Electronic structure of semiconductor nanowires

Y. M. Niquet,[1,*] A. Lherbier,[1] N. H. Quang,[1,2] M. V. Fernández-Serra,[3] X. Blase,[3] and C. Delerue[4]
[1]*Département de Recherche Fondamentale sur la Matière Condensée, SP2M/L_Sim, CEA Grenoble, 38054 Grenoble Cedex 9, France*
[2]*Institute of Physics and Electronics, Vietnamese Academy of Science and Technology, 10 Dao-Tan, Ba-Dinh, Hanoi 10000, Vietnam*
[3]*Laboratoire de Physique de la Matière Condensée et Nanostructures, Université Claude Bernard Lyon 1 and UMR CNRS 5586, Bâtiment Brillouin, 43 boulevard du 11 Novembre 1918, 69622 Villeurbanne, France*
[4]*Institut d'Électronique, de Micro-électronique et de Nanotechnologie (UMR CNRS 8520), Département ISEN, 41 boulevard Vauban, F-59046 Lille Cedex, France*
(Received 9 November 2005; revised manuscript received 13 February 2006; published 18 April 2006)

We compute the subband structure of several group IV and III-V ⟨001⟩-, ⟨110⟩-, and ⟨111⟩-oriented nanowires using sp^3 and $sp^3d^5s^*$ tight-binding models. In particular, we provide the band gap energy of the nanowires as a function of their radius R in the range $R=1-20$ nm. We then discuss the self-energy corrections to the tight-binding subband structure, that arise from the dielectric mismatch between the nanowires (with dielectric constant ε_{in}) and their environment (with dielectric constant ε_{out}). These self-energy corrections substantially open the band gap of the nanowires when $\varepsilon_{\text{in}}>\varepsilon_{\text{out}}$, and decrease slower ($\propto 1/R$) than quantum confinement with increasing R. They are thus far from negligible in most experimental setups. We introduce a semi-analytical model for practical use. This semianalytical model is found in very good agreement with tight-binding calculations when $\varepsilon_{\text{in}}>\varepsilon_{\text{out}}$.

DOI: 10.1103/PhysRevB.73.165319 PACS number(s): 73.21.Hb, 73.22.Dj

I. INTRODUCTION

The vapor-liquid-solid[1] (VLS) and related[2,3] growth mechanisms have allowed the synthesis of high-quality, free-standing nanowires of almost every usual group IV, III-V, or II-VI semiconductor,[4] including Si,[5,6] Ge,[5] InAs,[3] GaAs,[2,7] InP,[8,9] and GaP,[10] etc. The diameter of these nanowires typically ranges from a few to a few tens of nanometers, while their length can exceed a few micrometers. Zinc-blende and diamondlike nanowires usually grow along ⟨001⟩, ⟨110⟩, or ⟨111⟩ crystallographic directions depending on the size and growth conditions. The VLS approach is very versatile. Indeed, the composition of the nanowires can be modulated along the growth axis,[11–13] which enables the synthesis of various nanowire "heterostructures" and superlattices with embedded quantum dots or tunnel barriers. The nanowires can also be encapsulated[14] in one or more shells of other material(s) (usually with larger band gaps), that move surface traps away from the cores and can be used as delta doping layers. VLS nanowires thus afford plenty of opportunities for new and original experiments probing one-dimensional physics. They have already unveiled a wide range of interesting optical and transport properties,[15] such as strong luminescence polarization[16] or clean Coulomb blockade features at low temperature.[17,18]

VLS nanowires are also very attractive for "bottom-up" nanoelectronics. They are polyvalent building blocks that can serve both as devices and connectors. As a matter of fact, many prototypes of such devices have been realized in just a few years, including *p-n* and resonant tunneling diodes,[11,19] bipolar[20] and field-effect transistors,[21–23] detectors,[24] etc. The advent of a nanowire-based electronics, however, calls for a better control of their sizes and positions, progress being made step by step along that way.[25–27]

The quasiparticle (or "one-particle") subband structure is a key to the understanding of charge transport in semiconducting nanowires. Most features of the $I(V)$ or conductance characteristics (such as the current onsets) are directly related to the band gap and subband energies. The latter depend on the size, shape, and environment of the nanowires both through quantum and "dielectric" confinement. Most VLS nanowires indeed exhibit large dielectric mismatches with their surroundings, that are responsible for so-called self-energy corrections to the subband structure.[28,29] As we shall see, these corrections are far from negligible in most experimental situations.

Many calculations have been reported on semiconducting nanowires. The structural, electronic, and optical properties of very small nanowires have been computed with *ab initio* methods[30–34] such as density-functional theory (DFT).[35,36] These studies mostly focused on silicon, starting with porous Si more than ten years ago. The band gap energies and optical properties of larger nanowires have been calculated with various semiempirical methods such as $\mathbf{k}\cdot\mathbf{p}$ theory,[37] pseudopotentials,[38,39] and tight-binding,[40–42] without, however, dealing with the self-energy problem. The transport properties of small nanowires have been addressed much more recently with *ab initio*[43,93] and semiempirical[44] methods, still neglecting the self-energy corrections. The latter have been computed last year[45] in very small silicon nanowires using the *ab initio GW* approach.[46] They had been discussed previously with semiclassical image charge models in the context of the exciton[47] or donor[48,49] binding energies, but a systematic investigation of their impact on the subband structure has not been carried out before.

In this paper, we compute the tight-binding (TB) subband structure of ⟨001⟩-, ⟨110⟩-, and ⟨111⟩-oriented group IV and III-V cylindrical nanowires. In particular, we compare sp^3

and $sp^3d^5s^*$ TB models in a wide range of diameters (1–40 nm), and give analytical fits to the conduction and valence band edge energies for practical use. We also check our results for silicon nanowires against the local density approximation (LDA). We then discuss the self-energy problem. We introduce a semianalytical model that yields the self-energy correction to the band gap energy of a nanowire with radius R embedded in a medium with dielectric constant ε_{out}. This model, supported by tight-binding calculations, shows that the self-energy corrections decrease as $1/R$, slower that quantum confinement.

The tight-binding subband structure of the nanowires is discussed in Sec. II, while the self-energy corrections are discussed in Sec. III.

II. QUANTUM CONFINEMENT

In this section, we first introduce the tight-binding models used throughout this work (Sec. II A), then discuss the subband structure of Si, Ge, InAs, GaAs, InP, and GaP nanowires (Sec. II B); we last compare our TB results with experimental data and *ab initio* calculations in Sec. II C.

A. Tight-binding models

The nanowires (NWs) are carved out of the bulk zinc-blende or diamond crystal keeping all atoms inside a cylinder of radius R centered on a cation. They are one-dimensional periodic structures with unit cell length $\ell=a$ ($\langle 001\rangle$ NWs), $\ell=a/\sqrt{2}$ ($\langle 110\rangle$ NWs), or $\ell=a\sqrt{3}$ ($\langle 111\rangle$ NWs), a being the bulk lattice parameter.[50] We define the actual R as the radius of the cylinder of length ℓ whose volume is the same as the average volume $\Omega=N_{\text{sc}}a^3/8$ occupied by the N_{sc} semiconductor atoms of the unit cell:

$$R=\sqrt{\frac{N_{\text{sc}}a^3}{8\pi\ell}}. \tag{1}$$

The dangling bonds at the surface of the nanowires are saturated with hydrogen atoms. This approach, though missing surface reconstruction effects, has already proved to be highly successful in describing intrinsic quantum confinement in semiconductor nanocrystals.[51–53]

The principle of the semiempirical tight-binding method is to expand the quasiparticle wave functions onto a basis of atomic orbitals.[29,54] The Hamiltonian matrix elements between neighboring orbitals are considered as adjustable parameters usually fitted to reproduce bulk band structures, then transferred to the nanostructures. A given TB model can thus be characterized by its atomic basis set (e.g., orthogonal sp^3) and by its range (matrix elements up to first, second, or third nearest neighbors with two or three center integrals[55]). We have compared (when possible) various TB models and parametrizations available in the literature to assess the robustness of our results. We have indeed used a first-nearest neighbor two center orthogonal $sp^3d^5s^*$ TB model and a second or third nearest neighbors three center orthogonal sp^3 TB model. The parameters of the $sp^3d^5s^*$ TB model are taken from Jancu *et al.* (JMJ model, Ref. 56 for all materials) and from Boykin *et al.* (TBB model, Ref. 57 for Si/Ge and Ref. 58 for InAs/GaAs), while those of the sp^3 TB model are taken from Niquet *et al.* (YMN model, Ref. 51 for Si, Ref. 52 for Ge, and Ref. 53 for InAs). These models reproduce the overall bulk band structures (and experimental bulk band gap energies); the TBB and YMN models, however, achieve better accuracy on the electron and hole effective masses than the JMJ model. Hydrogen TB parameters are taken from Ref. 51 for Si, Ref. 52 for Ge, and Ref. 53 for III-V materials. Spin-orbit coupling is taken into account in all calculations, unless otherwise stated.

Due to translational symmetry, the quasiparticle energies can be sorted into subbands labeled by a wave vector $\mathbf{k}=k\mathbf{u}$ and a band index n, where $\mathbf{u}$ is the unit vector oriented along the nanowire and $k\in[-\pi/\ell,\pi/\ell[$. The TB problem then reduces (for each k) to the diagonalization of a sparse matrix of order $n\sim N_{\text{sc}}\times N_{\text{orb}}$, where N_{orb} is the number of orbitals per atom. Practically, a finite number of conduction and valence subbands are computed around the gap using an iterative Jacobi-Davidson algorithm[59,60] as described in Appendix A.

B. Results

For practical use, the conduction and valence band edge energies $\varepsilon_{\text{c}}(R)$ and $\varepsilon_{\text{v}}(R)$ have been fitted for each material and nanowire orientation with the following expression:

$$\varepsilon(R)-\varepsilon(\infty)=\frac{K}{R^2+aR+b}, \tag{2}$$

where K, a, and b are adjustable parameters, and $\varepsilon(\infty)$ is the bulk band edge [here $\varepsilon_{\text{v}}(\infty)=0$ and $\varepsilon_{\text{c}}(\infty)$ is the bulk band gap energy $E_{\text{g,b}}$[50,61]]. This expression, while having the $\propto 1/R^2$ behavior at large R expected from $\mathbf{k}\cdot\mathbf{p}$ theory,[62] allows for a slower confinement at small R. The values of K, a, and b are reported in Tables I and II. Unless specified, the fit was made simultaneously on all TB data (JMJ, TBB, and YMN models when available) in the range $R=1-20$ nm. It thus represents an average of the TB results, the scattering being usually well within reasonable bounds (see discussion in Sec. II C). We would like to emphasize that Tables I and II should also hold[51] for quasicircular (e.g., facetted) nanowires provided R is still chosen to match the cross-sectional area of the nanowire.

