General atomistic approach for modeling metal-semiconductor interfaces using density functional theory and nonequilibrium Green's function

Stradi, Daniele; Martinez, Umberto; Blom, Anders; Brandbyge, Mads; Stokbro, Kurt

Published in:
Physical Review B

Link to article, DOI:
10.1103/PhysRevB.93.155302

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
General atomistic approach for modeling metal-semiconductor interfaces using density functional theory and nonequilibrium Green's function

Daniele Stradi,1,2,* Umberto Martinez,2 Anders Blom,2 Mads Brandbyge,1 and Kurt Stokbro2
1Center for Nanostructured Graphene (CNG), Department of Micro- and Nanotechnology (DTU Nanotech), Ørsteds Plads, Building 345B, DK-2800 Kongens Lyngby, Denmark
2QuantumWise A/S, Fruebjergvej 3, Postbox 4, DK-2100 Copenhagen, Denmark

Metal-semiconductor contacts are a pillar of modern semiconductor technology. Historically, their microscopic understanding has been hampered by the inability of traditional analytical and numerical methods to fully capture the complex physics governing their operating principles. Here we introduce an atomistic approach based on density functional theory and nonequilibrium Green’s function, which includes all the relevant ingredients required to model realistic metal-semiconductor interfaces and allows for a direct comparison between theory and experiments via I-Vbias curve simulations. We apply this method to characterize an Ag/Si interface relevant for photovoltaic applications and study the rectifying-to-Ohmic transition as a function of the semiconductor doping. We also demonstrate that the standard “activation energy” method for the analysis of I-Vbias data might be inaccurate for nonideal interfaces as it neglects electron tunneling, and that finite-size atomistic models have problems in describing these interfaces in the presence of doping due to a poor representation of space-charge effects. Conversely, the present method deals effectively with both issues, thus representing a valid alternative to conventional procedures for the accurate characterization of metal-semiconductor interfaces.

DOI: 10.1103/PhysRevB.93.155302

I. INTRODUCTION

Metal-semiconductor (M-S) contacts play a pivotal role in almost any semiconductor-based technology. They are an integral part of a broad range of devices with applications as diverse as photovoltaics (PVs) [1], transistors and diodes [2,3], and fuel cells [4,5].

The requirement of M-S interfaces with tailored characteristics, such as a specific resistance at the contact, has fueled research on the topic for decades [6,7]. Nevertheless, despite the high degree of sophistication of current semiconductor technology, the understanding of M-S interfaces at the microscopic level still constitutes a considerable challenge [8,9]. Even the structure of the interface itself, which is buried in the macroscopic bulk metal and semiconductor materials, represents a serious impediment, as it makes the direct exploration of the interface properties cumbersome.

A measure of the device current I as a function of the applied bias \( V_{\text{bias}} \) is a standard procedure to probe a M-S interface [2], despite the drawback that the measured I-Vbias curves do not provide any direct information on the interface itself but rather on the full device characteristics. As such, it is common practice to interpret the I-Vbias curves by fitting the data with analytical models, which are then used to extract the relevant interface parameters such as the Schottky barrier height \( \Phi \) [10]. As no general analytical model exists, the accuracy of this procedure critically depends on whether the model describes well the physical regime of the interface under scrutiny. Furthermore, most models disregard the atomistic aspect of the interface, although it is nowadays accepted that chemistry plays a dominant role in determining the electronic characteristics of the interface [7,9,11–13]. These ambiguities complicate the assignment of the features observed in the measured spectra to specific characteristics of the M-S interface.

Conversely, atomistic electronic structure methods [14] are an ideal tool for the characterization of M-S interfaces and have been successfully employed over the years for their analysis [15–22]. However, due to their computational cost, these studies have focused on model interfaces described using finite-size models formed by few atomic layers (e.g., slabs), the validity of which is justified in terms of the local nature of the electronic perturbation due to the interface. For similar reasons, most studies have considered nondoped interfaces, as the models required to describe a statistically meaningful distribution of dopants in the semiconductor would be excessively demanding [23,24]. Last but not least, these model calculations describe the system only at equilibrium (i.e., at \( V_{\text{bias}} = 0 \) V), thereby missing a direct connection with the I-Vbias measurements.

Here, we develop a general framework which attempts to overcome the limitations inherent in conventional electronic structure methods for simulating M-S interfaces. We employ density functional theory (DFT) [25] together with the nonequilibrium Green’s function (NEGF) method [26] to describe the infinite, nonperiodic interface exactly. The DFT+NEGF scheme allows us to predict the behavior of the M-S interface under working conditions by simulating the I-Vbias characteristics of the interface at zero and at finite \( V_{\text{bias}} \). To describe correctly the electronic structure of the doped semiconductor, we employ an exchange-correlation (xc) functional designed ad hoc to reproduce the experimental semiconductor band gap [27], and a spatially dependent effective scheme to account for the doping on the semiconductor side.

We apply this DFT+NEGF approach to study the characteristics of an Ag/Si interface relevant for PV applications [12,23,28–38]. Specifically, we focus on the (100)/(100) interface [28,37] and on the dependence of I-Vbias...
characteristics on the semiconductor doping—notice that the method is completely general and can be used to describe other M-S interfaces with different crystalline orientations. We consider a range of doping densities for which the interface changes from rectifying to Ohmic. We demonstrate that the “activation energy” (AE) method routinely employed to analyze M-S contacts systematically overestimates the value of $\Phi$, with an error that is both bias and doping dependent, due to the assumption of a purely thermionic transport mechanism across the barrier. Conversely, we show how an analysis of the $I-V$ characteristics based on the DFT+NEGF electronic structure data provides a coherent picture of the rectifying-to-Ohmic transition as the doping is varied. Finally, we also show that finite-size slab models do not provide a good representation of the interface electronic structure when doping in the semiconductor is taken into account. This is due to the inability of the semiconductor side of the slab to screen the electric field resulting from the formation of the interface.

