Convergence study of neutral and charged defect formation energies in Si nanowires

Riccardo Rurali,1,2 Maurizia Palumbo,3 and Xavier Cartoixà1,*

1Departament d’Enginyeria Electrònica, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
2Institut de Ciència de Materials de Barcelona (CSIC), Campus de Bellaterra, 08193 Bellaterra, Barcelona, Spain
3Dipartimento di Fisica, European Theoretical Spectroscopy Facility (ETSF), Università di Roma “Tor Vergata,” Via della Ricerca Scientifica 1, 00133 Roma, Italy

(Received 24 November 2009; revised manuscript received 30 April 2010; published 3 June 2010)

We present a systematic study of the convergence with the supercell size of the defect formation energy of a substitutional Al impurity in a Si nanowire within density-functional theory. We discuss the case of neutral and charged impurities, highlighting the importance, in the presence of ionized defects and periodic boundary conditions, of including a Madelung-like correction to the total energy in order to get correct results, which cannot be obtained by the simple use of a large vacuum buffer layer or an extrapolation thereof.

DOI: 10.1103/PhysRevB.81.235304 PACS number(s): 73.20.Hb, 61.72.jj, 61.46.Km

Theoretical modeling of Si nanowires (SiNWs) has received considerable interest in recent years.1 The study of impurities, in particular, has been the subject of many investigations,2–18 due to their important role in several applications such as electron devices,19–21 photovoltaics,22,23 and photonics.24,25 Unfortunately, it is often difficult to compare results obtained by different authors. Although it can be generally assumed that all the published calculations are converged in terms of basis set and Brillouin zone sampling, the supercell sizes used are sometimes very different and may lead to qualitatively different results.

When using periodic boundary conditions with a defective system, the supercell has to be large enough to allow neglecting the interaction of the impurity with its periodic replicas.26 While in bulk Si there are well established criteria for the most common defects (see for instance the detailed convergence study of the electronic structure of the Si monovacancy in Ref. 27), and the spurious electrostatic interactions between slab instances have been recently addressed,28 no systematic studies have been performed for nanowires. There are two main differences with bulk that should be submitted to careful scrutiny: (i) in very thin wires the extension of the defect wave function is affected by quantum confinement; (ii) a vacuum buffer must be used in the non-periodic directions to separate the nanowire from its periodic replicas.

In this work, we have studied a substitutional Al impurity in the innermost part of a 1.0 nm SiNW grown along the (110) direction, focusing on the requirements on the computational cell size needed to obtain converged results. We also discuss the effect of image charges on the energetics of the calculation, showing that the problem of the charged impurity in a nanowire is substantially more difficult than in bulk.

First-principles electronic structure calculations have been performed within the density-functional theory (DFT), as implemented in the SIESTA package.29 We have used an optimized double-$\zeta$ polarized basis set30 for the valence electrons, modeling the core electrons with norm-conserving pseudopotentials of the Troullier-Martins type.31 We have used the generalized gradient approximation (GGA) (Ref. 32) for the exchange-correlation energy. The Brillouin zone has been sampled with 2 $k$ points in the supercell, which we have verified to yield converged results for the smallest of the unit cells used in this work. Full structural relaxations were carried until the maximum force on the atoms was lower than 0.02 eV/Å, considering both the neutral and the ionized impurity (Al$^{+1}$ and Al$^{-1}$). In the case of the Al$^{+1}$ substitutional, where the computational supercell should have a finite net charge we use a compensating background to achieve charge neutrality.

For unit cells with a net charge, the presence of the replica charges and the compensating background introduces a spurious electrostatic contribution to the total energy that must be subtracted. In bulk, a Madelung energy is computed in order to remove the leading order contribution.34 However, reduced dimensionality systems feature, in principle, a dielectric tensor rather than a constant, preventing the straightforward use of the Madelung correction. We have recently generalized the Madelung correction to the case of an arbitrary static dielectric tensor $\varepsilon$, and this generalization will be used throughout this paper.