We next discuss specific features (effective masses, etc.) of the subband structure of Si, Ge, and III-V nanowires.

1. Si nanowires

The subband structure of a [111]-oriented Si nanowire with radius $R=3.75$ nm is shown as a typical example in Fig. 1. While the valence band maximum lies at $k=0$, the conduction band minimum falls around $k=0.4\pi/\ell$, except in the smallest nanowires ($R<1$ nm) where it shifts to $k\sim 0$. As a matter of fact, bulk silicon has six equivalent conduction band minima located around $\pm 0.8\Gamma X$ in the bulk Brillouin zone[50,63] (i.e., around $\pm 1.6\pi/a$ along $\langle 001\rangle$ directions). The energy isosurfaces around these minima are ellipsoids elongated along the ΓX axes (electrons have a heavy mass $m_l^*=0.92\,m_0$ along ΓX and a light mass $m_t^*=0.19m_0$ perpendicu-

TABLE I. The parameters K, a, and b [see Eq. (2)] for silicon and germanium nanowires. Δ is the splitting between the main conduction band valleys in silicon nanowires (see Sec. II B 1).

Material	NW		K (eV nm^2)	a (nm)	b (nm^2)
Si	⟨001⟩	ε_v	−0.8825	1.245	0.488
		ε_c	0.6589	0.235	0.142
		Δ[a]	0.4546	0.614	0.457
	⟨110⟩	ε_v	−0.6825	2.062	0.996
		ε_c	0.6470	0.123	0.849
		Δ[a]	0.2213	−0.016	−0.038
	⟨112⟩	ε_v	−0.7075	2.616	−0.083
		ε_c	0.7273	0.246	0.313
	⟨111⟩	ε_v	−0.6964	3.664	−0.374
		ε_c	0.8010	0.342	0.212
Ge	⟨001⟩	ε_v	−1.8294	2.938	0.159
		ε_c	1.8505	0.930	0.640
	⟨110⟩	ε_v	−1.6159	4.310	0.579
		ε_c	1.3299	0.825	0.873
	⟨111⟩	ε_v	−1.5720	5.007	0.640
		ε_c	1.5161	0.746	0.717

[a]Fitted on TBB and YMN models.

TABLE II. The parameters K, a, and b [see Eq. (2)] for InAs, GaAs, InP, and GaP nanowires.

Material	NW		K (eV nm^2)	a (nm)	b (nm^2)
InAs	⟨001⟩	ε_v	−1.2243	2.181	0.583
		ε_c	10.6948	6.512	2.406
	⟨110⟩	ε_v	−1.0051	2.548	1.379
		ε_c	10.5845	6.645	2.879
	⟨111⟩	ε_v	−0.9887	2.720	0.898
		ε_c	10.6951	6.512	2.773
GaAs	⟨001⟩	ε_v	−1.0320	2.026	0.401
		ε_c	3.4198	1.786	1.103
	⟨110⟩	ε_v	−0.8035	2.608	1.079
		ε_c	3.3448	1.615	1.915
	⟨111⟩	ε_v	−0.8678	2.731	0.935
		ε_c	3.4098	1.806	1.467
InP	⟨001⟩	ε_v	−0.9179	2.058	0.008
		ε_c	3.0422	1.620	1.144
	⟨110⟩	ε_v	−0.8109	3.327	1.148
		ε_c	3.0365	1.710	1.460
	⟨111⟩	ε_v	−0.8273	4.465	0.223
		ε_c	3.0696	1.798	1.214
GaP	⟨001⟩	ε_v	−0.8314	1.469	0.126
		ε_c	0.3780	−0.397	0.287
	⟨110⟩	ε_v	−0.6794	2.203	1.669
		ε_c	0.3810	−0.203	0.486
	⟨111⟩	ε_v	−0.6942	3.557	0.287
		ε_c	0.5589	0.000	0.285

lar to ΓX, m_0 being the free electron mass). These six minima all project onto $k \simeq \pm 1.6\pi/\ell$ in ⟨111⟩-oriented nanowires, or equivalently onto $k \simeq \pm 0.4\pi/\ell$ since the subband structure is periodic in reciprocal space (with period $L=2\pi/\ell$). Note that intervalley couplings actually split the sixfold degenerate bulk conduction band minima into three subbands. The splittings are, however, fairly small and highly dependent on the detailed geometry of the nanowire, the three subbands lying within $\simeq 0.1$ meV for $R=5$ nm and up to $\simeq 15$ meV for $R=1$ nm. The conduction and valence band edge energies $\varepsilon_c(R)$ and $\varepsilon_v(R)$ are plotted as a function of R in Fig. 2 for three different TB models. The solid line is the fit to the TB data [Eq. (2) and Table I]. As expected, the band gap opens with decreasing R due to quantum confinement. The three TB models are found in reasonable agreement with each other when $R \gtrsim 1.5$ nm.

The effective mass (lowest conduction subband) of the electrons along ⟨111⟩-oriented Si nanowires is $m^* \simeq 0.4 m_0$, increasing for $R \lesssim 2$ nm (see Fig. 3). The hole mass (highest valence subband) is close to the light hole mass in bulk Si, $m^* \simeq 0.15 m_0$. Again, the sp^3 YMN model and $sp^3d^5s^*$ TBB model, that correctly reproduce the bulk effective masses, are found in good agreement. The $sp^3d^5s^*$ JMJ model (not shown), that does not reproduce the bulk effective masses as well, is slightly off, but shows exactly the same trends.

In [001]-oriented Si nanowires four of the six bulk conduction band minima (those along [100] and [010]) project onto $k=0$, while the last two minima (along [001]) again fold back at $k \simeq \pm 0.4\pi/\ell$. The actual conduction band minimum falls at $k=0$ because the electrons around the [001] minima are light in the plane normal to the wire and thus have a higher energy than the electrons around the [100] and [010] minima with mixed heavy and light character. Accordingly, electrons with $k \sim 0$ are light along the nanowire ($m^* \simeq m_t^*$), while electrons with $k \sim \pm 0.4\pi/\ell$ are heavy ($m^* \simeq m_l^*$). The contribution of each valley to the transport properties of a doped ⟨001⟩-oriented Si nanowire depends on their

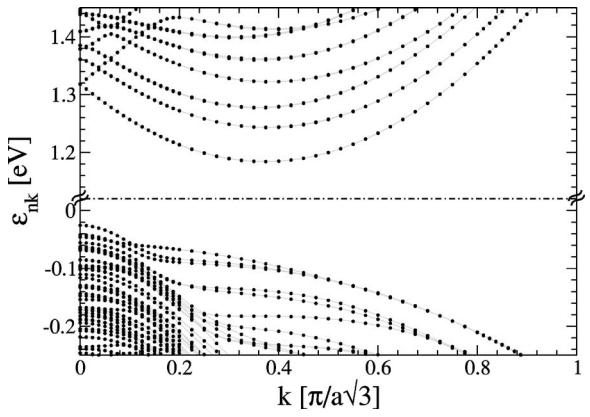

FIG. 1. Band structure of a ⟨111⟩-oriented Si nanowire with radius $R=3.75$ nm ($sp^3d^5s^*$ TBB model).

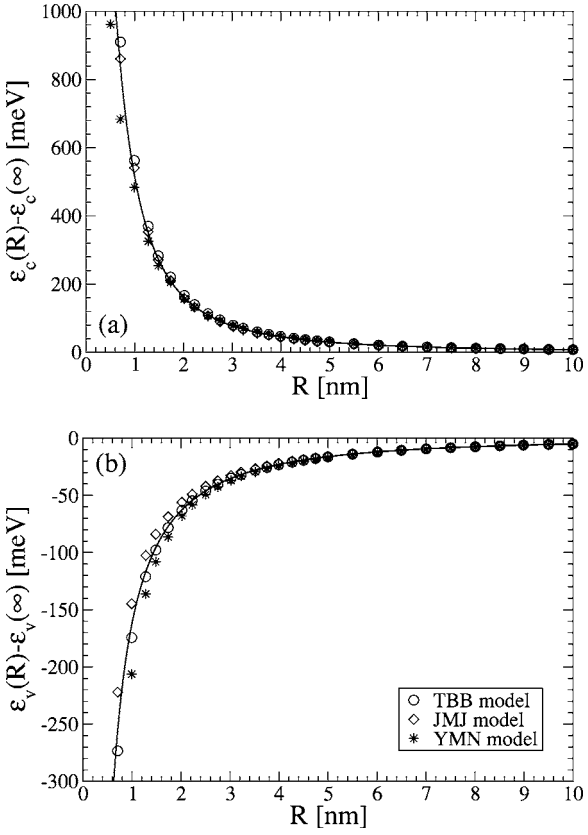

FIG. 2. (a) Electron and (b) hole confinement energies in ⟨111⟩-oriented Si nanowires, in the $sp^3d^5s^*$ TBB and JMJ models, and in the sp^3 YMN model. $\varepsilon_c(\infty)$ and $\varepsilon_v(\infty)$ are the bulk band edges. The solid lines are fits to the TB data according to Eq. (2) and Table I.

splitting[64] Δ (plotted in Fig. 4) and on the temperature T. At low enough T only the $k=0$ valley is occupied by the electrons whereas the $k \simeq \pm 0.4\pi/\ell$ valleys fill as soon as $kT \gtrsim \Delta$. At a variance with Fig. 3, the holes in ⟨001⟩-oriented NWs are quite heavy ($m^* \simeq m_0$, increasing up to $\simeq 1.5m_0$ for $R=2$ nm). This enhancement of the hole mass far above bulk values is due to confinement-induced couplings between the

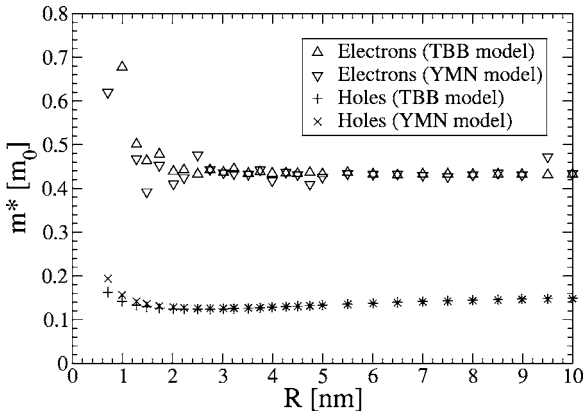

FIG. 3. Electron (lowest conduction subband) and hole (highest valence subband) effective masses in ⟨111⟩-oriented Si nanowires, in the $sp^3d^5s^*$ TBB model and in the sp^3 YMN model.