The paper is organized as follows. Sections II and III describe the computational methods and the system models employed in this work, respectively. Section IV A presents the calculated $I-V$ characteristics and the validation of the AE method based on the calculated data. Section IV B deals with the analysis of the $I-V$ curves in terms of the electronic structure of the interface as obtained from the DFT+NEGF calculations. In Sec. IV C, the simulations are compared to finite-size slab models. The main conclusions are drawn in Sec. V.

II. COMPUTATIONAL METHODS

The Ag(100)/Si(100) interface has been simulated using Kohn-Sham (KS) DFT as implemented in the ATOMISTIX TOOLKIT (ATK). DFT [25] and DFT+NEGF [26] simulations have been performed using a formalism based on a nonorthogonal pseudoatomic orbitals [40] (PAOs) basis set. The one-electron KS valence orbitals have been expanded using a linear combination of double-$\zeta$ PAOs including polarization functions (DZP). The confinement radii $r_c$ employed are 4.39 Bohr, 7.16 Bohr, and 7.16 Bohr for the Ag 4$d$, 5$s$, and 5$p$ orbitals, respectively, and 5.40 Bohr, 6.83 Bohr, and 6.83 Bohr for the Si 3$s$, 3$p$, and 3$d$ orbitals, respectively. The ionic cores have been described using Troullier-Martins [41] norm-conserving pseudopotentials [42]. The energy cutoff for the real-space grid used to evaluate the Hartree and xc contributions of the KS Hamiltonian has been set to 150 Ry. Monkhorst-Pack [43] grids of $k$ points have been used to sample the 3D (2D) Brillouin zone in the DFT (DFT+NEGF) simulations. We have used an $11 \times 11 \times 11$ grid of $k$ points for the bulk calculations, and a $k$ points grid of $18 \times 9 \times 1$ ($18 \times 9$) for the DFT (DFT+NEGF) simulations of the interface. Geometry optimizations have been performed by setting the convergence threshold for the forces of the moving atoms to $2 \times 10^{-2}$ eV/Å. In all the simulations, periodic boundary conditions (PBCs) have been used to describe the periodic structure extending along the directions parallel to the interface plane. In the slab DFT simulation, Dirichlet and Neumann boundary conditions have been applied in the direction normal to the interface on the silver and silicon sides of the simulation cell, respectively, whereas in the DFT+NEGF simulations the same direction has been described using Dirichlet boundary conditions at the two boundaries between the interface and the bulklike electrodes.

A. “Spill-in” terms

As described in Ref. [26], the DFT+NEGF method used to simulate the infinite, nonperiodic Ag(100)/Si(100) interface relies on a two-probe setup in which a left (L) and a right (R) semi-infinite electron reservoirs are connected through a central (C) region containing the interface. Once the chemical potentials $\mu_{L,R}$ of the reservoirs have been defined, a self-consistent (SCF) KS procedure is used to obtain the electronic density of the C region. The main quantity being evaluated in the SCF cycle is the density matrix required to express the electronic density of the C region in the basis of PAOs centered in the same region, $\tilde{D}^{CC}$. Assuming $\mu_L > \mu_R$, $\tilde{D}^{CC}$ takes the form

$$\tilde{D}^{CC} = -\frac{1}{\pi} \int_{-\infty}^{\mu_R} \text{Im} [\tilde{G}^{CC}] d\epsilon$$

$$-\frac{1}{\pi} \int_{\mu_L}^{\mu_R} \tilde{G}^{CC} \text{Im} [\tilde{\Sigma}^{LL}] \tilde{G}^{CC} d\epsilon,$$

where $\tilde{\Sigma}^{LL}$ is the self-energy matrix describing the coupling of the central region to the semi-infinite L reservoir, and the Green’s function of the central region $\tilde{G}^{CC}$ is obtained by

$$\tilde{G}^{CC}(\epsilon) = ((\epsilon + i\delta)\tilde{S}^{CC} - \tilde{H}^{CC} - \tilde{\Sigma}^{LL} - \tilde{\Sigma}^{RR})^{-1},$$

with $\tilde{S}^{CC}$ and $\tilde{H}^{CC}$ being the overlap and Hamiltonian matrices associated with the PAOs centered at the C region, and $\tilde{\Sigma}^{RR}$ being the self-energy matrix of the R reservoir.

