in the starting material. Hall measurements of the starting material also show a peak mobility

of 5950 $cm^2/(V \cdot s)$. An estimated surface roughness and correlation length of 2.2 Å and 24 Å, respectively, were also obtained from the high-density dependence [25].

APPENDIX B: LATERAL QUANTUM CONFINEMENT

In a simple square potential well or particle in a box, the discrete energy levels increase as the width of the quantum well is reduced. Similarly, quantum confinement in a nanowire causes the lowest sub-band to increase in energy. In this appendix, we consider the effect of lateral quantum confinement on the threshold voltage. This effect causes a shift in the threshold voltage, resulting in the threshold of a thin wire gate being larger than the threshold of a much wider gate (the field threshold). This theory was recently developed for nanowires in Si/SiGe heterostructures [13]; here, we apply it to the MOS geometry shown in Fig. 6(a).

To understand this effect, we consider a range of wire-gate voltages and wire widths. For each wire width, we compute the electrostatic potential for each corresponding wire-gate voltage by solving Poisson's equation in 2D with the finite-element method in COMSOL Multiphysics. We then calculate the ground-state energy of the resulting confinement potential by solving the Schrödinger equation. This gives us the energy of the lowest sub-band as a function of the voltage. We find that the ground-state energy is linear with gate voltage, that is,

$$E_0^L = m_L V_G, \tag{B1}$$

where E_0^L is the ground-state energy of a wire of width L, m_L is the slope or lever arm, and V_G is the wire-gate voltage. As the wire width increases, the slopes m_L saturate. We then define the lever arm for an infinitely wide wire as approximately equal to the lever arm for a 5000-nm-wide wire, that is, $m_\infty \approx m_{5000} = -0.91$ meV/mV for the MOS geometry shown in Fig. 6(a).

Assuming that conduction occurs when the ground-state energy is aligned with some external reference energy level, E^0 (e.g., the quasi-Fermi level), we define the field threshold voltage as follows:

$$V_\infty = \frac{E^0}{m_\infty}, \tag{B2}$$

where V_∞ is the threshold for an infinitely wide wire (the field threshold).

If the threshold voltage of a wire of width L is V_L and we define the threshold shift as $\Delta V_L = V_L - V_\infty$, then we may write

$$\begin{aligned} \Delta V_L &= V_L - V_\infty \\ &= \frac{E^0}{m_L} - \frac{E^0}{m_\infty} \\ &= \frac{E^0 m_\infty}{m_L m_\infty} - \frac{E^0 m_L}{m_\infty m_L} \\ &= \frac{E^0}{m_L m_\infty}(m_\infty - m_L). \end{aligned} \tag{B3}$$

We can rewrite the expression for the threshold shift explicitly in terms of the field threshold, using Eq. (B2), as follows:

$$\Delta V_L = V_\infty \left(\frac{m_\infty}{m_L} - 1 \right). \tag{B4}$$

Moreover, using the definition $\Delta V_L = V_L - V_\infty$, we can write a direct expression for the wire threshold as follows:

$$V_L = V_\infty \frac{m_\infty}{m_L}. \tag{B5}$$

We find that the lever arm m_L as a function of the wire width L is fitted well by a decaying exponential function, that is, $m_L \approx m_\infty + ae^{-bL}$, where $a = 0.32$ meV/mV and