The dielectric tensor of the SiNWs we study will only have diagonal components: $\varepsilon_x$, $\varepsilon_y$, and $\varepsilon_z$, for the directions perpendicular ($x, y \rightarrow L_x$) and parallel to the wire ($z \rightarrow L_z$). Since a spacer layer between the SiNW instances is present (see Fig. 1), the dielectric tensor components obtained from the calculation, which we call $\varepsilon_i$ and $\varepsilon_i'$, refer to the vacuum+SiNW composite. While heuristic arguments can be used to estimate the elements of the dielectric tensor, an $ab initio$ calculation of the values of $\varepsilon_z$ and $\varepsilon_z'$ is required for two reasons: (i) both values in principle depend on the supercell size in the nonperiodic directions, $L_x$ and $L_y$; (ii) while in the limit of isolated nanowire ($L_x, L_y \rightarrow \infty$) one can assume $\varepsilon_i = 1$ (the impurities interact through vacuum, Fig. 1), $\varepsilon_{i'}$, whose value can depend in principle on the wire diameter (through, i.e., quantum confinement effects), is not known and must be calculated.

Here, we calculate the elements of $\varepsilon$ within the random phase approximation (RPA),35,36 taking into account local fields, by means of a plane-wave approach.37,38 Due to its small contribution,39 the GGA exchange-correlation kernel is not included in the present calculations of the macroscopic dielectric function. The possibility to include many-body effects, such as the self-energy corrections to DFT-GGA electronic band structure and excitonic effects has not been ex-
explored here, due to the very high computational cost of this kind of calculations and to the probable compensation of these two contributions.\textsuperscript{40,41} Due to the use of nonlocal norm-conserving pseudopotentials, the contribution of $[V_{\text{NL}}, \mathbf{r}]$ (where $V_{\text{NL}}$ is the nonlocal part of the pseudopotentials) has been taken into account\textsuperscript{39} in the calculation of the velocity operator $v$. We verified that neglecting this term leads to an overestimation of the static dielectric function for bulk silicon of about 28%. Overestimation ranging from 4% to 20% has been found for $\varepsilon_{\parallel}$ and $\varepsilon_{\perp}$ of the SiNW (it has been checked that $\varepsilon_{\parallel,\text{SiNW}} = \varepsilon_{\perp,\text{SiNW}}$). The obtained values of $\varepsilon_{\parallel}$ and $\varepsilon_{\perp}$ are reported in Fig. 2 as a function of the supercell size along the directions $L_x$ and $L_y$, perpendicular to the wire axis. Similar to other works on SiNWs,\textsuperscript{42-44} we have found that the main part of local-field effects can be taken into account using the effective medium theory (EMT) by means of the Maxwell-Garnett (MG) formulas for the macroscopic dielectric function.\textsuperscript{45} Nevertheless, differently from these previous works, we are assuming here an intrinsic anisotropy of the dielectric response of the isolated NWs, and use the MG formulas, for SiNWs embedded in vacuum, in the following form:

$$\varepsilon_{\parallel,\text{MG}} = f \varepsilon_{\parallel,\text{SiNW}} + (1-f),$$

$$\varepsilon_{\perp,\text{MG}} = 1 + \frac{2f\alpha}{1-f\alpha},$$

where $f$ is the filling factor defined as the ratio between the wire volume and the total volume, $\alpha = \varepsilon_{\parallel,\text{SiNW}}^{-1}$ is the wire polarizability, and $\varepsilon_{\parallel,\text{SiNW}}(\varepsilon_{\perp,\text{SiNW}})$ is the parallel (perpendicular) component of the SiNW dielectric tensor, and they are also shown in Fig. 2. The best fitting values for $\varepsilon_{\parallel,\text{SiNW}}$ and $\varepsilon_{\perp,\text{SiNW}}$ correspond to 7.75 and 4.84, respectively,\textsuperscript{46} indicating that the hydrogenation termination as well as the contribution of quantum confinement effects may play a significant role in the description of the static macroscopic dielectric function of the considered [110] thin SiNW.