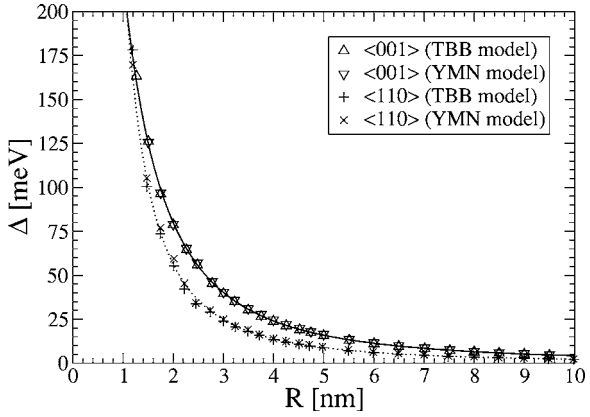

FIG. 4. The splitting Δ between the conduction band minimum at $k=0$ and the conduction band minimum at $k \simeq \pm 0.4\pi/\ell$ (⟨001⟩-oriented Si nanowires) or at $k \simeq \pm 0.8\pi/\ell$ (⟨110⟩-oriented Si nanowires). The solid and dashed lines are fits to the TB data according to Eq. (2) and Table I (TBB and YMN models).

$J=\frac{1}{2}$ and $J=\frac{3}{2}$ angular momentum components of the valence band wave functions and is reminiscent of the "camel back" structure to be discussed later in other materials. It may increase the sensitivity of the hole to disorder and fluctuations along the wire (making localization easier), which could show up in transport properties. A similar enhancement of the hole mass was found in small ⟨001⟩-oriented Si nanowires with square shapes.[44]

Last, in ⟨110⟩-oriented Si nanowires, two of the six bulk conduction band minima project onto $k=0$ (which is the actual conduction band minimum of the nanowire), while the four others fold onto $k \simeq \pm 0.8\pi/\ell$. The splitting between the two valleys is also plotted in Fig. 4. The electrons around $k=0$ exhibit a lighter ($m^* \simeq m_t^*$) mass along the nanowire than the electrons around $k=\pm 0.8\pi/\ell$ ($m^* \simeq 0.55m_0$), while the holes are rather light ($m^* \simeq 0.2m_0$). We emphasize that the optical matrix elements between the conduction and valence band edges of ⟨001⟩- and ⟨110⟩-oriented Si nanowires remain small in spite of their (pseudo-) direct band gap.

2. Ge nanowires

The subband structure of a ⟨111⟩-oriented Ge nanowire with radius $R=3.75$ nm is plotted in Fig. 5. Bulk germanium has three families of conduction band minima lying within a 250 meV energy range.[50] There are four equivalent conduction band minima with energy $E=0.74$ eV at L points (⟨111⟩ directions), one at Γ with energy $E=0.90$ eV, and six equivalent ones along ΓX directions (like in silicon) with energy $E \simeq 1$ eV. The energy isosurfaces around the L minima are again ellipsoids elongated along the ⟨111⟩ axes ($m_l^* = 1.59m_0$, $m_t^* = 0.08m_0$). In [111]-oriented Ge nanowires, the four L minima fold onto $k=\pi/\ell$, although nonequivalently. Indeed, electrons near the L[111] minimum are confined at a higher energy because they are light in the plane normal to the wire [and heavy ($m^* \sim m_l^*$) along the wire axis]. The other minima give rise to lighter ($m^* \simeq 0.25m_0$) bands with lower energy. This is clearly evidenced in Fig. 5. The Γ and Si-like minima are also visible in this figure. Note that the

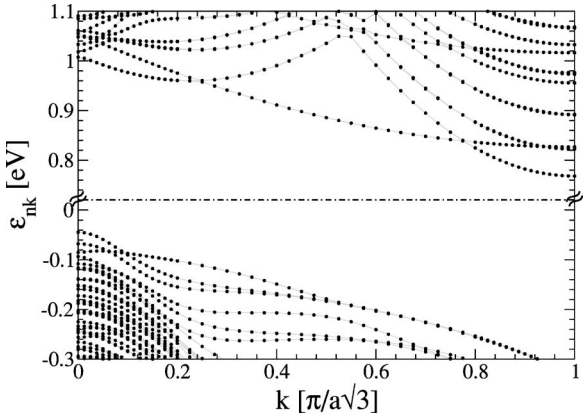

FIG. 5. Band structure of a ⟨111⟩-oriented Ge nanowire with radius $R=3.75$ nm ($sp^3d^5s^*$ TBB model).

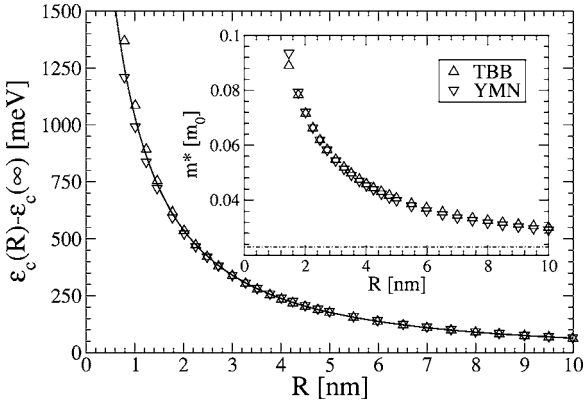

FIG. 6. Electron confinement energies and masses (inset) in ⟨111⟩-oriented InAs nanowires, in the $sp^3d^5s^*$ TBB model, and in the sp^3 YMN model. The solid line is a fit to $\varepsilon_c(R)-\varepsilon_c(\infty)$ according to Eq. (2) and Table I. The bulk conduction band effective mass in InAs is $m^*=0.023m_0$ (horizontal dash-dotted line).

actual conduction band minimum shifts to Γ in the smallest nanowires ($R<1$ nm). The conduction band minimum falls at $k=\pi/\ell$ in ⟨001⟩-oriented and at $k=0$ in ⟨110⟩-oriented Ge nanowires (with a clear bulk L character in both cases except possibly in the smallest nanowires that can mix contributions from all bulk minima). The electron mass along the wire is around $0.55m_0$ in ⟨001⟩-oriented and around m_t^* in ⟨110⟩-oriented nanowires.

The hole mass in small ⟨111⟩-oriented Ge nanowires ranges from 0.07 to $0.1m_0$, but rapidly increases above $R=5$ nm until the onset of a so-called "camel back" structure[62] beyond $R\simeq 15$ nm: the valence band maximum indeed shifts from $k=0$ to $k=\pm k_{cb}$. This shift as well as the height δE_{cb} of the camel back (the difference between the valence band edges at $k=\pm k_{cb}$ and $k=0$) are, however, pretty small: at most $k_{cb}\simeq\pm 0.004\pi/\ell$ and $\delta E_{cb}\simeq 30$ μeV. Camel back structures also appear above $R\simeq 10$ nm in ⟨001⟩-oriented Ge nanowires and in the whole investigated range in ⟨110⟩-oriented Ge nanowires. They contribute to an overall flattening of the highest valence subband around $k=0$, but are likely too small to be evidenced experimentally.

3. III-V nanowires

In all ⟨001⟩-, ⟨110⟩-, and ⟨111⟩-oriented III-V nanowires considered in this work the conduction band minimum falls at $k=0$ (or very near $k=0$ for GaP, where the X point folds[61,65]). The lowest conduction band of bulk InAs, GaAs, and InP is very dispersive but pretty isotropic around Γ; as a consequence $\varepsilon_c(R)$ increases very rapidly with decreasing R (see Fig. 6), the values of K_c, a_c, and b_c (Table II) being, however, nearly, the same as the nanowire orientation in these materials. The effective mass m^* of the electrons along the wire, though very close to the bulk value for large R, also dramatically increases with confinement because the bulk conduction band shows significant nonparabolicity at large energy.[41,66] It can be reproduced with a linear law $m^*[\varepsilon_c(R)]=m^*(0)\{1+\alpha[\varepsilon_c(R)-\varepsilon_c(\infty)]\}$, where $m^*(0)$ is the bulk effective mass and α has been fitted on the TB data in the range $\varepsilon_c(R)-\varepsilon_c(\infty)<0.5$ eV (see Table III). m^* typically shows nonlinear variations outside this range.

The valence band maximum falls at or very near to $k=0$. All III-V nanowires considered in this work (except ⟨001⟩ and ⟨111⟩-oriented GaP nanowires) indeed exhibit camel back structures[41] beyond some critical R_{cb} in the range $R=1-20$ nm. k_{cb} first rapidly increases then slowly decreases with $R>R_{cb}$. It is expected to vanish for large enough R (bulklike wires). The most proeminent camel back structure was found in a ⟨110⟩-oriented InAs nanowire with radius $R=1$ nm, where $k_{cb}\sim\pm 0.023\pi/\ell$ and $\delta E_{cb}=5.9$ meV (YMN model). R_{cb}, k_{cb}, and δE_{cb} are, however, very dependent on the TB model and details of the geometry.

C. Discussion: comparison with *ab initio* calculations and experiment

We finally compare the different TB models with each other as well as with experimental data and *ab initio* calculations. We mostly focus on silicon, because of its practical importance.

Figures 2–4 show that the three TB models used for Si are in good agreement, the scattering being significant only in the smallest nanowires ($R\lesssim 1.5$ nm). In this strong confinement regime (large $d\varepsilon_c/dR$ and $d\varepsilon_v/dR$), the quasiparticle energies indeed become very sensitive to the details of the TB model and geometry. Although semiempirical methods are also expected to break down somewhere in this range, we point out that the symmetry of the low-lying TB wave functions is usually correct and that the quasiparticle energies remain close to *ab initio* results[51] (see a later comparison). At

TABLE III. The bulk conduction band effective mass $m^*(0)$ (in units of m_0) and the parameter $\alpha=[1/m^*(0)]dm^*/d\varepsilon_c$ (eV^{-1}) for ⟨001⟩-, ⟨110⟩-, and ⟨111⟩-oriented InAs, GaAs, and InP nanowires.