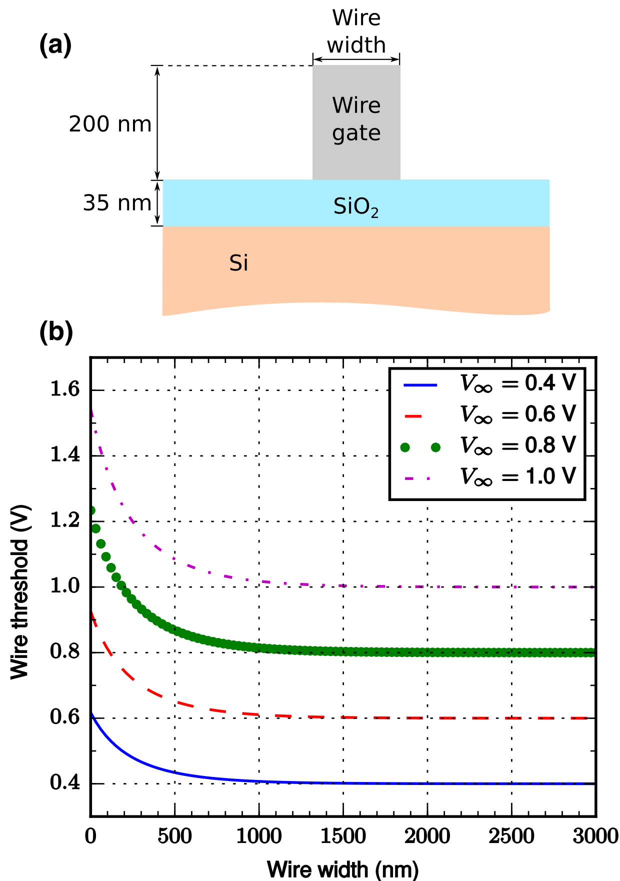

FIG. 6. (a) A cross-section schematic of the MOS geometry used to model the effect of quantum confinement on the threshold voltage. (b) The wire threshold voltage as a function of wire width for different field threshold voltages.

$b = 0.003\ \text{nm}^{-1}$. The expression for the wire threshold in Eq. (B5) then becomes

$$V_L = V_\infty \frac{m_\infty}{m_\infty + ae^{-bL}}. \tag{B6}$$

Thus we can calculate the wire threshold given a field threshold voltage and wire width using this compact model. The wire threshold as a function of the wire width for different field thresholds is shown in Fig. 6(b). We note this assumes that lateral quantum confinement is the dominant effect and that the exact increase due to confinement is geometry dependent.

APPENDIX C: THE TRANSPORT FORMALISM FOR A GENERAL TRAPEZOIDAL BARRIER

In this appendix, we present a detailed calculation of tunneling through a "1D" trapezoidal barrier. We will use the Landauer-Büttiker formalism, computing the transmission function via a semianalytic approach.

The Landauer-Büttiker formalism is a general framework for computing tunneling quantum transport through a device. It can handle devices with many leads and high source-drain biases. The method breaks down when transport becomes nonballistic, through "vertical" scattering events. This occurs when energy is not conserved throughout the scattering region; in this instance, more sophisticated techniques, such as nonequilibrium Green's functions, must be employed.

The case we consider here is a simple, two-terminal device. Within the Landauer-Büttiker formalism, the total current through the scattering region is given by

$$I = \int dE \frac{2e}{h} \bar{T}(E) \left[f_s(E) - f_d(E) \right], \tag{C1}$$

where $\bar{T}(E)$ is the total transmission function at a given energy (summed up over transverse modes), f_s is the Fermi function of the source, and f_d is the Fermi function of the drain. The difficult part of computing I is to calculate the total transmission function $\bar{T}(E)$. To facilitate this, we decompose $\bar{T}$ as follows:

$$\begin{aligned} \bar{T}(E) &= \sum_{m \in \text{modes}} T_m(E) \\ &= T(E) M(E), \end{aligned} \tag{C2}$$

where $T_m(E)$ is the transmission of a given transverse mode and $M(E)$ is the total number of accessible modes at a given energy E. Hence, we have decomposed the problem into computing $T(E)$ and $M(E)$ independently. First, we examine the computation of $T(E)$.

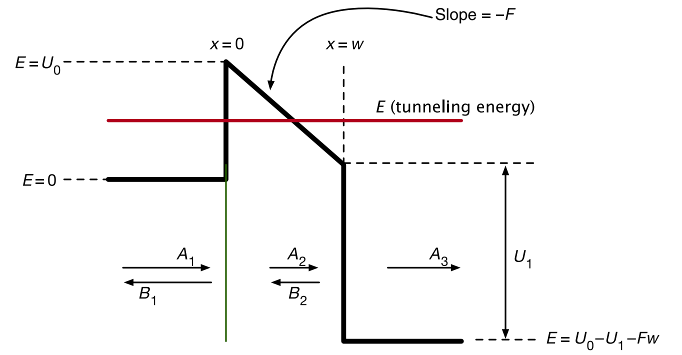

FIG. 7. A schematic of the transmission problem.

1. Computing the transmission function

The transmission function gives the probability of an incident mode tunneling through the barrier. Schematically, we consider the case shown in Fig. 7. There, the transmission function is a function of the energy E and is parametrized by the left-side barrier height U_0, the field F between the source and drain, the barrier width w, and the right-side barrier height U_1.