However, even if the DFT+NEGF method provides an elegant scheme to evaluate $\tilde{D}^{CC}$, it should be noticed that solving Eqs. (1) and (2) is not sufficient to obtain the correct Hamiltonian and the electronic density of the C region. The reason is that the relevant integrals involved in Eqs. (1) and (2) are evaluated only in the region of space encompassing the C region, and only for the atoms localized in that region. As a consequence, the tails of the PAOs located close to both sides of the L/C and R/C boundaries, which penetrate into the neighboring regions, are not accounted for (see Fig. 1). To correct this behavior, we introduce additional corrective terms that we name “spill-in”. For the Hamiltonian, corrective terms are applied to both the two-center and three-center integrals. Specifically, if two PAOs $\phi_i$ and $\phi_j$ centered in the C region lie close to a boundary, e.g., to the L/C one, the corrected Hamiltonian in Eq. (2) will also include the matrix element $H_{i,j} = \langle \phi_i | V_{\text{eff}}^{LL}(r) | \phi_j \rangle$ associated with the tails of the PAOs protruding into the L region, $V_{\text{eff}}^{LL}(r)$ being the periodic KS potential of the semi-infinite L reservoir [Fig. 1(a)]. Similar arguments hold also for the Hamiltonian three-center nonlocal terms. For the electronic density of the C region, additional contributions are included for each pair of PAOs $\phi_i$ and $\phi_j$ located close to a boundary when at least one of them is centered at the neighboring reservoir region. In total, two new contributions must be added to the electronic density evaluated using Eqs. (1) and (2) for each pair of PAOs at each boundary.
FIG. 1. Scheme showing the two-center “spill-in” terms used for the evaluation of the Hamiltonian (a) and the electronic density (b) of the C region for a pair of s-type PAOs $\phi_j$ and $\phi_i$ located close to the L/C boundary. In (a,b) the blue shaded regions indicate the integrals performed in the C region, whereas the red and green regions indicate the “spill-in” terms for the Hamiltonian and for the electronic density, respectively.

For the L/C boundary, these are [Fig. 1(b)]

$$n^{LL} = \sum_{i,j} D_{ij}^{LL} \phi_i^\dagger \phi_j^\dagger,$$

$$n^{LC} = \sum_{i,j} D_{ij}^{LC} \phi_i^\dagger \phi_j^\dagger,$$

which can be further distinguished based on whether both ($D^{LL}$) or just one ($D^{LC}$) of the two PAOs involved is centered at the L region. In the calculations presented in this work, the spill-in terms are independent of the applied bias voltage. This is justified, as we checked that the nonperiodic KS potential at the boundary of the C region for each value of the applied bias matches smoothly with the periodic KS potential of the neighboring reservoir, i.e., that the “screening approximation” is verified. (See Ref. [26] for additional details). We stress that including these spill-in terms is essential to ensure a stable and well-behaved convergence behavior of the SCF cycle, which turns out to be especially important for heterogeneous systems such as the Ag(100)/Si(100) interface investigated in this work.

B. Exchange-correlation potential

Further complications in describing the Ag(100)/Si(100) interface arise from the fact that one of its sides is semiconducting. In fact, a major problem affecting the description of metal-semiconductor interfaces is the severe underestimation of the semiconductor band gap in DFT calculations using (semi-)local xc functionals based on the local density approximation (LDA) or on the generalized gradient approximation (GGA) [45]. For model calculations based on few-layers-thick fully periodic systems, such an underestimation has been shown to result in unrealistically low Schottky barriers at the interface [15,46]. In order to remedy this drawback, we have evaluated the electronic structure of the LDA-optimized interface geometries using the Tran-Blaha meta-GGA xc functional (TB09) [27]. The TB09 xc functional has been shown to provide band gaps in excellent agreement with the experiments for a wide range of semiconductors including silicon, at a computational cost comparable to that of conventional (semi-)local functionals. In the TB09 xc functional, the exchange potential $\upsilon_{TB09}^{ex}(r)$ depends explicitly on the electron kinetic energy $\tau(r)$,

$$\upsilon_{TB09}^{ex}(r) = c \upsilon_{BR}^{ex}(r) + \frac{3c - 2}{\pi} \sqrt{\frac{4\tau(r)}{6\rho(r)}},$$

with $r(r) = 1/2 \sum_{i=1}^{N} |\nabla \psi_i(r)|^2$, $N$ being the total number of KS orbitals, $\psi_i(r)$ the $i$th orbital, $\rho(r)$ the electronic density, and $\upsilon_{BR}^{ex}(r)$ the Becke-Roussel exchange potential [47]. The parameter $c$ in Eq. (4) is evaluated self-consistently and takes the form

$$c = \alpha + \beta \left[ \frac{1}{\Omega} \int_{-\infty}^{\infty} \frac{|\nabla \rho(r)|}{\rho(r)} dr \right]^\frac{1}{2},$$

where $\Omega$ is the volume of the simulation cell and the two empirical parameters $\alpha = -0.012$ (dimensionless) and $\beta = 1.023$ Bohr$^4$ have been fitted to reproduce the experimental band gaps of a large set of semiconductors [27]. To obtain a description as accurate as possible of the semiconductor band gap at the Si(100) side of the interface, we have tuned the value of the $c$ parameter in order to reproduce the experimentally measured band gap of bulk silicon, $E_{gap}^{exp} = 1.17$ eV [44]. This has been accomplished by calculating the band gap of bulk silicon at fixed values of the $c$ parameter in a range around the self-consistently computed value in which the variation of $E_{gap}^{TB09}$ with $c$ is linear. Then, the optimal value of $c$ has
been determined as the intersect between the value of $E_{\text{gap}}^{\exp}$ and a linear fit to the computed values of $E_{\text{gap}}$ [Fig. 2(a)]. Using the TB09 xc functional with the $c$ parameter fixed at the optimal value determined using this procedure (hereafter, TB09-o), we calculate the indirect band gap of bulk silicon to be $E_{\text{gap}}^{\text{TB09-o}} = 1.169$ eV [Fig. 2(b)], in excellent agreement with the value 1.17 eV. The TB09-o functional has been used for all the electronic structure and transport analyses of the Ag(100)/Si(100) interface reported in this work. We have checked that the band structure of bulk silver calculated using the TB09-o functional is very similar to that calculated using the LDA, which is known to perform well for noble metals.