Material	$m^*(0)$	$\alpha_{\langle 001\rangle}$	$\alpha_{\langle 110\rangle}$	$\alpha_{\langle 111\rangle}$
InAs	0.023	4.96	7.38	3.86
GaAs	0.066	2.30	5.41	1.18
InP	0.079	2.43	4.67	1.25

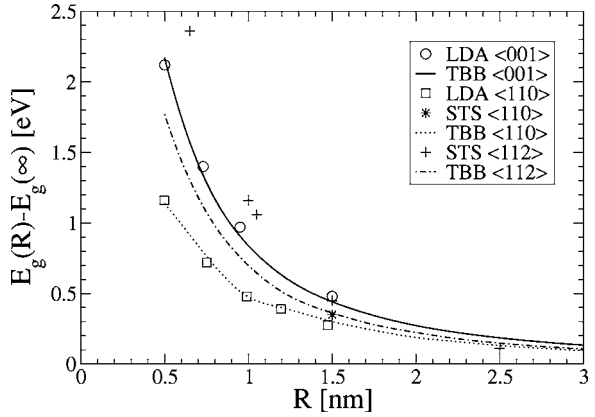

FIG. 7. Comparison between the TB band gap energies ($sp^3d^5s^*$ TBB model) and (i) the local density approximation (LDA) for ⟨001⟩-oriented (Ref. 32) and ⟨110⟩-oriented Si nanowires; (ii) the scanning tunneling spectroscopy (STS) data of Ref. 74 for ⟨110⟩- and ⟨112⟩-oriented Si nanowires. Spin-orbit coupling was not taken into account in the LDA and TB calculations, which only affects the band gap energies by a few meV.

variance, $\mathbf{k}\cdot\mathbf{p}$ and effective mass approximations[62] suffer from increasing overconfinement with decreasing R.[51,67] The agreement between the sp^3 and $sp^3d^5s^*$ models supports the transferability of both parametrizations from bulk materials to nanostructures. The TB models were found in similar agreement for the other materials considered in this work (see, e.g., Fig. 6), although slightly worse on the valence band than on the conduction band side. The holes, being typically heavier than the electrons, are indeed much more sensitive to the quality of the parametrization. In particular, the curvature of the valence band extrema should be carefully adjusted when fitting TB parameters.[51,57,58,68]

We now compare our TB results for Si with *ab initio* calculations and experimental data. The band gap energies of small ⟨001⟩-oriented Si nanowires computed by Delley and Steigmeier[32] in the LDA are plotted in Fig. 7. The total confinement energy $\Delta E_g(R)=E_g(R)-E_g(\infty)$ should be correctly given by the LDA although the latter is known to fail for the bulk band gap energy $E_g(\infty)\equiv E_{g,b}$. The LDA band gap energies of ⟨110⟩-oriented Si nanowires have been computed with SIESTA (Ref. 69) using norm-conserving pseudopotentials[70] in a nonlocal separable representation.[71] Numerical atomic orbitals were used as basis sets.[72,73] The geometry was optimized before the band structure calculation. The Si-Si bonds are shorter at the surface of the NWs than in the bulk (up to 4% in the smallest ones), while the inner bonds experience a minor enlargement. The relaxed nanowire structures thus remain quite close to their TB bulk-like counterparts. The LDA data are compared with the $sp^3d^5s^*$ TBB model, that is likely the most accurate in small nanostructures (because its basis set allows greater flexibility and because it provides the best description of the bulk bands). As a matter of fact, the TBB and LDA ΔE_g's are found in very good agreement down to the smallest ⟨001⟩- and ⟨110⟩-oriented nanowires. This further supports the validity of the TB approach to the electronic properties of silicon nanostructures. The large band gap differences between ⟨001⟩- and ⟨110⟩-oriented Si nanowires can be traced back to the anisotropy of the conduction and valence bands of bulk Si. Ma *et al.*[74] have measured the band gap energy of a few ⟨112⟩- and one ⟨110⟩-oriented Si nanowires using scanning tunneling spectroscopy (STS). Their data are plotted in Fig. 7. We have also computed the electronic structure of ⟨112⟩-oriented Si nanowires for comparison. The TBB model underestimates the opening of the STS band gap, especially in the smallest nanowires. As mentioned in Ref. 74, the diameter of the nanowires may be overestimated by the scanning tunneling microscope (STM). We emphasize, however, that the band gap energy of the nanowires may be substantially affected by the dielectric environment of the STS experiment, as discussed in the next section.

III. SELF-ENERGY CORRECTIONS

VLS nanowires usually exhibit sharp dielectric interfaces with their surroundings (e.g., vacuum or metallic electrodes). In most practical arrangements, these built-in dielectric mismatches are responsible for significant self-energy corrections to the TB band structure of the nanowire.[29] In this section, we discuss the underlying physics (Sec. III A) as well as the trends and magnitude of this effect on a simple, semianalytical effective mass model (Sec. III B). We then show that this simple model accurately reproduces the results of more detailed TB calculations when $\varepsilon_{in}>\varepsilon_{out}$ (Sec. III C). We finally compare our results with other calculations and experiments in Sec. III D.

A. Theory

When an additional electron is injected into a solid, it repells nearby valence electrons, thus dragging a so-called Coulomb hole around.[75] The work needed to form this (short-range) Coulomb hole makes a significant *self-energy* contribution to the band gap energy of semiconductors. We assume that the TB parameters account for this effect in bulk materials (because they reproduce the experimental band gap) as well as nanostructures. In bulk materials, the charge $q\sim 1/\varepsilon_{in}-1$ cast out from the Coulomb hole is repelled to "infinity." In nanostructures, however, this charge builds up around the surfaces or dielectric interfaces of the system. The interaction between the additional electron and these so-called "image" charges is responsible for additional self-energy corrections to the TB band structure. Of course, the above arguments also hold when removing one electron (adding a hole) to the system. We stress, however, that the self-energy correction to the TB band structure differs from the self-energy correction to the DFT band structure, that also accounts for the formation of the Coulomb hole in bulk materials.

Self-energy effects are usually addressed within a many-body framework,[75] using, for example, a Green's function approach such as the *GW* method.[46,76] This method treats the short-range (SR) and long-range (image charges) parts of the Coulomb hole on the same footing and takes retardation effects into account (the Coulomb hole does not follow the additional particle instantaneously). Unfortunately, *ab*

initio[77–79] *GW* codes cannot handle more than a few tens of atoms at the present time due to computational limitations. Zhao *et al.*,[45] for example, have computed the *GW* self-energy corrections in very small ($R \leqslant 0.8$ nm) free-standing ⟨110⟩-oriented Si nanowires. It has, however, been shown using semiempirical *GW* codes[28,80] that a simpler, semiclassical treatment of the image charge effects is accurate enough for low-lying quasiparticle states in nanocrystals[81–84] and thin films.[85]

In this semiclassical model—that can be derived from suitable approximations to the *GW* method—each material is characterized by its macroscopic dielectric constant[86] ε. An additional electron at point $\mathbf{r}$ produces a potential $V(\mathbf{r};\mathbf{r}')$ at point $\mathbf{r}'$ that is the solution of Poisson's equation:

$$\nabla_{\mathbf{r}'}[\varepsilon(\mathbf{r}')\nabla_{\mathbf{r}'}V(\mathbf{r};\mathbf{r}')] = 4\pi\delta(\mathbf{r}-\mathbf{r}'). \tag{3}$$

$V(\mathbf{r};\mathbf{r}') = V_b(\mathbf{r};\mathbf{r}') + V_s(\mathbf{r};\mathbf{r}')$ can be split in two parts, where $V_b(\mathbf{r};\mathbf{r}') = -1/[\varepsilon(\mathbf{r})|\mathbf{r}-\mathbf{r}'|]$ is the potential created by the additional electron plus the SR part of the Coulomb hole, and $V_s(\mathbf{r};\mathbf{r}')$ is the potential created by the image charges on the dielectric interfaces. This image charge distribution thus acts back on the electron with a potential $\Sigma(\mathbf{r}) = -\frac{1}{2}V_s(\mathbf{r};\mathbf{r})$ (the one-half factor following from the adiabatic building of the charge distribution[46]). The first-order correction to the TB conduction band energies ε_{cnk} reads

$$E_{cnk} = \varepsilon_{cnk} + \langle\varphi_{cnk}|\Sigma|\varphi_{cnk}\rangle, \tag{4}$$

φ_{cnk} being the corresponding wave functions. It can be shown that the potential acting on a hole is formally the opposite.[87] The first-order correction to the TB valence band energies ε_{vnk} is likewise:

$$E_{vnk} = \varepsilon_{vnk} - \langle\varphi_{vnk}|\Sigma|\varphi_{vnk}\rangle. \tag{5}$$

The trends in self-energy corrections most easily show up on a single dielectric interface. We therefore consider a nanowire of radius R with dielectric constant ε_{in} embedded in a medium with uniform dielectric constant ε_{out}. The image charge potential $\Sigma(r)$ then reads[47–49,88]

$$\Sigma(r) = \left[1 - \frac{\varepsilon_{out}}{\varepsilon_{in}}\right]\sum_{n\geqslant 0}\int_{-\infty}^{+\infty}\frac{dk}{2\pi}(2-\delta_{n,0}) \times \frac{K_n(|k|R)K_n'(|k|R)I_n^2(|k|r)}{D_n(\varepsilon_{in},\varepsilon_{out},|k|R)} \quad \text{if } r < R \tag{6a}$$

$$= \left[\frac{\varepsilon_{in}}{\varepsilon_{out}} - 1\right]\sum_{n\geqslant 0}\int_{-\infty}^{+\infty}\frac{dk}{2\pi}(2-\delta_{n,0}) \times \frac{I_n(|k|R)I_n'(|k|R)K_n^2(|k|r)}{D_n(\varepsilon_{in},\varepsilon_{out},|k|R)} \quad \text{if } r > R, \tag{6b}$$

where

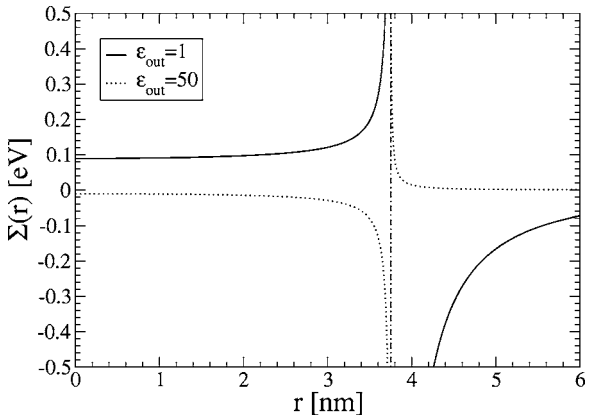

FIG. 8. The semiclassical self-energy potential $\Sigma(r)$ in a Si nanowire with radius $R=3.75$ nm ($\varepsilon_{in}=11.7$), for $\varepsilon_{out}=1$ and $\varepsilon_{out}=50$.