2. The form of the Schrödinger equation

To compute the transmission function, we need to solve the Schrödinger equation for an incoming plane wave at a specified energy. Within each region, we can write down the solution analytically:

$$\psi_1(x) = A_1 e^{ik_1 x} + B_1 e^{-ik_1 x}, \tag{C3}$$

$$\psi_2(x) = A_2 \text{Ai}(a(x)) + B_2 \text{Bi}(a(x)), \tag{C4}$$

$$\psi_3(x) = A_3 e^{ik_3 x}. \tag{C5}$$

Here, we have $k_1 = \sqrt{2mE}/\hbar$, $k_3 = \sqrt{2m(E - (U_0 - U_1 - Fw))}/\hbar$, and $a(x) = [2mF/\hbar^2]^{1/3} [(U_0 - E)/F - x]$. The transmission function is given by $T = |A_3/A_1|^2$. To solve for this quantity, we need to impose boundary conditions: $\psi_1(0) = \psi_2(0)$, $\psi_1'(0) = \psi_2'(0)$, $\psi_2(w) = \psi_3(w)$, and $\psi_2'(w) = \psi_3'(w)$. To satisfy the boundary conditions, we will also need to know the formulas for the derivatives:

$$\begin{aligned} \psi_1'(x) &= A_1 ik_1 e^{ik_1 x} - B_1 ik_1 e^{-ik_1 x}, \\ \psi_2'(x) &= -\left(\frac{2Fm}{\hbar^2}\right)^{1/3} \left[A_2 \text{Ai}'(a(x)) + B_2 \text{Bi}'(a(x))\right], \\ \psi_3'(x) &= A_3 ik_3 e^{ik_3 x}. \end{aligned} \tag{C6}$$

3. Renormalizing the Airy functions

One issue that arises immediately is that the Airy functions that form the solution in region 2 are not well defined when $F \to 0$. This poses a difficulty, since we want to

recover the original square barrier problem in this case. Since the argument of the Airy functions is real, we can use the asymptotic form for $F \to 0$:

$$\mathrm{Ai}(z) \sim \lim_{N\to\infty} \mathrm{Ai}_N(z) \equiv \frac{e^{-2/3z^{3/2}}}{2\sqrt{\pi}z^{1/4}} \sum_{n=0}^{N} \frac{(-1)^n \Gamma\left(n+\frac{5}{6}\right)\Gamma\left(n+\frac{1}{6}\right)\left(\frac{3}{4}\right)^n}{2\pi n! z^{3n/2}}, \tag{C7}$$

$$\mathrm{Bi}(z) \sim \lim_{N\to\infty} \mathrm{Bi}_N(z) \equiv \frac{e^{2/3z^{3/2}}}{\sqrt{\pi}z^{1/4}} \sum_{n=0}^{N} \frac{\Gamma\left(n+\frac{5}{6}\right)\Gamma\left(n+\frac{1}{6}\right)\left(\frac{3}{4}\right)^n}{2\pi n! z^{3n/2}}. \tag{C8}$$

These expressions are valid for real large-magnitude z. We want to use this to work out expressions for the boundary conditions of $\psi_2(x)$. We will use these to renormalize the solution coefficients so that we do not have undefined expressions. Numerically, we find that we need only go to $N = 0$ to obtain good accuracy.

We write down the asymptotic result at $x = 0$:

$$\mathrm{Ai}_{N=0}(a(0)) \equiv \alpha_0 = \frac{(F\hbar)^{1/6}\, e^{-2\sqrt{2m}(U_0-E)^{3/2}/3F\hbar}}{2\sqrt{\pi}\,(2m)^{1/12}\,(\mathrm{U}_0-E)^{1/4}}, \tag{C9}$$

$$\mathrm{Bi}_{N=0}(a(0)) \equiv \beta_0 = \frac{(F\hbar)^{1/6}\, e^{2\sqrt{2m}(U_0-E)^{3/2}/3F\hbar}}{\sqrt{\pi}\,(2m)^{1/12}\,(\mathrm{U}_0-E)^{1/4}}. \tag{C10}$$

Using this, we rewrite the ψ_2 solution as follows:

$$\psi_2(x) = (A_2\alpha_0)\frac{\mathrm{Ai}\,(a(x))}{\alpha_0} + (B_2\beta_0)\frac{\mathrm{Bi}\,(a(x))}{\beta_0} \equiv \mathcal{A}_2\mathcal{A}i\,(a(x)) + \mathcal{B}_2\mathcal{B}i\,(a(x)), \tag{C11}$$

where the new functions $\mathcal{A}i$ and $\mathcal{B}i$ have the property that

$$\lim_{F\to 0} \mathcal{A}i\,(a(x)) = e^{(x/\hbar)\sqrt{2m(U_0-E)}}, \tag{C12}$$

$$\lim_{F\to 0} \mathcal{B}i\,(a(x)) = e^{-(x/\hbar)\sqrt{2m(U_0-E)}}, \tag{C13}$$

which are, of course, just the growing and decaying exponentials that span the solution for a flat barrier.