Although the results discussed in Fig. 2, are interesting per se, the complexity of the interaction of a charge with its images cannot be captured by these single effective macroscopic dielectric constants. In fact, the choice of the values for the elements of $\hat{\varepsilon}$ that will enter the Madelung-like correction needs to be carefully considered. Since the interaction between the replicas in different wires takes place through vacuum and the SiNW, it is logical to choose the computed $\varepsilon_\perp$ for the $xx$ and $yy$ entries of the tensor. On the other hand, it must be noted that the electrostatic energy between two point charges located at the center of a dielectric cylinder of radius $R$ and separated a distance $z$ is shown in the Appendix to be, for large $z/R$,

$$W(z) \sim \frac{e^2}{4\pi\varepsilon_2 z} \left(1 + \frac{R^2}{z^2} \left(\varepsilon_2 - 1\right) \left(2 \ln \frac{2z}{R} - 3\right)\right),$$

where $\varepsilon_2$ is the cylinder (surrounding medium) dielectric constant. Thus, a rigorous correction of the electrostatic energy between the replicas would ask for a $z$ dependence of the tensor $zz$ component, which is not considered in our treatment in Ref. 18 nor will be considered here. Given that the vast majority of replicas lie far away from the charged defect, we set the tensor $zz$ component to 1 (i.e., $\varepsilon_2$) in what remains of this article (cf. the Appendix).

We start the convergence by studying the behavior of the total energy with respect to the supercell size in the wire periodic direction, $L_z$. To describe the periodicity of the wire along its axis, $L_z$, must be an integer multiple of the lattice parameter of the primitive cell, which we previously optimized and amounts to 3.943 Å. We considered supercells...
made from 4 to 16 primitive cells, corresponding to a separation between the impurity and its closest replicas along the wire axis from ~16 to ~63 Å. The results are summarized in Fig. 3, where we plot the (Madelung corrected) formation energy $E_F$ of the Al$_{Si}$ and Al$_{Si}^{-1}$ defects—the formation energies are evaluated at the top of the valence band (i.e., electron Fermi level $\mu_e$ set to zero) of the undoped wire, $E_{VBM}$, and $q\mu_e$ must be added to obtain the formation energies at different $\mu_e$s—and the transition level (0/−1). We use the definition of $E_F$ for a one-dimensional system given in Ref. 18, while the transition level is defined as the $\mu_e$ for which the formation energies of the neutral and the charged impurity are equal, and is simply obtained as

$$E(0/−1) = E(Al_{Si}^{-1}) - E(Al_{Si}) - E_{VBM},$$

where $E(Al_{Si})$ and $E(Al_{Si}^{-1})$ are the total energy of the system with a neutral and a charged Al defect, respectively.

As can be seen [Fig. 3(a)], the formation energy of the neutral defect is readily obtained with a supercell made of five primitive cells (amounting to a spacing of 24 Å between adjacent defects), whereas the ionized defect converges more slowly and 15–16 primitive cells are needed. Note that the 1/L$_z$ behavior of the convergence is not completely removed with our Madelung-like correction, due to the previously mentioned point that the effective dielectric constant has a $z$ dependence.

Figure 3(b) displays the ionization energy referred to the valence-band edge (black squares), whose convergence is limited by the slowest term in Eq. (4), $E(Al_{Si}^{-1})$ in this case. In the same panel we show the same quantity calculated without the Madelung correction (red triangles), converging more slowly and to an incorrect value (cf. the discussion of Fig. 4 below). The purpose of the analysis of Fig. 3 was to identify the effective range of perturbation in the electron charge due to the impurity, checking the convergence of the formation energy with respect to the size of the supercell in the axial direction. For this reason, we have used a very thick vacuum buffer (~100 Å) to separate the wire from its periodic replicas in the $x$ and $y$ directions.