$$D_n(\varepsilon_{in},\varepsilon_{out},|k|R) = \varepsilon_{out}K_n'(|k|R)I_n(|k|R) - \varepsilon_{in}K_n(|k|R)I_n'(|k|R). \tag{7}$$

$I_n(x)$ and $K_n(x)$, are the modified Bessel functions of the first and second kind, respectively. $\Sigma(r)$ is plotted in Fig. 8 for a Si nanowire with radius $R=3.75$ nm ($\varepsilon_{in}=11.7$, $\varepsilon_{out}=1$, or 50). It is positive inside the nanowire when $\varepsilon_{in} > \varepsilon_{out}$ because the image charges repelled by an inner electron are negative, while it is negative outside the nanowire because an outer electron will polarize the nanowire and attract positive charges on its side. $\Sigma(r)$ diverges when $r \to R$. As shown later, this (spurious) divergence does not have much impact on the energy of the low-lying quasiparticles that have negligible amplitudes near the surface. The effects of the image charge potential will be twofold: (i) it will open the intrinsic band gap; (ii) it may bind a series of image surface states outside the nanowire. A proper account of the latter effect would, however, require a more elaborate *GW* calculation.[89] In the opposite case $\varepsilon_{out} > \varepsilon_{in}$, $\Sigma(r)$ is negative inside the wire and positive outside, digging a well on the inner side of the surface. It closes the band gap and pushes low-lying quasiparticles states outward. We will come back to this point later.

B. Semianalytical effective mass model

We now introduce a semianalytical model for the self-energy corrections $\langle\Sigma_c\rangle(R) = +\langle\varphi_c|\Sigma|\varphi_c\rangle$ and $\langle\Sigma_v\rangle(R) = -\langle\varphi_v|\Sigma|\varphi_v\rangle$, where φ_v and φ_c are, respectively, the highest valence band and lowest conduction band wave functions. This semianalytical model, intended for practical applications, will be checked against TB calculations in the next paragraph.

We can hopefully get a good estimate[83,84] of $\langle\Sigma_c\rangle(R)$ and $\langle\Sigma_v\rangle(R)$ using an effective mass ansatz for φ_c and φ_v in Eqs. (4) and (5). Assuming hard wall boundary conditions at $r=R$, the envelope functions read in the single band, isotropic effective mass approximation:[62]

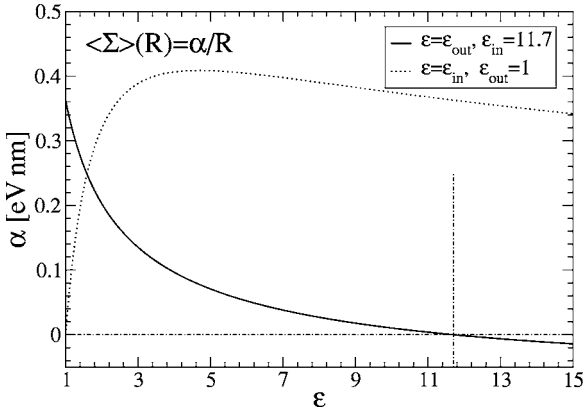

FIG. 9. The prefactor α of the self-energy correction $\langle\Sigma\rangle(R)=\alpha/R$ [Eq. (10)] as a function of ε_{out} for $\varepsilon_{\text{in}}=11.7$, and as a function of ε_{in} for $\varepsilon_{\text{out}}=1$. The vertical line is the dielectric constant of bulk Si, $\varepsilon_{\text{in}}=11.7$.

$$\varphi(r)=\frac{1}{\sqrt{2\pi KR}}J_0\left(\alpha_0\frac{r}{R}\right), \tag{8}$$

where $J_0(x)$ is the zeroth-order Bessel function of the first kind, with first zero $\alpha_0=0.240(4)$, and $K=\int_0^1 xJ_0^2(\alpha_0 x)dx=0.134(7)$. Hence,

$$\langle\Sigma\rangle(R)=\langle\varphi|\Sigma|\varphi\rangle=\frac{1}{KR^2}\int_0^R r\Sigma(r)J_0^2\left(\alpha_0\frac{r}{R}\right)dr. \tag{9}$$

The change of variables $r=Rx$ and $k=y/R$ takes out any dependence on R from the integrals in Eqs. (6a) and (9). This yields

$$\langle\Sigma\rangle(R)=\frac{1}{\varepsilon_{\text{in}}R}\frac{\varepsilon_{\text{in}}-\varepsilon_{\text{out}}}{\varepsilon_{\text{in}}+\varepsilon_{\text{out}}}F\left(\frac{\varepsilon_{\text{in}}}{\varepsilon_{\text{out}}}\right), \tag{10}$$

where $F(x)$ is a positive function whose expression follows from the former equations. We have numerically calculated $F(x)$ in the range $10^{-2}-10^2$ using 128 Bessel functions and integrating over k up to $k_{\max}=512/R$, the estimated accuracy on F being around 1 meV nm. We expect $\langle\Sigma\rangle(R)$ to depend on ε_{out} only in the limit $\varepsilon_{\text{in}}/\varepsilon_{\text{out}}\to\infty$, i.e., $F(x)\sim\alpha x$ when $x\to\infty$. For practical applications, we thus fit $F(x)$ with the following Padé approximant:

$$F(x)=\frac{0.0949x^3+17.395x^2+175.739x+200.674}{x^2+50.841x+219.091}\ [\text{eV nm}]. \tag{11}$$

$\langle\Sigma\rangle(R)\equiv\alpha/R$ has a clear $1/R$ behavior: *self-energy corrections to the quasiparticle band gap energy decrease slower with R than quantum confinement*, a conclusion already drawn in nanocrystals and films.[28,83] The prefactor α is plotted as a function of ε_{out} in Fig. 9 for a Si nanowire ($\varepsilon_{\text{in}}=11.7$). The self-energy corrections decrease very rapidly with increasing ε_{out}, being much larger in the limit $\varepsilon_{\text{in}}\gg\varepsilon_{\text{out}}$ than in the limit $\varepsilon_{\text{in}}\ll\varepsilon_{\text{out}}$. They are actually the most important contribution to the opening of the band gap in free-standing ($\varepsilon_{\text{out}}=1$) Si nanowires with radius $R\gtrsim 1$ nm. For example, Eq. (10) yields $\langle\Sigma\rangle(R)=97$ meV for $R=3.75$ nm, while Eq. (2) and Table I yield $\varepsilon_c(R)-\varepsilon_c(\infty)=51$ meV and $\varepsilon_v(R)-\varepsilon_v(\infty)=25$ meV.

α is also plotted as a function of ε_{in} in Fig. 9, for a nanowire in vacuum. It slightly increases down to $\varepsilon_{\text{in}}\simeq 5$ then abruptly falls to zero. The self-energy corrections are indeed driven by two competing mechanisms. On one hand, the total image charge grows up with increasing ε_{in}, which rises $\langle\Sigma\rangle(R)$. On the other hand, the image charge distribution also spreads farther and farther along the nanowire, which decreases $\langle\Sigma\rangle(R)$. The latter mechanism oversteps the former as soon as ε_{in} is large enough compared to ε_{out}.

Finally, we would like to recall that the one-particle band gap energy is relevant for charge transport but not for optical spectroscopy experiments. In the latter case, the enhancement of the exciton binding energy with decreasing R cancels the self-energy corrections,[47,83] the exciton being a neutral excitation (no net charging of the nanowire).

C. Tight-binding calculations

The results of the above effective mass model have been checked against TB calculations, either using Eqs. (4) and (5), or consistently including the potential $\pm\Sigma(r)$ in the Hamiltonian before calculating electron (+) or hole (−) states. The latter choice ("full" calculation) yields nonorthogonal electron and hole wave functions (because we use different potentials for the two kinds of quasiparticles), which would indeed be the outcome[46,76] of any more elaborate quasiparticle theory such as *GW*. The actual overlap between the lowest conduction band and highest valence band wave functions (at $k=0$) in ⟨111⟩-oriented Si nanowires was found lower than 10^{-3}. The TB treatment of the divergence of the image charge potential $\Sigma(r)$ is detailed in Appendix B

We now discuss the cases $\varepsilon_{\text{out}}=1$ and $\varepsilon_{\text{out}}=50$ in ⟨111⟩-oriented Si nanowires.

1. Case $\varepsilon_{\text{out}}=1$

The conduction and valence band edge energies $E_c(R)$ and $E_v(R)$ as well as the self-energy corrections $\langle\Sigma_c\rangle(R)$ and $\langle\Sigma_v\rangle(R)$ are plotted in Fig. 10 (TBB model). The results from the full calculation [including $\pm\Sigma(r)$ in the Hamiltonian] are compared with first-order perturbation theory [Eqs. (4) and (5) using TB wave functions and energies] and with the semianalytical model [Eqs. (2), (10), and (11), and Table I]. The three sets of curves are in very good agreement with each other. $\langle\Sigma\rangle(R)$ can indeed be split in two contributions, the first one being a rigid shift of value $\Sigma(r=0)$ [the baseline of $\Sigma(r)$, see Fig. 8], the second one, $\langle\delta\Sigma\rangle(R)$, being the remainder. The image charge potential $\Sigma(r)$ is, however, quite flat in the nanowire, only increasing near the surface where the low-lying quasiparticle wave functions have negligible amplitudes. $\langle\delta\Sigma\rangle(R)$, the unique wave-function-dependent contribution, therefore hardly represents $\simeq 10\%$ of the total self-energy correction. Moreover, the full calculation only leads to a moderate contraction of the low-lying quasiparticle wave functions ($\langle\varphi_v|r^2|\varphi_v\rangle^{1/2}$ decreases by up to 15% in the