Hence, this renormalization allows us to make contact with the simple limiting case. We thus replace our region 2 wave function and derivative with following piecewise definitions:

$$\psi_2(x) = \begin{cases} A_2\dfrac{\mathrm{Ai}\,(a(x))}{\alpha_0} + B_2\dfrac{\mathrm{Bi}\,(a(x))}{\beta_0}, & F \ge F_0, \\ A_2 e^{(x/\hbar)\sqrt{2m(U_0-E)}} + B_2 e^{-(x/\hbar)\sqrt{2m(U_0-E)}}, & F < F_0, \end{cases} \tag{C14}$$

$$\psi_2'(x) = \begin{cases} -\left(\dfrac{2Fm}{\hbar^2}\right)^{1/3}\left[A_2\dfrac{\mathrm{Ai}'\,(a(x))}{\alpha_0} + B_2\dfrac{\mathrm{Bi}'\,(a(x))}{\beta_0}\right], & F \ge F_0, \\ \dfrac{\sqrt{2m(U_0-E)}}{\hbar} \times \left[A_2 e^{(x/\hbar)\sqrt{2m(U_0-E)}} - B_2 e^{-(x/\hbar)\sqrt{2m(U_0-E)}}\right], & F < F_0. \end{cases} \tag{C15}$$

Here, we just relabel the unknown coefficients A_2 and B_2 for which we are solving to include the renormalization. We identify a good value of a threshold field F_0 numerically.

4. Construction of the system of equations

Now that we have worked out the wave functions for all three regions, we are ready to compute the transmission function. However, note that we have five unknown coefficients with only four constraints. This is because the full wave function needs to be normalized, which provides the fifth constraint. However, the transmission function only needs a ratio of coefficients, so we do not need to worry about this additional constraint. Specifically, we divide through everywhere by A_1, giving the following:

$$\psi_1(x) = e^{ik_1x} + \tilde{B}_1 e^{-ik_1x}, \tag{C16}$$

$$\psi_2(x) = \begin{cases} \tilde{A}_2\dfrac{\mathrm{Ai}\,(a(x))}{\alpha_0} + \tilde{B}_2\dfrac{\mathrm{Bi}\,(a(x))}{\beta_0}, & F \ge F_0, \\ \tilde{A}_2 e^{(x/\hbar)\sqrt{2m(U_0-E)}} + \tilde{B}_2 e^{-(x/\hbar)\sqrt{2m(U_0-E)}}, & F < F_0, \end{cases} \tag{C17}$$

$$\psi_3(x) = \tilde{A}_3 e^{ik_3x}, \tag{C18}$$

$$\psi_1'(x) = ik_1 e^{ik_1x} - \tilde{B}_1 ik_1 e^{-ik_1x}, \tag{C19}$$

$$\psi_2'(x) = \begin{cases} -\left(\dfrac{2Fm}{\hbar^2}\right)^{1/3}\left[\tilde{A}_2\dfrac{\mathrm{Ai}'\,(a(x))}{\alpha_0} + \tilde{B}_2\dfrac{\mathrm{Bi}'\,(a(x))}{\beta_0}\right], & F \ge F_0, \\ \dfrac{\sqrt{2m(U_0-E)}}{\hbar} \times \left[\tilde{A}_2 e^{(x/\hbar)\sqrt{2m(U_0-E)}} - \tilde{B}_2 e^{-(x/\hbar)\sqrt{2m(U_0-E)}}\right], & F < F_0, \end{cases} \tag{C20}$$

$$\psi_3'(x) = \tilde{A}_3 ik_3 e^{ik_3x}. \tag{C21}$$

Here, the variables with tildes indicate scaling by A_1, so $T = |\tilde{A}_3|^2 k_3/k_1$, where the last factor comes from the difference in velocity between incoming and outgoing modes. We will now drop the tildes for convenience.