C. Semiconductor doping

The last requirement to describe realistically the electronic structure of the Si(100)/Ag(100) interface is to account for the doping on the silicon side of the interface. Here, doping is achieved in an effective scheme by introducing localized charges bound to the individual silicon atoms. More specifically, in ATK [39] the total self-consistent electronic density $\rho_{\text{tot}}(r)$ is defined as [40]

$$\rho_{\text{tot}}(r) = \delta \rho(r) + \sum_{l}^{N_{\text{atoms}}} \rho_{l}(r),$$

where $\sum_{l}^{N_{\text{atoms}}} \rho_{l}(r)$ is the sum of the atomic densities of the individual neutral atoms of the system. As each atomic density $\rho_{l}(r)$ is a constant term, it can be augmented with a localized “compensation” charge having the opposite sign of the desired doping density, which acts as a carrier attractor by modifying the electrostatic potential on the atom. Using these “compensation” charges, an effective doping can be achieved both in the DFT and in the DFT+NEGF simulations. In the former, the “compensation” charge added to each silicon atom is neutralized by explicitly adding a valence charge of the opposite sign, so that the system remains charge neutral. In the latter, the “compensation” charge is neutralized implicitly by the carriers provided by the reservoirs, and the system is maintained charge neutral under the condition that the intrinsic electric field in the system is zero. This effective doping scheme has the advantage of (i) not depending on the precise atomistic details of the doping impurities, and (ii) being completely independent of the size and exact geometry of the system.

III. SYSTEM

In order to obtain a reliable description of the Ag(100)/Si(100) interface, we have followed a stepwise procedure. Initially, we have carried out a preliminary screening of the interface geometries and bonding configurations by using a 2×1 slab model formed by a six-layer Ag(100) slab interfaced with a nine-layer unreconstructed Si(100) slab. The calculated bulk lattice constants of silicon ($a_{\text{Ag}} = 5.41$ Å) and silver ($a_{\text{Ag}} = 4.15$ Å) are in good agreement with those reported in the literature [23]. To match the Ag(100) and the Si(100) surface, we have applied an isotropic compressive strain $\epsilon_{xx} = \epsilon_{yy} = -0.0793$ along the surface lattice vectors $\mathbf{v}_{1,2}$ of the Ag(100) surface. We have checked that in the compressed Ag(100) structure, the dispersion of the $s$ band, and its position with respect to the $d$ band are very similar to those calculated using the equilibrium value of $a_{\text{Ag}}$. The Si(100) surface opposite to the interface has been passivated with hydrogen atoms. The geometry of the resulting 15-layer slab has then been optimized using the LDA by keeping the farthest (with respect to the interface plane) four layers of the Ag(100) surface frozen, and by allowing the farthest (with respect to the interface plane) four layers of the Si(100) slab to move as a rigid body, thereby freezing only the interatomic distances and angles. All the remaining atoms have been allowed to fully relax. Different starting guesses for the interface structure have been tested, corresponding to different configurations of the Si(100) dangling bonds with respect to the high-symmetry fcc sites of the Ag(100) surface. The lowest energy configuration among those considered, corresponding to the Si(100) dangling bonds
sitting above the “hollow” fcc sites of the Ag(100) surface, has then been used as a representative model of the interface.

Starting from the lowest energy configuration obtained using the 15-layer slab, we then constructed more realistic models of the interface. Specifically, we have expanded the bulk regions of the 15-layer slab to create two-probe setups effectively describing the infinite, nonperiodic interface [Fig. 3(a)]. A final geometry optimization has been carried out using a two-probe setup in which the C region has been described by eight Ag(100) layers and an undoped silicon layer having a total width $W_{\text{Si}(100)}^\text{CC} = 47.84 \, \text{Å}$. The optimized geometry has been used to construct two-probe setups in which the doping of the silicon side has been taken into account using the effective doping method described in Sec. II C. We have considered doping densities of $n$-type carriers ($n_d$) in the experimentally relevant range [$10^{18}$--$10^{20} \, \text{cm}^{-3}$]. As discussed in more detail in Sec. IV, the width of the Si(100) layer needed to describe accurately the interface in the two-probe simulations depends on the size of the depletion region ($W_D$) on the silicon side of the interface. The relation between $W_D$ and $n_d$ is $1/W_D \propto n_d$, so that progressively narrower C regions can be used as the doping level is increased without any loss in accuracy. Therefore, in the following, the results presented for $n_d = 10^{20} \, \text{cm}^{-3}$, $n_d = 10^{19} \, \text{cm}^{-3}$, and $n_d = 10^{18} \, \text{cm}^{-3}$ refer to calculations performed with C regions of widths $W_{\text{Si}(100)}^\text{CC} = 47.84 \, \text{Å}$, $W_{\text{Si}(100)}^\text{CC} = 197.436 \, \text{Å}$, and $W_{\text{Si}(100)}^\text{CC} = 447.92 \, \text{Å}$, respectively. We have checked that reducing the width of the C region does not have any effect on the results, as long as all the space-charge effects due to the presence of the interface take place within the screening region. Furthermore, we notice how using a two-probe setup also allows us to simulate the characteristics of the interface when the L and R reservoirs are set at two different chemical potentials $\mu_L \neq \mu_R$ due to an applied bias voltage $qV_{\text{bias}} = \mu_R - \mu_L$. As will become clear later, this allows for a direct comparison to experiments and for analyzing the electronic structure of the interface under working conditions.