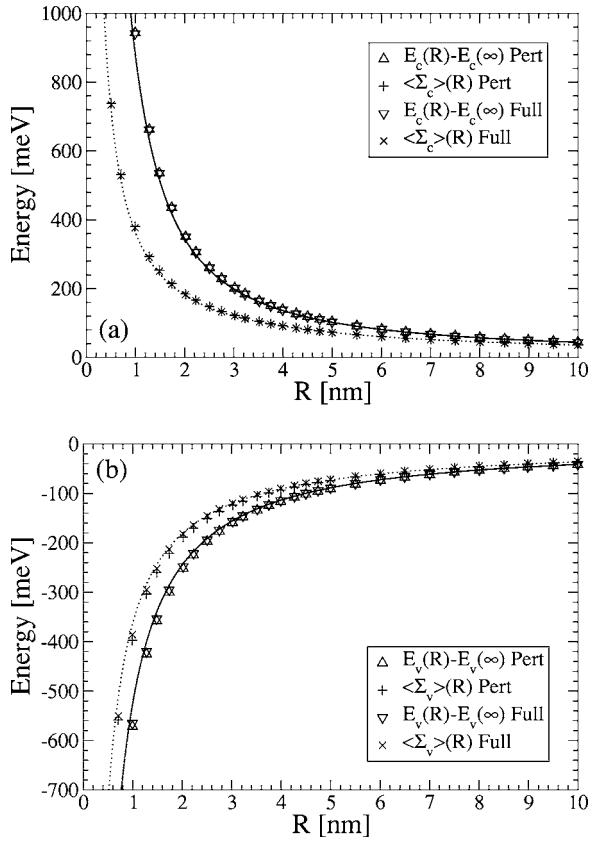

FIG. 10. The conduction (a) and valence (b) band edge energies $E_c(R)$ and $E_v(R)$, as well as the self-energy corrections $\langle\Sigma_c\rangle(R)$ and $\langle\Sigma_v\rangle(R)$ in $\langle 111\rangle$-oriented Si nanowires ($\varepsilon_{out}=1$, $sp^3d^5s^*$ TBB model). The results from a "full' calculation [including $\pm\Sigma(r)$ in the Hamiltonian] and from first-order perturbation theory (pert) are shown. The solid and dashed lines are the results from the fits of Sec. II B [Eq. (2) and Table I], and from the semianalytical model [Eqs. (10) and (11)].

largest nanowires). Equations (10) and (11) can thus be safely used in the case $\varepsilon_{in} \gg \varepsilon_{out}$.

The image charge potential shifts the conduction and valence bands in a nearly rigid way, as shown in Fig. 11, where the self-energy correction is plotted as a function of the bare subband energy $\varepsilon_{nk}=\langle\varphi_{nk}|H_{TB}|\varphi_{nk}\rangle$, H_{TB} being the TB Hamiltonian without image charge potential. In particular, the conduction and valence band effective masses are little affected by $\Sigma(r)$. This was to be expected from the above arguments. The self-energy correction nonetheless tends to increase with the subband index n, as the wave functions spread farther and farther from the nanowire axis. The scattering is stronger on the valence than on the conduction band side, the hole wave functions rapidly showing rich and complex features.

2. Case $\varepsilon_{out}=50$

The value $\varepsilon_{out}=50$ has been chosen as a representative case of the limit $\varepsilon_{out} \gg \varepsilon_{in}$. The comparison between the TB results and the semianalytical model is not as favorable as in the former case, especially on the valence band side (see Fig. 12). Indeed, the self-energy potential $\Sigma(r)$, plotted in Fig. 8,

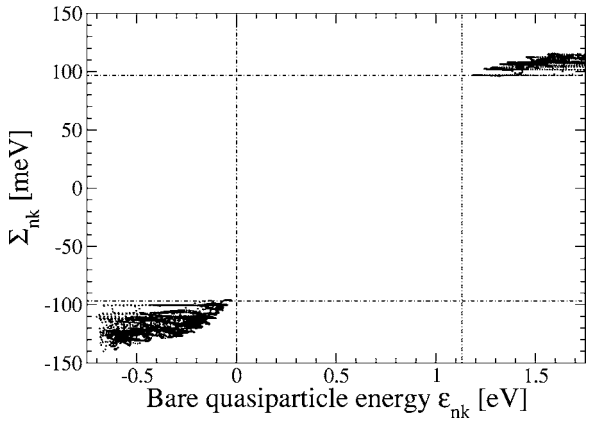

FIG. 11. The self-energy correction $\Sigma_{nk}=\pm\langle\varphi_{nk}|\Sigma|\varphi_{nk}\rangle$ as a function of the bare quasiparticle energy $\varepsilon_{nk}=\langle\varphi_{nk}|H_{TB}|\varphi_{nk}\rangle$ for a $\langle 111\rangle$-oriented Si nanowire with radius $R=3.75$ nm ($\varepsilon_{out}=1$, full calculation, TBB model). The 48 lowest conduction subbands and the 48 highest valence subbands at 57 k points in $[0,\pi/\ell[$ are represented. The vertical lines are the bulk conduction and valence band edges, while the horizontal lines are the self-energy corrections computed from Eqs. (10) and (11).

is nearly zero inside the nanowire but rapidly decreases close to the surface. We therefore expect $\langle\delta\Sigma\rangle(R)\sim\Sigma(r=0)$, the overall self-energy correction being however much smaller than in the limit $\varepsilon_{in} \gg \varepsilon_{out}$. The TB valence band wave function slightly differs from the single band effective mass approximation, which explains the increasing discrepancy between first-order TB and the semianalytical model for $\langle\Sigma_v\rangle(R)$. As a matter of fact, such $\lesssim$10 meV discrepancies also exist when $\varepsilon_{in} \gg \varepsilon_{out}$, but are negligible on the scale of Fig. 10(b). The image charge potential digs a well close to the surface of the nanowire that tends to attract the electrons and holes. The latter are much more sensitive to $\Sigma(r)$: $\langle\varphi_v|r^2|\varphi_v\rangle^{1/2}$indeed increases by up to 30% in the largest nanowires, while the full $\langle\Sigma_v\rangle(R)$ nearly doubles with respect to first-order perturbation theory. This is somewhat compensated by an increase of the kinetic energy of the hole, the differences between the full and first-order $E_v(R)$ being smaller than the differences between the full and first-order $\langle\Sigma_v\rangle(R)$. The self-energy correction overcompensates quantum confinement $[E_v(R)>0]$ above $R=6$ nm. The image charge potential does not, however, bind the holes in the range $R=1-10$ nm, but the highest valence band wave function might be bound in other materials, orientations, or diameter ranges. We stress that these results for $\varepsilon_{out} \gg \varepsilon_{in}$, though certainly showing the correct trends, are of limited quantitative accuracy. Indeed, a quantitative description of the self-energy effects close to the surface would require a far more elaborate, complete *GW* calculation free of singularities. This is, unfortunately, far beyond present computational capabilities above $R \simeq 1$ nm.

D. Comparison with *ab initio* calculations and experiment

We now compare our total self-energy correction $\Sigma_g(R)=\langle\Sigma_c\rangle(R)-\langle\Sigma_v\rangle(R)$ with the *ab initio GW* results of Zhao *et*

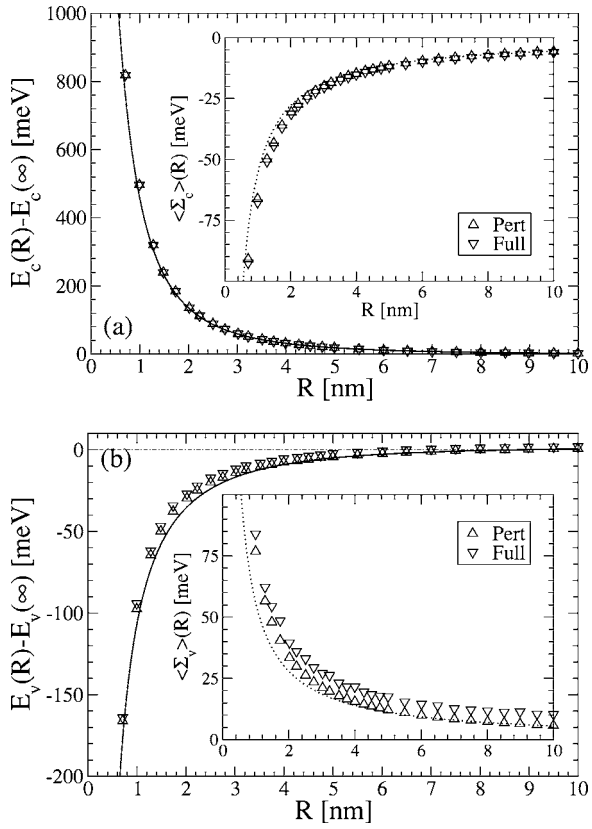

FIG. 12. The conduction (a) and valence (b) band edge energies $E_c(R)$ and $E_v(R)$, as well as the self-energy corrections $\langle\Sigma_c\rangle(R)$ and $\langle\Sigma_v\rangle(R)$ (insets) in $\langle 111\rangle$-oriented Si nanowires ($\varepsilon_{out}=50$, $sp^3d^5s^*$ TBB model). The results from a "full' calculation [including $\pm\Sigma(r)$ in the Hamiltonian] and from first-order perturbation theory (pert) are shown. The solid and dashed lines are the results from the fits of Sec. II B [Eq. (2) and Table I], and from the semianalytical model [Eqs. (10) and (11)].

al.[45] on $\langle 110\rangle$-oriented Si nanowires. The formation of the Coulomb hole in bulk materials is also accounted for by the *ab initio GW* self-energy correction, that is thus nonzero [$\Sigma_g(\infty)=0.5$ eV] in that limit. It is however included in the TB parameters (that yield the experimental bulk band gap), so that the TB self-energy correction (due to the image charges) is zero in the bulk. We shall therefore compare the respective $\Delta\Sigma_g(R)=\Sigma_g(R)-\Sigma_g(\infty)$. Zhao *et al.* give[90] $\Delta\Sigma_g(5.2\ \text{Å})=1.12$ eV and $\Delta\Sigma_g(8.3\ \text{Å})=0.79$ eV, while the full TB calculation (on $\langle 110\rangle$ NWs) yields $\Delta\Sigma_g(5.3\ \text{Å})=1.48$ eV and $\Delta\Sigma_g(8.5\ \text{Å})=0.91$ eV. The agreement is satisfactory, given the very small size of the nanowires. The applicability of our model, based on classical electrostatics with the bulk dielectric constant, is questionable in the nanometer range. By the way, we stress that the interaction between neighboring wires must be carefully cut off in supercell calculations. Figure 9 indeed suggests that the self-energy corrections rapidly decrease as soon as the outer medium can screen Coulomb interactions. Image charge effects are in this respect much longer ranged than quantum confinement.