Next, we turn to the study of the convergence of the total energy with respect to the vacuum buffer thickness surrounding the wire, i.e., the supercell size in the non-periodic directions, $L_x$, $L_y$. To isolate the convergence effect solely due to the vacuum, we use a 15-PC supercell, which gives converged values of $E_F$ with respect to $L_z$. We observe in Fig. 4(a) that, provided that the proper correction is applied, a vacuum buffer separation of 50 Å is enough to obtain well-converged results for the charged impurity. On the other hand, the formation energy of Al$_{Si}$ is virtually insensitive to $L_x$ ($L_y$).

Again, in the bottom panel the ionization energy with and without the Madelung correction is displayed. Since in the lateral direction we do not have a $z$ dependence of the xx and yy components of the effective dielectric tensor, we are capable of completely removing the $1/L_z$ dependence of the convergence. An $L_x, L_y \rightarrow \infty$ extrapolation for the ionization energy of the uncorrected case would reveal that the ionization energy tends to a different value (1.31 eV) from the Madelung-corrected case (0.99 eV). This is due to the limiting process, as $L_z$ is kept to the 15-PC value while $L_x$ and $L_y$ increase, causing the interaction between the replicas along the wire to dominate over the decreasing effect of the compensating background. Thus, the uncorrected value will include a spurious contribution to the total energy even at large
vacuum buffer separations. Of course, the limiting behavior when all $L_{x,y,z} \to \infty$ at the same time is the expected one, namely, that the uncorrected and corrected values converge toward the same magnitude, although the slow convergence rate for the uncorrected calculations renders them impractical and stresses the importance of using our proposed corrections.

In summary, we have carried out a systematic study of the convergence of the formation energy of a representative point defect in a thin SiNWs within DFT. For neutral defects we have found that supercells must allow a longitudinal separation between the impurity and its closest periodic replica of at least 24 Å, which is larger—and sometimes even much larger—that the values chosen in the results published so far. The vacuum separating the wire from its replicas along the perpendicular directions, on the other hand, is not apparent.

The integral over the semipositive real axis to an integral over the semipositive imaginary axis, on the other hand, is not a very critical parameter and the smallest value considered in this study is sufficient to provide converged values of the formation energy. Things change significantly with charged impurities, where a vacuum buffer of ~50 Å and a 15-PC supercell (59.1 Å) are required for satisfactory convergence. Finally, our results stress the importance to correct for the spurious interactions induced by the introduction of the compensating background, although feasible alternatives to the crude Madelung correction, already explored for bulk systems, should be investigated.

Financial support by the Ramón y Cajal program of the Ministerio de Ciencia e Innovación and funding under Contracts No. TEC2009-06986 and No. FIS2009-12721-C04-03 are greatly acknowledged. M.P. acknowledges R. Del Sole for useful discussions, funding from FP7/2007-2013 under Grant No. 211956 and from MIUR-PRIN 2007.

**APPENDIX**

The long range behavior of the electrostatic interaction energy between two point charges inside a dielectric cylinder is not completely clear. Lugo et al. provided expressions based on the image charge method, but a closer analysis seems to indicate that their expressions are valid for a dielectric sphere, rather than a cylinder. Based on physical arguments, Diarra et al. argued that, in a 1D system, screening is completely inefficient at distances $z$ large compared to the cylinder radius $R$, and that the effective $\epsilon$ is thus one. Bányaï et al. and Slachmuylders and co-workers provided analytical expressions for the interaction energy in the general case, but these are in series form, and the asymptotic behavior is not apparent.

In this appendix, we will carry out a derivation of the large $z/R$ behavior of the electrostatic interaction energy $W$ between two point charges separated a vertical distance $z$, both of them at the center of a dielectric cylinder with radius $R$, dielectric constant $\epsilon_1$ and surrounding medium with $\epsilon_2$. We start with a rewriting of the general result given in integral form by

\[
W(z) = \frac{\epsilon^2}{4\pi\epsilon_1} \left( \frac{1}{z} \right) \frac{\epsilon_1}{\epsilon_2} - \frac{1}{R} \times \int_0^\infty dx \cos \left( \frac{z}{R} x \right) \frac{K_0(x)K_1(x)}{I_0(x)K_1(x) + \frac{\epsilon_1}{\epsilon_2} I_1(x)K_0(x)}.
\]

(A1)

where $I_0(x)$ and $K_0(x)$ are the modified Bessel functions of the first and second kind, respectively.