Considering first the case where the Airy functions are well defined, we have the following system of equations to solve:

$$\begin{aligned}
1 &= -1B_1 + \frac{\mathrm{Ai}\,(a(0))}{\alpha_0}A_2 + \frac{\mathrm{Bi}\,(a(0))}{\beta_0}B_2 + 0A_3,\\
ik_1 &= ik_1B_1 - \left[\left(\frac{2Fm}{\hbar^2}\right)^{1/3}\frac{\mathrm{Ai}'\,(a(0))}{\alpha_0}\right]A_2\\
&\quad - \left[\left(\frac{2Fm}{\hbar^2}\right)^{1/3}\frac{\mathrm{Bi}'\,(a(0))}{\beta_0}\right]B_2 + 0A_3,\\
0 &= 0B_1 + \left[\frac{\mathrm{Ai}\,(a(w))}{\alpha_0}\right]A_2 + \left[\frac{\mathrm{Bi}\,(a(w))}{\beta_0}\right]B_2\\
&\quad - \left[e^{ik_3w}\right]A_3,\\
0 &= 0B_1 + \left[\left(\frac{2Fm}{\hbar^2}\right)^{1/3}\frac{\mathrm{Ai}'\,(a(w))}{\alpha_0}\right]A_2\\
&\quad + \left[\left(\frac{2Fm}{\hbar^2}\right)^{1/3}\frac{\mathrm{Bi}'\,(a(w))}{\beta_0}\right]B_2 + \left[ik_3e^{ik_3w}\right]A_3.
\end{aligned} \tag{C22}$$

If, instead, the Airy functions are not well defined and we have to use the asymptotic forms, we have the following:

$$\begin{aligned}
1 &= -1B_1 + 1A_2 + 1B_2 + 0A_3,\\
ik_1 &= ik_1B_1 + \left[\frac{\sqrt{2m(U_0-E)}}{\hbar}\right]A_2\\
&\quad - \left[\frac{\sqrt{2m(U_0-E)}}{\hbar}\right]B_2 + 0A_3,\\
0 &= 0B_1 + \left[e^{(w/\hbar)\sqrt{2m(U_0-E)}}\right]A_2\\
&\quad + \left[e^{-(w/\hbar)\sqrt{2m(U_0-E)}}\right]B_2 - \left[e^{ik_3w}\right]A_3,\\
0 &= 0B_1 - \left[\frac{\sqrt{2m(U_0-E)}}{\hbar}e^{(w/\hbar)\sqrt{2m(U_0-E)}}\right]A_2\\
&\quad + \left[\frac{\sqrt{2m(U_0-E)}}{\hbar}e^{-(w/\hbar)\sqrt{2m(U_0-E)}}\right]B_2\\
&\quad + \left[ik_3e^{ik_3w}\right]A_3.
\end{aligned} \tag{C23}$$

We cast these equations as a simple matrix problem solution for A_3.

5. Computing the total current

Now that we can reliably compute the transmission function, we need to use it to compute the total current

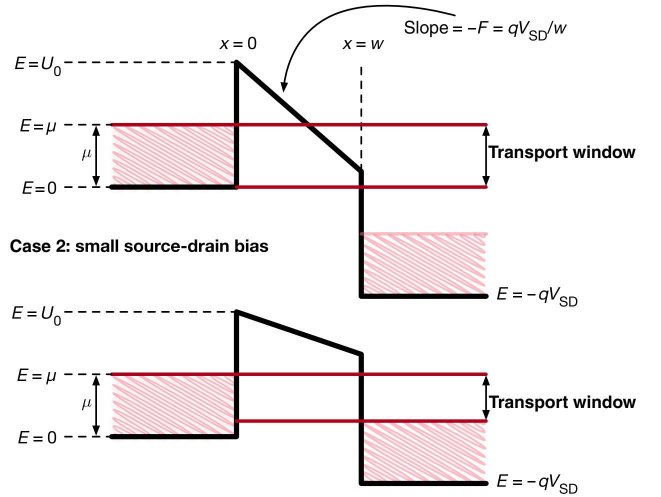

FIG. 8. A schematic for computing the total current.

within the Landauer-Büttiker formalism. To do this, we modify the schematic slightly to the form shown in Fig. 8.

6. The density of the transverse modes

The total transmission function $\bar{T}(E)$ is given by

$$\begin{aligned}
\bar{T}(E) &= \sum_{m\in\text{modes}} T_m(E)\\
&= T(E)M(E)\\
&= T(E)L^d\int_0^E dE' g_d(E'),
\end{aligned} \tag{C24}$$

where $T_m(E)$ is the transmission of a given transverse mode, $M(E)$ is the total number of accessible modes at a given energy E, g is the density of states, L is the cross-section length scale of the scattering region, and d is the dimension of the lead. The three cases of interest are for 0D (one mode), 1D (sheet contacts), and 2D (volume contacts) densities of states, given by the following:

$$\begin{aligned}
g_0(E) &= \delta(E_0 - E),\\
g_1(E) &= \frac{2}{\pi\hbar}\sqrt{\frac{2m}{E}},\\
g_2(E) &= \frac{2m}{\pi\hbar^2},
\end{aligned} \tag{C25}$$