We would finally like to comment again on the STS data[74] of Fig. 7. In this experiment, the nanowires are laid down on a (semi) metallic highly ordered pyrolitic graphite (HOPG) substrate and the current is collected by a nearby metallic STM tip. We therefore expect the STS setup to act as a rather high ε_{out} medium, and thus small self-energy corrections. This seems hardly compatible with Fig. 7, that suggests instead $\varepsilon_{out}=1-2$ in the smallest nanowires. As a matter of fact, the difference between the STS and TB data decreases much faster than $1/R$. Of course, a STS setup is a highly inhomogeneous environment that may not be reproduced so easily with a single "effective" dielectric constant. We would have expected however the self-energy corrections to decrease slower than $1/R$ as screening becomes less efficient with increasing tip-substrate distance. There are, though, uncertainties on the experimental diameters and band gap energies, as well as on the physics of the tip-nanowire interaction. In particular, the image charge potential might dig a well under the STM tip (or under a metallic ring surrounding the nanowire, for example), that could trap a few electrons or holes, increasing confinement energy.

IV. CONCLUSION

We have computed the subband structure and quasiparticle band gap energy of several group IV and III-V $\langle 001\rangle$-, $\langle 110\rangle$, and $\langle 111\rangle$-oriented nanowires using various sp^3 and $sp^3d^5s^*$ tight-binding models. These models are in very good agreement one with each other, showing the robustness of our results. The results obtained for Si nanowires were also successfully checked against LDA calculations. We have provided analytical fits to the conduction and valence band edge energies for practical use. We have also shown that the self-energy corrections, which arise from the dielectric mismatch between the nanowires and their environment, are usually far from negligible when $\varepsilon_{in}\gg\varepsilon_{out}$ and decrease like $1/R$, slower than the quantum confinement (R being the radius of the nanowire). Many important features of the transport through nanowires (such as current onsets) depend on both these quantum confinement and self-energy effects.

ACKNOWLEDGMENTS

N. H. Quang thanks the CEA and the Laboratory of Atomistic Simulation (L_Sim) for their hospitality and for a grant. He also thanks the VAST and the Vietnamese National basic research program for partially funding his visit to the CEA. This work was supported by the French "Action Concertée Incitative" (ACI) "TransNanofils." The authors are indebted to Région Rhônes-Alpes and CNRS for partial funding and to the supercomputing centers CDCSP (University of Lyon) and IDRIS (Orsay, CNRS).

APPENDIX A: APPLICATION OF THE JACOBI-DAVIDSON ALGORITHM TO TIGHT-BINDING PROBLEMS

The eigenstates of TB Hamiltonian H_{TB} were computed using a Jacobi-Davidson algorithm (JDA) with harmonic Ritz values as described in Refs. 59 and 60. This algorithm proved to be much more efficient than the folded spectrum

method[91] (FSM) used in previous studies,[51–53] because preconditioning of the JDA is much easier than preconditioning of the FSM (see below) for TB problems. Time-reversal and/or spatial symmetries are used to speed up the search for the eigenstates. Indeed, symmetry operations are applied once an eigenvalue has converged to find all degenerate eigenvectors.

Each Jacobi-Davidson iteration involves an *approximate* solution of a linear system of the form[59]

$$[I-|\psi\rangle\langle\psi|][H_{\mathrm{TB}}-\theta I][I-|\psi\rangle\langle\psi|]|u\rangle=-|r\rangle, \qquad \text{(A1)}$$

where $(\theta,|\psi\rangle)$ is the best possible approximation to an eigenpair of H_{TB} and $|r\rangle=H_{\mathrm{TB}}|\psi\rangle-\theta|\psi\rangle$ is the residual. This linear system is solved with a few generalized minimal residual (GMRES) iterations.[92] The accuracy of the solution $|u\rangle$ for a given number of GMRES iterations can be improved with a preconditioner,[92] i.e., an approximate inverse of $H_{\mathrm{TB}}-\theta I$. Here we used a so-called "bond orbital model[54]" as a preconditioner. It is based on the idea that conduction band wave functions are mostly antibonding combinations of atomic orbitals while valence band wavefunctions are mostly bonding combinations. This bond orbital model is built as follows for sp^3 TB models.

(1) For each atom, compute the sp^3 hybrids pointing toward the four nearest neighbors.

(2) For each pair of first nearest neighbors, compute the bonding and antibonding combinations of the two sp^3 hybrids aligned with the bond. Let $\mathcal{B}$ be the basis of these bonding and antibonding combinations of sp^3 hybrids, that are centered on bonds rather than atoms.

(3) Give each bonding combination an energy $E=E_0-\beta$, and each antibonding combination an energy $E=E_0+\beta$. The resulting bond orbital model Hamiltonian $\tilde{H}_{\mathrm{TB}}-\theta I$ is diagonal in $\mathcal{B}$ and thus easily invertible in this basis.

(4) Transform $(\tilde{H}_{\mathrm{TB}}-\theta I)^{-1}$ back to the original sp^3 basis, which yields an effective first nearest neighbor model for the preconditioner.

The sp^3 hybrids are replaced by the s orbital for hydrogen atoms. Typical values for E_0 and β are $E_0=0$ eV and $\beta\simeq 5$ eV, irrespective of the material. In practice, we fix θ in the midgap range and compute the preconditioner once for all.

The bond orbital model is built in the same way for the sp^3 orbitals of $sp^3d^5s^*$ TB models. The d and s^* orbitals are preconditioned "on-site," just by setting $[(\tilde{H}_{\mathrm{TB}}-\theta I)^{-1}]_{ii}=1/(E_{ds^*}-\theta)$ for these orbitals, where $E_{ds^*}\simeq 15$ eV. This preconditioner, though crude, precisely discriminates between the bonding and antibonding combinations of atomic orbitals as needed for the computation of valence or conduction band states, allowing fast convergence of the JDA.

APPENDIX B: TIGHT-BINDING TREATMENT OF THE DIVERGENCE OF $\Sigma(\mathbf{r})$

It is customary[85] to handle the divergence of the semiclassical image charge potential in atomistic calculations with a shift δR of the dielectric interface so that all atoms fall within $R+\delta R$. This, however, significantly affects the self-energy profile far inside the structure. Here we adopt another strategy: we first assume that the image charge potential at $\mathbf{r}$ is created by a charge distribution (the electron plus the short-range part of the Coulomb hole) with a finite (Gaussian) extension along z and θ. We thus now define $\Sigma(r)$ as (r, θ, and z being the cylindrical coordinates):

$$\Sigma(r)=\frac{1}{4\pi\sigma_\theta\sigma_z}\int d\theta\int dz\, V_s(r,\theta,z;r,0,0)e^{-\theta^2/(2\sigma_\theta^2)}e^{-z^2/(2\sigma_z^2)} \qquad \text{(B1)}$$

which amounts to multiplying the integrand in Eqs. (6) by $e^{-\sigma_\theta n^2}\times e^{-\sigma_z k^2}$. We use $\sigma_z=R\sigma_\theta=1$ Å, a reasonnable estimate for the Coulomb hole size.[79] This effectively replaces the divergence at $r=R$ by a discontinuity, but leaves $\Sigma(r)$ unchanged a few σ_z from the interface. Second, we extrapolate Eq. (6a) for $r>R$, which is straightforward once the divergence has been removed. We do so because the TB basis sets are not designed to tackle the image surface states that may bind outside the nanowire, whose proper description would anyway require a complete *GW* calculation.

The tight-binding results are almost insensitive to σ_z (in the range 0.25–1 Å) when $\varepsilon_{\mathrm{in}}>\varepsilon_{\mathrm{out}}$. The valence band self-energy corrections are, however, much more sensitive to σ_z in the limit $\varepsilon_{\mathrm{in}}\ll\varepsilon_{\mathrm{out}}$, the hole wave functions slightly spreading outwards as discussed in Sec. III C. A detailed description of the image charge effects near the surface would again require a complete *GW* calculation, far beyond present computational capabilities.

*Electronic address: yniquet@cea.fr

[1] R. S. Wagner and W. C. Ellis, Appl. Phys. Lett. **4**, 89 (1964).
[2] A. I. Persson, M. W. Larsson, S. Stenström, B. J. Ohlsson, L. Samuelson, and L. R. Wallenberg, Nat. Mater. **3**, 677 (2004).
[3] K. A. Dick, K. Deppert, T. Mårtensson, B. Mandl, L. Samuelson, and W. Seifert, Nano Lett. **5**, 761 (2005).
[4] X. Duan and C. M. Lieber, Adv. Mater. (Weinheim, Ger.) **12**, 298 (2000).
[5] A. M. Morales and C. M. Lieber, Science **279**, 208 (1998).
[6] Y. Wu, Y. Cui, L. Huynh, C. J. Barrelet, D. C. Bell, and C. M. Lieber, Nano Lett. **4**, 433 (2004).
[7] X. Duan, J. Wang, and C. M. Lieber, Appl. Phys. Lett. **76**, 1116 (2000).
[8] X. Duan, Y. Huang, Y. Cui, J. Wang, and C. M. Lieber, Nature (London) **409**, 66 (2001).
[9] U. Krishnamachari, M. Borgström, B. J. Ohlsson, N. Panev, L. Samuelson, W. Seifert, M. W. Larsson, and L. R. Wallenberg, Appl. Phys. Lett. **85**, 2077 (2004).
[10] M. S. Gudiksen and C. M. Lieber, J. Am. Chem. Soc. **122**, 8801 (2000).