The integral above can be approximated for large $z/R$ through the following procedure. We first write the cosine as the real part of an imaginary exponent, and the real part can be obtained through the following procedure. We first write the cosine as a real part of an imaginary exponent, and the real part can be obtained through the following procedure.

\[
I(z/R) = \int_0^\infty dx \cos \left( \frac{z}{R} x \right) \frac{K_0(x)K_1(x)}{I_0(x)K_1(x) + \frac{\epsilon_1}{\epsilon_2} I_1(x)K_0(x)}
\]

\[
= \Re \left( \int_0^{\infty} dx e^{i\frac{z}{R} x} \frac{K_0(x)K_1(x)}{I_0(x)K_1(x) + \frac{\epsilon_1}{\epsilon_2} I_1(x)K_0(x)} \right).
\]

(A2)

We can now make the change of variable $x = iu$ and, since we are integrating over real $u$, transform the modified Bessel functions of pure imaginary argument into Bessel functions of the first, second and third kinds of pure real argument.

![Graphical evaluation of the denominator shows that it pre-](image)
\[
I(z/R) = \pi \left\{ \int_{0}^{\infty} d\epsilon \frac{u e^{-\epsilon z/R}}{2} \left[ \frac{H_{0}^{(1)}(-\epsilon/u)H_{1}^{(1)}(-\epsilon/u)}{J_{0}(u)H_{1}^{(1)}(-\epsilon/u)} + \frac{\epsilon_{1}}{\epsilon_{2}} \frac{J_{1}(u)H_{0}^{(1)}(-\epsilon/u)}{J_{0}(u)H_{1}^{(1)}(-\epsilon/u)} \right] \right. \\
= -\frac{\pi}{2} \int_{0}^{\infty} du e^{-\epsilon z/R} \left\{ 1 - \frac{\epsilon_{1}}{\epsilon_{2}} \left[ \gamma + \ln(\epsilon/u) \right] u^{2} \right\} ,
\]

(A3)

where, in the last step, we have Taylor expanded the quotient since, for large enough \(z/R\), only a small \(u\) region close to the origin will contribute to the integral, \(\gamma\) is Euler’s constant and we have kept only the real terms.

The integral is now readily evaluated, yielding the asymptotic behavior for the interaction energy [see Eq. (3)], which, when keeping only the slowest decaying term, goes like \(e^{2}/(4\pi \epsilon_{2} z)\), showing that the presence of the dielectric cylinder is not felt at large \(z/R\).

The goodness of Eq. (3) is shown in Fig. 5, where the ratio of Eq. (3) to a numerical evaluation of the exact Eq. (A1) is represented. The value used for the cylinder dielectric constant is as found from the first principles calculations, \(\epsilon_{1}=7.75\) (cf. Fig. 2), while the outer medium is chosen to be vacuum with \(\epsilon_{2}=1\). It is clearly seen that the third order expression Eq. (3) provides a good approximation for \(z/R \geq 11\), while the first order expression \(e^{2}/(4\pi \epsilon_{2} z)\) is only good for larger \(z/R\). Note that, if we rewrite Eq. (A1) as \(e^{2}/[4\pi \epsilon(z) z]\), the solid line in Fig. 5 provides the effective epsilon \(\epsilon(z)\) seen by the point charge due to the image charges.
[46] These values of $\varepsilon_{\perp,\text{SiNW}}$ and $\varepsilon_{\parallel,\text{SiNW}}$ clearly depend on the estimated filling factor or, in other words, on the value of the NW’s occupied area. Including the hydrogen atoms in this area, the NW area amounts to 83 Å². A change of about 5% in $f$ induces a variation of $\sim0.5$ in the fitting values of the dielectric components.