where we have multiplied the usual formulas for the 1D and 2D DOS expressions by 2 for the extra valley degeneracy in silicon. Using these density-of-states formulas, we

compute $M(E)$:

$$M_0(E) = \begin{cases} 1, & E_0 \leq E, \\ 0, & E_0 > E, \end{cases}$$

$$M_1(E) = \frac{4L\sqrt{2mE}}{\pi\hbar}, \tag{C26}$$

$$M_2(E) = \frac{2mL^2E}{\pi\hbar^2}. \tag{C27}$$

7. Computation of the current

Equipped with the number of modes, we can compute the current:

$$I = \int dE \frac{2e}{h} \bar{T}(E) \left[f_s(E) - f_d(E)\right] \approx \frac{2e}{h} \int_{\mu_0}^{\mu} dE T(E) M(E), \tag{C28}$$

where

$$\mu_0 = \begin{cases} 0, & \mu - qV_{\text{SD}} < 0, \\ \mu - qV_{\text{SD}}, & \mu - qV_{\text{SD}} \geq 0. \end{cases}$$

The two cases here correspond to cases 1 and 2 in Fig. 8. From the above formulas, we can write three different current models:

$$I_0 = \frac{2e}{h} \int_{\mu_0}^{\mu} dE T(E), \tag{C29}$$

$$I_1 = \frac{2e}{h} \int_{\mu_0}^{\mu} dE \frac{4LT(E)\sqrt{2mE}}{\pi\hbar}, \tag{C30}$$

$$I_2 = \frac{2e}{h} \int_{\mu_0}^{\mu} dE \frac{2mL^2ET(E)}{\pi\hbar^2}, \tag{C31}$$

where we assume that for I_0 the single mode is in the transport window.

APPENDIX D: THE THREE-STEP MODEL

In the three-step barrier model, we approximate a square barrier under bias using three step potentials as shown in Fig. 9. For $V_{SD} = 0$, the barrier reduces to a square barrier of height U_0 and width w. The first step occurs at $x = -w/2$ and has a fixed height of U_0. The next step occurs at $x = 0$ and has a variable step height that depends on the source-drain bias V_{SD}. The final step occurs at $x = w/2$ and also has a variable height that depends on V_{SD}. We use the transfer or T-matrix formalism to compute the transmission coefficient [26,27]. For simplicity, we only consider positive V_{SD} but this approach can be easily applied to negative V_{SD} as well.

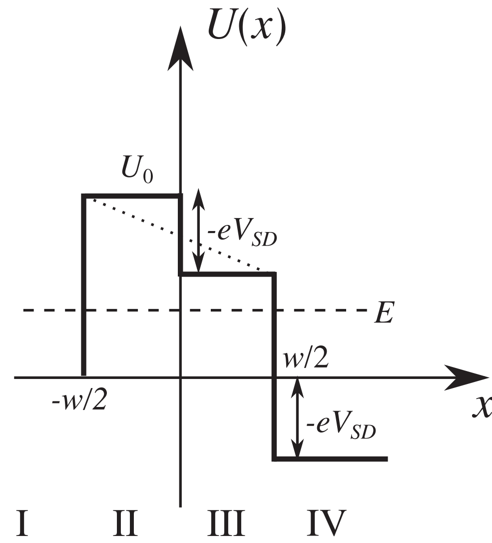

FIG. 9. The 1D potential energy $U(x)$ diagram, showing the three-step barrier model. The energy of the impinging electron is E.

First, consider the case where $E > U_0$. There are four wave numbers, each corresponding to a particular region of the barrier as indicated by the roman numerals in Fig. 9. The wave numbers are given as follows:

$$k_1^2 = 2mE/\hbar^2, \tag{D1}$$

$$k_2^2 = 2m\left(E - U_0\right)/\hbar^2, \tag{D2}$$

$$k_3^2 = 2m\left(E - U_0 + eV_{SD}\right)/\hbar^2, \tag{D3}$$

$$k_4^2 = 2m\left(E + eV_{SD}\right)/\hbar^2. \tag{D4}$$

Here, m is the effective mass of the electron. Using the T-matrix formalism, we find the transmission coefficient to be as follows:

$$T = \frac{4k_4/k_1}{\left[f_1 \cos\theta \cos\phi - f_2 \sin\theta \sin\phi\right]^2 + \left[f_3 \sin\theta \cos\phi + f_4 \cos\theta \sin\phi\right]^2}, \tag{D5}$$

where the coefficients f_1, f_2, f_3, and f_4 are functions of the wave numbers, given by

$$f_1 = 1 + \frac{k_4}{k_1}, \tag{D6}$$

$$f_2 = \frac{k_3}{k_2} + \frac{k_2 k_4}{k_1 k_3}, \tag{D7}$$

$$f_3 = \frac{k_4}{k_2} + \frac{k_2}{k_1}, \tag{D8}$$

$$f_4 = \frac{k_4}{k_3} + \frac{k_3}{k_1}. \quad \text{(D9)}$$

The arguments of the trigonometric functions depend on the width of the barrier as $\theta = k_2 w/2$ and $\phi = k_3 w/2$. This is the general form of the transmission coefficient.