[11] M. S. Gudiksen, L. J. Lauhon, J. Wang, D. C. Smith, and C. M. Lieber, Nature (London) **415**, 617 (2002).
[12] M. T. Björk, B. J. Ohlsson, T. Sass, A. I. Persson, C. Thelander, M. H. Magnusson, K. Deppert, L. R. Wallenberg, and L. Samuelson, Appl. Phys. Lett. **80**, 1058 (2002).
[13] Y. Wu, R. Fan and P. Yang, Nano Lett. **2**, 83 (2002).
[14] L. J. Lauhon, M. S. Gudiksen, D. Wang, and C. M. Lieber, Nature (London) **420**, 57 (2002).
[15] C. Thelander, M. T. Björk, M. W. Larsson, A. E. Hansen, L. R. Wallenberg, and L. Samuelson, Solid State Commun. **131**, 573 (2004).
[16] J. Wang, M. S. Gudiksen, X. Duan, Y. Cui, and C. M. Lieber, Science **293**, 1455 (2001).
[17] S. De Franceschi, J. A. Van Dam, E. P. A. M. Bakkers, L. F. Feiner, L. Gurevich, and L. P. Kouwenhoven, Appl. Phys. Lett. **83**, 344 (2003).
[18] Z. Zhong, Y. Fang, W. Lu, and C. M. Lieber, Nano Lett. **5**, 1143 (2005).
[19] M. T. Björk, B. J. Ohlsson, C. Thelander, A. I. Persson, K. Deppert, L. R. Wallenberg, and L. Samuelson, Appl. Phys. Lett. **81**, 4458 (2002).
[20] Y. Cui and C. M. Lieber, Science **291**, 851 (2001).
[21] X. Duan, C. Niu, V. Sahi, J. Chen, J. W. Parce, S. Empedocles, and J. L. Goldman, Nature (London) **425**, 274 (2003).
[22] Y. Cui, Z. Zhong, D. Wang, W. U. Wang, and C. M. Lieber, Nano Lett. **3**, 149 (2003).
[23] D. Wang, Q. Wang, A. Javey, R. Tu, H. Dai, H. Kim, P. C. McIntyre, T. Krishnamohan, and K. C. Saraswat, Appl. Phys. Lett. **83**, 2432 (2003).
[24] Y. Cui, Q. Wei, H. Park, and C. M. Lieber, Science **293**, 1289 (2001).
[25] Y. Huang, X. Duan, Q. Wei, and C. M. Lieber, Science **291**, 630 (2001).
[26] T. Mårtensson, P. Carlberg, M. Borgström, L. Montelius, W. Seifert, and L. Samuelson, Nano Lett. **4**, 699 (2004).
[27] S. Jin, D. Whang, M. C. McAlpine, R. S. Friedman, Y. Wu, and C. M. Lieber, Nano Lett. **4**, 915 (2004).
[28] C. Delerue, G. Allan, and M. Lannoo, Phys. Rev. Lett. **90**, 076803 (2003).
[29] C. Delerue and M. Lannoo, *Nanostructures: Theory and Modelling* (Springer, New York, 2004).
[30] A. J. Read, R. J. Needs, K. J. Nash, L. T. Canham, P. D. J. Calcott, and A. Qteish, Phys. Rev. Lett. **69**, 1232 (1992).
[31] F. Buda, J. Kohanoff, and M. Parrinello, Phys. Rev. Lett. **69**, 1272 (1992).
[32] B. Delley and E. F. Steigmeier, Appl. Phys. Lett. **67**, 2370 (1995).
[33] R. Rurali and N. Lorente, Phys. Rev. Lett. **94**, 026805 (2005).
[34] R. N. Musin and X.-Q. Wang, Phys. Rev. B **71**, 155318(R) (2005).
[35] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[36] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[37] U. Bockelmann and G. Bastard, Phys. Rev. B **45**, 1688 (1992).
[38] C. Y. Yeh, S. B. Zhang, and A. Zunger, Phys. Rev. B **50**, 14405 (1994).
[39] M. Califano and A. Zunger, Phys. Rev. B **70**, 165317 (2004).
[40] M. P. Persson and H. Q. Xu, Appl. Phys. Lett. **81**, 1309 (2002).
[41] M. P. Persson and H. Q. Xu, Nano Lett. **4**, 2409 (2004).
[42] M. P. Persson and H. Q. Xu, Phys. Rev. B **70**, 161310(R) (2004).
[43] U. Landman, R. N. Barnett, A. G. Scherbakov, and Ph. Avouris, Phys. Rev. Lett. **85**, 1958 (2000).
[44] Y. Zheng, C. Rivas, R. Lake, K. Alam, T. B. Boykin, and G. Klimeck, IEEE Trans. Electron Devices **52**, 1097 (2005).
[45] X. Zhao, C. M. Wei, L. Yang, and M. Y. Chou, Phys. Rev. Lett. **92**, 236805 (2004); F. Bruneval, S. Botti, and L. Reining, *ibid.* **94**, 219701 (2005).
[46] L. Hedin and S. Lundqvist, Solid State Phys. **23**, 1 (1969).
[47] E. A. Muljarov, E. A. Zhukov, V. S. Dneprovskii, and Y. Masumoto, Phys. Rev. B **62**, 7420 (2000).
[48] A. Shik, J. Appl. Phys. **74**, 2951 (1993).
[49] J.-N. Chazalviel, F. Ozanam, and V. M. Dubin, J. Phys. I **4**, 1325 (1994).
[50] *Physics of Group IV Elements and III-V Compounds*, edited by O. Madelung, M. Schulz, and H. Weiss, Landolt-Börnstein, New Series, Group III, Vol. 17, Part A (Springer-Verlag, New York, 1982).
[51] Y. M. Niquet, C. Delerue, G. Allan, and M. Lannoo, Phys. Rev. B **62**, 5109 (2000).
[52] Y. M. Niquet, G. Allan, C. Delerue, and M. Lannoo, Appl. Phys. Lett. **77**, 1182 (2000).
[53] Y. M. Niquet, C. Delerue, G. Allan, and M. Lannoo, Phys. Rev. B **65**, 165334 (2002).
[54] W. A. Harrison, *Elementary Electronic Structure* (World Scientific, London, 1999).
[55] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).
[56] J. M. Jancu, R. Scholz, F. Beltram, and F. Bassani, Phys. Rev. B **57**, 6493 (1998).
[57] T. B. Boykin, G. Klimeck, and F. Oyafuso, Phys. Rev. B **69**, 115201 (2004).
[58] T. B. Boykin, G. Klimeck, R. C. Bowen, and F. Oyafuso, Phys. Rev. B **66**, 125207 (2002).
[59] G. L. G. Sleijpen and H. A. Van der Vorst, SIAM J. Matrix Anal. Appl. **17**, 401 (1996); SIAM Rev. **42**, 267 (2000).
[60] G. Sleijpen and H. Van der Vorst, in *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, edited by Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. Van der Vorst (SIAM, Philadelphia, 2000).
[61] The bulk band gap energies $E_{g,b}$ are 1.17 eV (Si, indirect band gap near X), 0.74 eV (Ge, indirect band gap at L), 0.42 eV (InAs), 1.52 eV (GaAs), 1.42 eV (InP), and 2.35 eV (GaP, indirect band gap near X).
[62] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Les Editions de Physique, Les Ulis, 1988).
[63] The minima are located at $0.813\Gamma X$ in the $sp^3d^5s^*$ TBB model. The experimental (Ref. 50) position of the minima is closer to $0.85\Gamma X$.
[64] Note that intervalley couplings again split the fourfold degenerate conduction band minima at $k=0$ into one twofold degenerate and two nondegenerate subbands (excluding spin degeneracy), this splitting being however much smaller than Δ.
[65] There can also be a very small ($|k_{cb}| \lesssim 10^{-3}\pi/\ell$) camel back structure in the lowest conduction band of ⟨110⟩-oriented III-V nanowires.
[66] R. Chen and K. K. Bajaj, Phys. Rev. B **50**, 1949 (1994).
[67] H. Fu, L. W. Wang, and A. Zunger, Phys. Rev. B **57**, 9971 (1998).
[68] T. B. Boykin, Phys. Rev. B **56**, 9613 (1997).
[69] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, J. Phys.: Condens. Matter **14**, 2745 (2002).

[70] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[71] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[72] E. Anglada, J. M. Soler, J. Junquera, and E. Artacho, Phys. Rev. B **66**, 205101 (2002).
[73] In particular, the nanowire structures were relaxed in a variationally optimized (Ref. 72) single-ζ polarized basis set, with orbital radii in the 7–8 bohrs range. The band structures were then calculated in a double-ζ polarized basis set with 5–7 bohrs orbital radii. The first Brillouinzone was sampled with four k-points along the nanowire axis.
[74] D. D. D. Ma, C. S. Lee, F. C. K. Au, S. Y. Tong, and S. T. Lee, Science **299**, 1874 (2003).
[75] J. C. Inkson, *Many-Body Theory of Solids* (Plenum, New York, 1984).
[76] W. G. Aulbur, L. Jönsson, and J. W. Wilkins, Solid State Phys. **54**, 1 (1999).
[77] M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985).
[78] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[79] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10159 (1988).
[80] C. Delerue, M. Lannoo, and G. Allan, Phys. Rev. Lett. **84**, 2457 (2000); **89**, 249901(E) (2002).
[81] L. E. Brus, J. Chem. Phys. **79**, 5566 (1983).
[82] L. E. Brus, J. Chem. Phys. **80**, 4403 (1984).
[83] M. Lannoo, C. Delerue, and G. Allan, Phys. Rev. Lett. **74**, 3415 (1995).
[84] G. Allan, C. Delerue, M. Lannoo, and E. Martin, Phys. Rev. B **52**, 11982 (1995).
[85] M. Kumagai and T. Takagahara, Phys. Rev. B **40**, 12359 (1989).
[86] In principle, the static dielectric constant ε can be split into an electronic contribution ε^{∞} and a ionic contribution (not relevant in Si/Ge). The former accounts for the screening of a point charge by the electrons, while the latter accounts for the screening by the ions. The ions typically respond on much larger time scales than the electrons and will not be able to follow an additional particle traveling in the nanowire. Therefore ε should be replaced with ε^{∞} in Sec. III when dealing with III-V materials.
[87] The change of sign in the self-energy correction when going from electrons to holes can be traced back to the so-called "screened exchange" contribution to the *GW* [or its simpler (Refs. 46 and 76), static COHSEX (Coulomb hole and screened exchange)] approximation.
[88] T. Stöckli, Z. L. Wang, J. M. Bonard, P. Stadelmann, and A. Châtelain, Philos. Mag. B **79**, 1531 (1999).
[89] M. Rohlfing, N. P. Wang, P. Krüger, and J. Pollmann, Phys. Rev. Lett. **91**, 256802 (2003).
[90] We have recomputed the radius of the nanowires according to Table I of Ref. 45 and Eq. (1).
[91] L. W. Wang and A. Zunger, J. Chem. Phys. **100**, 2394 (1994).
[92] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, 2nd ed. (SIAM, Philadelphia, 1994).
[93] M. V. Fernández-Serra, C. Adessi, and X. Blase, Phys. Rev. Lett. (to be published).