Finally, to understand to which extent the slab model is able to describe accurately the electronic structure of the infinite, nonperiodic interface, we have also considered slab models having a similar interface geometry as that used in the two-probe setup [Fig. 3(b)]. Both short and long slab models have been constructed, in which the width of the Si(100) layer used to describe the silicon side of the interface has been set to either $W_{\text{slab(short)}}^\text{CC} = 38.33 \, \text{Å}$ or $W_{\text{slab(long)}}^\text{CC} = 98.62 \, \text{Å}$. Notice how these values of $W_{\text{slab}}^\text{CC}$ are many times larger than those used for similar studies of the Ag(100)/Si(100) interface reported in the literature [37].

IV. RESULTS

A. Device characteristics and validation of the activation energy model

Figure 4(a) shows the current-voltage ($I-V_{\text{bias}}$) characteristics calculated for the two-probe setup at low ($n_d = 10^{18} \, \text{cm}^{-3}$), intermediate ($n_d = 10^{19} \, \text{cm}^{-3}$), and high ($n_d = 10^{20} \, \text{cm}^{-3}$) doping densities of the Si(100) side of the interface. A strong dependence on the doping density is evident. At low doping, the interface shows a well-defined Schottky diode-like behavior: the forward bias ($V_{\text{bias}} > 0 \, \text{V}$) current increases about 6 orders of magnitude in the range of $V_{\text{bias}}$ [0.02--0.5 V], whereas the reverse bias one ($V_{\text{bias}} < 0 \, \text{V}$) varies only within 1 order of magnitude in the corresponding range. The diode-like asymmetry in the $I-V_{\text{bias}}$ curves persists at intermediate doping, although it is less pronounced than at low doping, the current at forward bias and reverse bias varying within 3 and 2 orders of magnitude, respectively. The scenario changes qualitatively at high doping as the $I-V_{\text{bias}}$ curve becomes highly symmetric, suggesting an Ohmic behavior of the interface.

According to thermionic emission theory, the $I-V_{\text{bias}}$ characteristics of a Schottky diode can be described by [2]

$$ I = I_0 \left( e^{qV_{\text{bias}}/n k_B T} - 1 \right), \tag{7} $$

where $q$ is the elementary charge, $k_B$ is the Boltzmann constant, $T$ is the temperature, $I_0$ is the saturation current, and $n$ is the so-called ideality factor. The latter accounts for the deviation of the $I-V_{\text{bias}}$ characteristics from those of an ideal diode, for which $n = 1$. Fitting the simulated data at forward bias to Eq. (7) allows to extract $n$ from the slope of the fitted curves. In Fig. 4(b) the fitted curves are compared to the forward bias data. The latter are presented using an
The $I-V_{\text{bias}}$ simulation allows us to test the reliability of the experimental procedures used to extract the Schottky barrier $\Phi$. In particular, we focus on the so-called activation-energy (AE) method, which does not require any a priori assumption on the electrically active interface area $A$ [2]. In the AE method the $I-T$ dependence is measured at a small constant $V_{\text{bias}}$. Over a limited range of $T$ around room temperature, assuming that the Richardson constant $A^*$ and $\Phi$ are constant, the $I-T$ characteristics can be described by the expression

$$IT^{-2} = A^* e^{-\Phi/\Delta_1} e^{\Phi_{\text{AE}}(n)/T}. \tag{9}$$

Following Eq. (9), the Schottky barrier height $\Phi^\text{AE}$ can be extracted from the ln$(I/T^2)$ vs 1/$T$ data using

$$-\frac{k_B}{q} \frac{d[\ln(I/T^2)]}{d(1/T)} = \frac{\Phi^\text{AE} - V_{\text{bias}}}{n}, \tag{10}$$

in which $n$ is the ideality factor extracted above.

Figures 5(a) and 5(b) show the simulated AE plots (as Arrhenius plots) at different values of $V_{\text{bias}}$ for low and intermediate doping densities, at which the interface still displays clear Schottky diodelike characteristics. The $I-T$ dependence has been evaluated in a linear response fashion, using the Landauer-Büttiker expression for the current, $I = \frac{2e}{h} \int T(E, \mu_L, \mu_R) \left[ f\left(\frac{E - \mu_L}{k_B T}\right) - f\left(\frac{E - \mu_R}{k_B T}\right)\right] dE$ with the transmission coefficient $T(E, \mu_L, \mu_R)$ evaluated self-consistently at an electron temperature of 300 K. Fully self-consistent simulations performed for selected temperatures show that this approach is valid within the range of $T$ considered, 250 K $\leq T \leq 400$ K.

Ideally, for a given doping the Schottky barrier depends exclusively on the M-S energy-level alignment at the interface and therefore, disregarding image-force effects, should remain constant with $V_{\text{bias}}$ [2]. This implies that in Eq. (10), the left-hand side should equal the right-hand side at any value of $V_{\text{bias}}$. However, in the present case this condition is not verified, as the variation of the right-hand-side term with $V_{\text{bias}}$ is larger than that of the left-hand-side term [see Figs. 5(c) and 5(d)]. Indeed, for $n_d = 10^{18} \text{ cm}^{-3}$ ($n_{d} = 10^{19} \text{ cm}^{-3}$), a linear fit to the calculated values of the left-hand side of Eq. (10) gives a slope of $-664 \text{ meV/V}$ ($-177 \text{ meV/V}$), whereas the slope associated to the variation of the right-hand-side term is $-917 \text{ meV/V}$ ($-549 \text{ meV/V}$).