Now consider the large-V_{SD} case such that $V_{SD} > U_0 - E$ and $U_0 > E$. The form of the transmission coefficient and the wave numbers remain the same, except that the wave number in region II is replaced in the following manner: $k_2 \rightarrow i\kappa_2$. Then, the trigonometric functions that have θ as the argument become hyperbolic functions, that is, $\sin k_2 w/2 \rightarrow i \sinh \kappa_2 w/2$ and $\cos k_2 w/2 \rightarrow \cosh \kappa_2 w/2$. The wave number in region II is now given by $\kappa_2^2 = 2m(U_0 - E)/\hbar^2$.

Then consider the small V_{SD} case such that $V_{SD} < U_0 - E$ and $U_0 > E$, as shown in Fig. 9. Again, the form of the transmission coefficient remains the same, except that now the wave numbers in both region II and region III are replaced: $k_2 \rightarrow i\kappa_2$ and $k_3 \rightarrow i\kappa_3$. The wave numbers in regions II and III are now given by $\kappa_2^2 = 2m(U_0 - E)/\hbar^2$ and $\kappa_3^2 = 2m(U_0 - E - eV_{SD})/\hbar^2$. Now all the trigonometric functions become hyperbolic functions, as discussed above.

Finally, Eq. (D5) and the appropriate wave numbers (based on the values of V_{SD}, U_0, and E) are used in Eq. (C1) to compute the current through the device.

[1] Adrian M. Ionescu and Heike Riel, Tunnel field-effect transistors as energy-efficient electronic switches, Nature **479**, 329 (2011).

[2] D. M. Zajac, T. M. Hazard, X. Mi, K. Wang, and J. R. Petta, A reconfigurable gate architecture for Si/SiGe quantum dots, Appl. Phys. Lett. **106**, 223507 (2015).

[3] A. O. Orlov, I. Amlani, G. H. Bernstein, C. S. Lent, and G. L. Snider, Realization of a functional cell for quantum-dot cellular automata, Science **277**, 928 (1997).

[4] K. Uchida, K. Junji, O. Ryuji, and A. Toriumi, Programmable single-electron transistor logic for future low-power intelligent LSI: Proposal and room-temperature operation, IEEE Trans. Electron Devices **50**, 1623 (2003).

[5] K. MacLean, S. Amasha, Iuliana P. Radu, D. M. Zumbühl, M. A. Kastner, M. P. Hanson, and A. C. Gossard, Energy-Dependent Tunneling in a Quantum Dot, Phys. Rev. Lett. **98**, 036802 (2007).

[6] M. Rudolph, P. Harvey-Collard, R. Jock, T. Jacobson, J. Wendt, T. Pluym, J. Domńguez, G. Ten-Eyck, R. Manginell, M. P. Lilly, and M. S. Carroll, in *2016 IEEE International Electron Devices Meeting(IEDM)* (IEEE, San Francisco, 2016), p. 34.1.1.

[7] M. Singh, J. L. Pacheco, D. Perry, E. Garratt, G. Ten Eyck, N. C. Bishop, J. R. Wendt, R. P. Manginell, J. Dominguez, T. Pluym, D. R. Luhman, E. Bielejec, M. P. Lilly, and M. S. Carroll, Electrostatically defined silicon quantum dots with counted antimony donor implants, Appl. Phys. Lett. **108**, 062101 (2016).

[8] E. P. Nordberg, G. A. Ten Eyck, H. L. Stalford, R. P. Muller, R. W. Young, K. Eng, L. A. Tracy, K. D. Childs, J. R. Wendt, R. K. Grubbs, J. Stevens, M. P. Lilly, M. A. Eriksson, and M. S. Carroll, Enhancement-mode double-top-gated metal-oxide-semiconductor nanostructures with tunable lateral geometry, Phys. Rev. B **80**, 115331 (2009).