Following the procedure in the AE method, we use the value of $n$ obtained from Eq. 4 to subtract the bias dependence. The result is shown in Figs. 5(e) and 5(f) and it can be seen that this leads to an unphysical increase of $\Phi^\text{AE}$ with $V_{\text{bias}}$. The error becomes more severe as $n_d$ is increased. At low (intermediate) doping, $\Phi^\text{AE}$ varies from by 30% (325%) in the range of $V_{\text{bias}}$ considered, leading to a change $\Delta \Phi^\text{AE} = 3.73 \! k_B T$ ($\Delta \Phi^\text{AE} = 7.31 \! k_B T$). Thus, the intrinsic accuracy of the AE method depends strongly on multiple factors. On the one side, the nonlinear increase in $\Phi^\text{AE}$ with $V_{\text{bias}}$ suggests that using a single value of $V_{\text{bias}}$ is not sufficient to obtain an accurate estimate of $\Phi$. On the other side, the change in $\Delta \Phi^\text{AE}$ with $n_d$ at a given $V_{\text{bias}}$ indicates that the AE method is unsuited for comparative analyses of the variation of $\Phi$ with doping. These facts call for a more direct and general strategy for the characterization of M-S interfaces under working conditions.

**Fig. 4.** Calculated $I-V_{\text{bias}}$ (a) and forward bias $I/(1-e^{-qV_{\text{bias}}/k_B T}) - V_{\text{bias}}$ (b) characteristics at $n_d = 10^{18} \text{ cm}^{-3}$ (blue triangles), $n_d = 10^{19} \text{ cm}^{-3}$ (green squares), $n_d = 10^{20} \text{ cm}^{-3}$ (red dots). In (a), the values of $I$ at $n_d = 10^{18} \text{ cm}^{-3}$ and $n_d = 10^{19} \text{ cm}^{-3}$ have been multiplied by a factor of 10 and 100, respectively. The solid lines in (b) are fit to the data in the range $0.02 \leq V_{\text{bias}} \leq 0.08 \text{ V}$ using Eq. (7). The ideality factor $n$ extracted from the slope of each fitted curve is reported using the same color as the corresponding curve.

 alternative form of Eq. (7),

$$I = I_0 \exp\left(\frac{qV_{\text{bias}}}{n_k T}\right) \left(1 - e^{-qV_{\text{bias}}/n_k T}\right), \tag{8}$$

which allows for a better comparison with the fitted curves as $I/(1-e^{-qV_{\text{bias}}/k_B T})$ varies exponentially with $V_{\text{bias}}$ in the fitting interval considered, viz. $0.02 \text{ V} \leq V_{\text{bias}} \leq 0.08 \text{ V}$.

At low doping, $n = 1.09$, indicating that the system behaves essentially as an ideal Schottky diode. At intermediate doping, $n = 1.82$, and the system deviates significantly from the ideal behavior. At high doping, $n = 2.40$, consistent with the observation that the system does not behave anymore as a Schottky diode.
FIG. 5. (a, b) Empty dots: calculated $I$-$T$ data at different bias voltages for $n_d = 10^{18}$ cm$^{-3}$ (a) and $n_d = 10^{19}$ cm$^{-3}$ (b). Solid lines: fit to the simulated data using Eq. (10). (c, d) Left-hand side (filled dots) and right-hand side (dashed line) of Eq. (10) as a function of $V_{\text{bias}}$. The values of the left-hand side have been extracted from the slope of the fitted $I$-$T$ curves in (a, b). The solid lines are linear fits to the data. The right-hand side of Eq. (10) has been plotted using the value of $\Phi_{1\text{e}}$ calculated at $V_{\text{bias}} = 0$ V, which approaches the value of $\Phi_{1\text{e}}$ at $V_{\text{bias}} = 0$ V. (e, f) Schottky barrier height $\Phi_{1\text{e}}$ evaluated using Eq. (10) as a function of $V_{\text{bias}}$.

B. Electronic properties of the interface

A strong advantage of the DFT+NEGF simulations is that they allow the visualization of the electronic structure of the interface and the direct tracking of its changes when $n_d$ and $V_{\text{bias}}$ are varied. This makes it possible to analyze the calculated $I$-$V_{\text{bias}}$ characteristics in terms of the electronic structure of the interface.

Figure 6 shows the local density of states [48] (LDOS) of the two-probe model at equilibrium (i.e., at $V_{\text{bias}} = 0$ V) along the direction normal to the interface at the different doping densities considered. Increasing the doping has a twofold effect on the electronic properties of the system: on the one side, $W_D$ decreases from $\sim 200$ nm to $\sim 20$ nm when the doping is increased from $n_d = 10^{18}$ cm$^{-3}$ to $n_d = 10^{20}$ cm$^{-3}$, as a direct consequence of the increased screening of the $n$-doped silicon. Increasing $n_d$ also shifts the Fermi level towards the silicon conduction bands. In particular, at $n_d = 10^{18}$ cm$^{-3}$ the conduction band minimum (CBM) of silicon at $Z > W_D$ lies at $E - \mu_{L,R} = +20$ meV, whereas at $n_d = 10^{19}$ cm$^{-3}$ and $n_d = 10^{20}$ cm$^{-3}$ it lies at $E - \mu_{L,R} = -40$ meV and $E - \mu_{L,R} = -100$ meV, respectively. It is also worth noticing how the macroscopic average [49] of the Hartree potential along the direction normal to the interface, $\langle V_H \rangle$ (blue lines in Fig. 6), follows the profile of the silicon CBM close to as well as far away from the interface. Similarly to what happens for the electronic bands, $\langle V_H \rangle$ becomes constant at $Z > W_D$, indicating that the electronic structure starts to resemble that of the infinite periodic bulk. A closer analysis also reveals that a finite density of states extends considerably on the

FIG. 6. Local density of states (LDOS) of the two-probe setup at equilibrium for $n_d = 10^{18}$ cm$^{-3}$ (a), $n_d = 10^{19}$ cm$^{-3}$ (b), and $n_d = 10^{20}$ cm$^{-3}$ (c). The energy on the vertical axis is relative to the system chemical potential $\mu_{L,R}$. Regions of low (high) LDOS are shown in dark (bright) color. The blue line in each panel indicates the macroscopic average of the Hartree potential $\langle V_H \rangle$ subtracted from the electron affinity of bulk Si and $\mu_{L,R}$. The yellow vertical line in each panel indicates the associated $\Phi_{1\text{e}}$.
semiconductor side of the interface due to penetration of the metallic states into the semiconductor side [50–52].

Due to the lack of a well-defined electronic separation between the metal and the semiconductor, it is difficult to provide an unambiguous value for $\Phi$ based on the electronic structure data only. However, due to the fact that $\langle V_H \rangle$ closely traces the CBM, it is still possible to estimate the Schottky barrier by defining $\Phi_{pot}$ as the difference between $\mu_L$ and the maximum of $\langle V_H \rangle$ on the semiconductor side of the interface, $\langle V_H^{(max)} \rangle$ (see Fig. 6).

We calculate $\Phi_{pot} = 412 \text{ meV}$ and $\Phi_{pot} = 342 \text{ meV}$ for $n_d = 10^{18} \text{ cm}^{-3}$ and $n_d = 10^{19} \text{ cm}^{-3}$, respectively. For $n_d = 10^{20} \text{ cm}^{-3}$ the barrier is considerably lower, $\Phi_{pot} = 133 \text{ meV}$, reflecting the more pronounced Ohmic behavior observed in the $I-V_{bias}$ curves. Focusing on the low and intermediate doping cases, it can be noticed how the values of $\Phi_{pot}$ are considerably larger than those of $\Phi_{AE}$ at $V_{bias} \rightarrow 0 \text{ V}$. In particular, at low doping $\Phi_{pot} - \Phi_{AE} = 112 \text{ meV}$, whereas at intermediate doping the difference is even larger, $\Phi_{pot} - \Phi_{AE} = 286 \text{ meV}$.

A consistent physical picture that rationalizes the $I-V_{bias}$ curves can be obtained by studying the doping dependence of the spectral current $I(E) = \frac{2e}{h} T(E, \mu_L, \mu_R) f(\frac{E-E_{\mu_L}}{k_B T}) - f(\frac{E-E_{\mu_R}}{k_B T})$. Figures 7(b) and 7(d) show the profiles of $\langle V_H \rangle$ obtained at forward bias in the bias range $0.02 \text{ V} < V_{bias} < 0.2 \text{ V}$ for low and intermediate doping densities. As $V_{bias}$ is increased, $\langle V_H^{(max)} \rangle$ shifts towards higher energies due to image-force effects [2], and $\langle V_H \rangle$ becomes progressively flatter on the semiconductor side. The overall result of these changes is a decrease of the barrier $\phi_F$ associated with the thermionic emission process from the Si(100) conduction band to Ag(100) [see Fig. 7(a)]:

$$\phi_F = \Phi - V_{bias}/n.$$  \hspace{1cm} (11)

The associated spectral currents $I(E)$ are shown in Figs. 7(c) and 7(e). For an interface in which the only contribution to transport comes from thermionic emission, $I(E)$ should be nonzero only at $E - \mu_L > \Phi_{pot}$. However, in the present case $I(E)$ is finite also at $E - \mu_L < \Phi_{pot}$, indicating that electron tunneling has a non-negligible contribution to $I$. This contribution is much larger for intermediate than for low doping densities. Indeed, at $V_{bias} \rightarrow 0 \text{ V}$, the position of $E(I^{(max)})$ lies very close to $\Phi_{pot}$ in the low-doping case, as expected in the case of a nearly ideal Schottky diode. Conversely, at intermediate doping $E(I^{(max)})$ lies well below $\Phi_{pot}$, indicating that electron tunneling has become the dominant transport process.

The trend of $I(E)$ with $V_{bias}$ is consistent with these considerations. At low doping, $E(I^{(max)})$ is pinned to $\langle V_H^{(max)} \rangle$, whereas the onset of finite $I(E)$ at $E - \mu_L < \Phi_{pot}$ moves towards higher energies, following the variation of $\langle V_H \rangle$. On the other hand, at intermediate doping the overall shape of $I(E)$ remains the same as $V_{bias}$ is increased, and the variation of $E(I^{(max)})$ follows closely that of $\langle V_H \rangle$. We also notice the presence of a narrow resonance at $E - \mu_L = +0.395 \text{ eV}$, whose position is independent of $n_d$ and $V_{bias}$. This is due to a localized electronic state at the interface which is pinned to $\mu_L$.