[9] G. P. Lansbergen, R. Rahman, C. J. Wellard, I. Woo, J. Caro, N. Collaert, S. Biesemans, G. Klimeck, L. C. L. Hollenberg, and S. Rogge, Gate-induced quantum-confinement transition of a single dopant atom in a silicon FinFET, Nat Phys **4**, 656 (2008).

[10] Ted Thorbeck and Neil M. Zimmerman, Formation of strain-induced quantum dots in gated semiconductor nanostructures, AIP Adv. **5**, 087107 (2015).

[11] C. W. J. Beenakker and H. vanHouten, in *Semiconductor Heterostructures and Nanostructures*, edited by Henry Ehrenreich and David Turnbull, Solid State Physics, Vol. 44 (Academic Press, Boston, 1991), p. 1.

[12] L. A. Tracy, E. H. Hwang, K. Eng, G. A. Ten Eyck, E. P. Nordberg, K. Childs, M. S. Carroll, M. P. Lilly, and S. Das Sarma, Observation of percolation-induced two-dimensional metal-insulator transition in a Si MOSFET, Phys. Rev. B **79**, 235307 (2009).

[13] T. M. Lu, J. K. Gamble, R. P. Muller, E. Nielsen, D. Bethke, G. A. Ten Eyck, T. Pluym, J. R. Wendt, J. Dominguez, M. P. Lilly, M. S. Carroll, and M. C. Wanke, Fabrication of quantum dots in undoped $Si/Si_{0.8}Ge_{0.2}$ heterostructures using a single metal-gate layer, Appl. Phys. Lett. **109**, 093102 (2016).

[14] S. Datta, *Electronic Transport in Mesoscopic Systems*, Cambridge Studies in Semiconductor Physics (Cambridge University Press, Cambridge, 1997).

[15] John G. Simmons, Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film, J. Appl. Phys. **34**, 1793 (1963).

[16] M. Lenzlinger and E. H. Snow, Fowler-Nordheim tunneling into thermally grown SiO_2, J. Appl. Phys. **40**, 278 (1969).

[17] Everardus H. Huisman, Constant M. Guédon, Bart J. van Wees, and Sense Jan van der Molen, Interpretation of transition voltage spectroscopy, Nano Lett. **9**, 3909 (2009).

[18] H. Ilatikhameneh, R. B. Salazar, G. Klimeck, R. Rahman, and J. Appenzeller, From Fowler-Nordheim to non-equilibrium Green's function modeling of tunneling, arXiv:1509.08170 [cond-mat.mes-hall].

[19] X. Gao, D. Mamaluy, E. Nielsen, R. W. Young, A. Shirkhorshidian, M. P. Lilly, N. C. Bishop, M. S. Carroll, and R. P. Muller, Efficient self-consistent quantum transport simulator for quantum devices, J. Appl. Phys. **115**, 133707 (2014).

[20] M. Friesen, M. Y. Simmons, and M. A. Eriksson, Tunneling in nanoscale devices, arXiv:1310.6115 [cond-mat.mes-hall].

[21] J. M. Elzerman, R. Hanson, L. H. Willems van Beveren, L. M. K. Vandersypen, and L. P. Kouwenhoven, Excited-state spectroscopy on a nearly closed quantum dot via charge detection, Appl. Phys. Lett. **84**, 4617 (2004).

[22] S. Amasha, K. MacLean, I. P. Radu, D. M. Zumbühl, M. A. Kastner, M. P. Hanson, and A. C. Gossard, Spin-dependent

tunneling of single electrons into an empty quantum dot, Phys. Rev. B **78**, 041306 (2008).

[23] Madhu Thalakulam, C. B. Simmons, B. J. Van Bael, B. M. Rosemeyer, D. E. Savage, M. G. Lagally, Mark Friesen, S. N. Coppersmith, and M. A. Eriksson, Single-shot measurement and tunnel-rate spectroscopy of a Si/SiGe few-electron quantum dot, Phys. Rev. B **84**, 045307 (2011).

[24] R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits* (Wiley, New York, 2002).

[25] G. Mazzoni, A. L. Lacaita, L. M. Perron, and A. Pirovano, On surface roughness-limited mobility in highly doped n-MOSFET's, IEEE Trans. Electron Devices **46**, 1423 (1999).

[26] John H. Davies, *The Physics of Low-Dimensional Semiconductors* (Cambridge University Press, New York, 1997). Cambridge Books Online.

[27] Yuji Ando and Tomohiro Itoh, Calculation of transmission tunneling current across arbitrary potential barriers, J. Appl. Phys. **61**, 1497 (1987).