The variation of $\phi_F$ with $V_{bias}$ can be related to the slope-dependent term of Eq. (10) by assuming $\Phi = \Phi_{AE}$ in Eqs. (10) and (11), thus allowing for a direct comparison with the AE data (see Fig. 8). Independently of the value of $V_{bias}$, the slope-dependent term lies always below $\phi_F$ due to the missing contribution of electron tunneling in the AE method: the latter assumes that the current has a purely thermionic origin, and consequently predicts a value of $\phi_F$ lower than the actual one. In agreement with the previous analyses, this deviation is considerably larger in the intermediate doping case, due to the much larger contribution of electron tunneling.

C. Comparison of the two-probe with the slab model

The results obtained using the two-probe model can be used as a reference to validate the use of finite-size models to describe the Ag(100)/Si(100) interface. Such models are integral parts of the band alignment method often used to...
evaluate $\Phi$ using conventional DFT [53–56]. The method relies on aligning the electronic band structures of the two bulk materials forming the interface on an absolute energy scale by using a reference quantity, often $\langle V_H \rangle$ [49]. The perturbation of the bulk electronic structure in each material due to the presence of the interface is accounted for by either a slab [37] or a fully periodic [57] model. $\langle V_H \rangle$ is then used as a common reference to align the electronic structure obtained from independent calculations of the two bulk materials. Despite its widespread use, this strategy relies on two drastic assumptions. First, it is implicitly assumed that the electronic properties of the interface are independent of the doping level of the semiconductor. Moreover, it is assumed that the electronic properties in the central part of each side of the interface model are a good approximation to those of the two bulk materials.

Figure 9 shows a comparison between $\langle V_H \rangle$ obtained at the different doping densities considered for the two slab models (short and long, see Sec. III) and for the two-probe setup. We notice that introducing an effective doping in the slab model, which was not taken into account in previous slab models for Ag/Si interfaces [37], attempts to better mimic the two-probe simulation in which silver is interfaced with n-doped silicon. The profiles of $\langle V_H \rangle$ of the three different systems have been aligned according to the value of $\mu_L$, the side at which Dirichlet boundary conditions are used for the three systems. Irrespective of the doping level, the doped short slab model provides a poor description of the variation of $\langle V_H \rangle$ at the interface. In particular, $\langle V_{H}^{\text{max}} \rangle$ is always $\sim200$ meV higher that obtained with the two-probe model. Furthermore, on the Si(100) side of the interface, $\langle V_H \rangle$
going from the short to the long slab model come at the expense of a much higher computational cost. In fact, each DFT calculation for the long slab model takes on average 338.6 s/step. Conversely, each DFT+NEGF calculation using the two-probe model is approximatively 1 order of magnitude faster, taking on average 46.6 s/step. This suggests that, in addition to computational efficiency, there are also more fundamental reasons for making DFT+NEGF the method of choice for describing M-S interfaces, as in the two-probe setup the two main assumptions of the band alignment method are naturally lifted. We emphasize that, although the results presented in this paper are specific to the Ag(100)/Si(100) interface only, similar conclusions are likely to hold true for all systems for which the poor screening on the semiconductor side of the interface results in space-charge effects that extend over widths of the order of several nanometers.

V. CONCLUSIONS

In this work, we have presented an approach based on density functional theory (DFT) and nonequilibrium Green’s functions (NEGF) for realistic metal-semiconductor (M-S) interfaces modeling. Our approach is designed to deal effectively and correctly with the nonperiodic nature of the interface, with the semiconductor band gap and with the doping on the semiconductor side of the contact, and allows for a direct theory-experiment comparison as it can simulate $I-V_{\text{bias}}$ characteristics. Using a Ag/Si interface relevant for photovoltaic applications as a model system, we have shown that our approach is a better alternative to (i) analytical approaches such as the “activation energy” (AE) method to analyze the $I-V_{\text{bias}}$ characteristics of nonideal rectifying systems with non-negligible tunneling contribution, and (ii) finite-size slab models to describe the interface between metals and doped semiconductors. This DFT+NEGF approach could pave the way for an understanding of M-S interfaces beyond the limitations imposed by traditional analytical and atomistic methods.

ACKNOWLEDGMENTS

QuantumWise acknowledges support from a Innovation Fund Denmark Grant No. 79-2013-1 “Nano-Scale Design Tools for the SemiConductor Industry”. The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under Grant Agreement III-V-MOS Project No. 619326. We thank Walter A. Harrison, Andreas Goebel, and Paul A. Clifton at Acorn Technologies for their input on this work. D.S. acknowledges support from the H.C. Ørsted-COFUND postdoc program at DTU.
The LDOS is obtained by projecting the density of states on each atomic layer in the $Z$ direction and then interpolating the layer-resolved density of states to obtain a continuous LDOS. An exponential scale is used for the LDOS intensity. While this procedure provides a good representation of the silicon band gap throughout the device, the finer details of the LDOS plots depend on (i) the interpolation procedure used to obtain a continuous LDOS and (ii) the exponential scale used in the interpolated LDOS plots.

[48] The LDOS is obtained by projecting the density of states on each atomic layer in the $Z$ direction and then interpolating the layer-resolved density of states to obtain a continuous LDOS. An exponential scale is used for the LDOS intensity. While this procedure provides a good representation of the silicon band gap throughout the device, the finer details of the LDOS plots depend on (i) the interpolation procedure used to obtain a continuous LDOS and (ii) the exponential scale used in the interpolated LDOS plots.