10.1098/rsta.2002.1127

THE ROYAL SOCIETY

# Confinement effects and tunnelling through quantum dots

By Michel Lannoo[1], Christophe Delerue[2], Guy Allan[2] and Yann-Michel Niquet[2]

[1]*Laboratoire Matériaux et Microélectronique de Provence, Institut Supérieur d'Electronique de la Méditerranée, Place Georges Pompidou, 88000 Toulon, France*
[2]*Institut d'Electronique et de Microélectronique du Nord, Département ISEN, 41 bld Vauban, 59046 Lille Cedex, France*

*Published online 17 December 2002*

Several recent theoretical advances concerning semiconductor quantum dots are reviewed. First of all, the effect of the quantum confinement on the energy gap is revisited on the basis of GW and Bethe–Salpeter calculations, showing that the excitonic gap is practically equal to the ordinary eigenvalue gap of single-particle approximations. The second part demonstrates that it is now possible to calculate the conductance peaks for the tunnelling current through a nanostructure. Finally, we discuss in some detail the concept of a macroscopic dielectric constant for nanostructures, showing that, except for a thin surface layer, the local dielectric constant still keeps its bulk value down to pretty small nanostructures.

**Keywords: quantum dots; nanocrystals; confinement**

## 1. Size dependence of the gap

We consider here the gap-versus-size problem for nanocrystallites. This can be obtained from three approaches as an eigenvalue gap, a quasi-particle gap or an excitonic gap. We discuss the results and relations between these three distinct definitions showing that the excitonic gap is almost equal to the eigenvalue gap.

### (*a*) *The eigenvalue gap,* $\varepsilon_g^0$

This is defined as a difference in eigenvalues of standard one-particle theories such as the Hartree–Fock theory, the local density approximation (LDA) or semi-empirical theories. We compare here different calculations of $\varepsilon_g^0$ for Si nanocrystals.

*Ab initio*, self-consistent methods such as the LDA can only be used for small clusters (less than 1000 atoms) with high symmetry (Delley & Steigmeier 1995). They are not suitable for the computation of optical and transport properties in realistic situations. Semi-empirical, non-self-consistent methods—such as pseudo-potentials (PP) (Wang & Zunger 1994*a*; Zunger & Wang 1996), tight binding (TB) (Slater & Koster 1954; Proot *et al*. 1992; Delerue *et al*. 1993; Ren 1997), or $\boldsymbol{k}\cdot\boldsymbol{p}$ (Takagahara & Takeda 1992)—can solve much larger problems. Semi-empirical methods are

One contribution of 13 to a Discussion Meeting 'Quantum dots: science on the smallest scale?'.

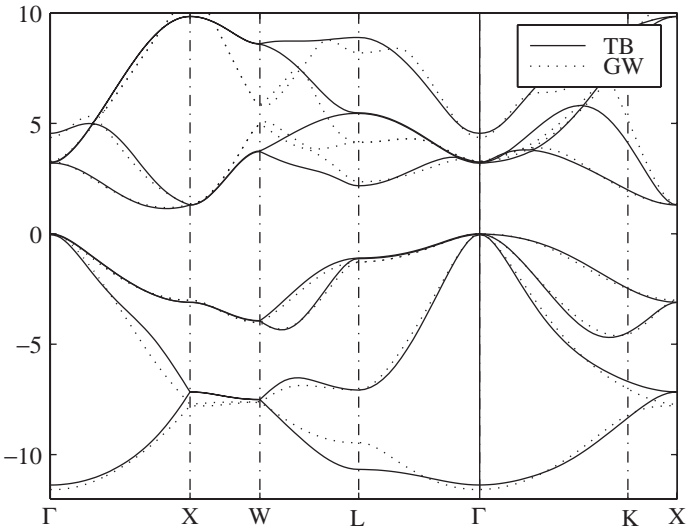


Figure 1. Band structure of bulk Si in the orthogonal third-nearest-neighbour TB and GW models. TB parameters are given in Niquet *et al*. (2000).

designed to make the best possible approximation to the self-consistent, one-particle Hamiltonian $H_0$ in bulk material, either in the whole first Brillouin zone (PP, TB) or around specific $k$-points ($\boldsymbol{k}\cdot\boldsymbol{p}$). They involve adjustable parameters (e.g. effective masses, TB interaction parameters, etc.) that are fitted to experimental data or *ab initio* bulk band structures. These parameters are then transferred to the nanostructures (i.e. $H = H_0$ inside the nanostructure) with appropriate boundary conditions. The better the bulk description and boundary conditions, the better the electronic structure we expect in nanostructures.

We present here an improved TB description designed to give accurate results over a whole range of sizes (Niquet *et al*. 2000). For this, we fit the TB parameters not only on bulk band energies in the first Brillouin zone but also on effective masses. This was not the case for previous TB treatments but is essential if one wants to obtain accurate results for large nanostructures, where the effective mass approximation (EMA) or $\boldsymbol{k}\cdot\boldsymbol{p}$ models become exact. We also include spin–orbit coupling. From improved computer codes we are able to apply TB to nanostructures with feature sizes in the range 1–12 nm, which has never been achieved previously. This extended range allows us to study precisely the overlap region where both TB and $\boldsymbol{k}\cdot\boldsymbol{p}$ or the EMA are accurate. We provide analytic fits to the confinement energies, which are practically exact over the whole range of sizes.

When fitting TB parameters on bulk band energies as well as conduction or valence-band effective masses, the total root-mean-squared (RMS) error is taken as a weighted average of the RMS error on bulk band energies and on effective masses. It is minimized with a conjugate-gradients algorithm. Bulk band energies are selected from a GW band structure (Hybertsen & Louie 1985; L. Reining 1999, personal communication), which is the best available at the moment. An orthogonal $sp^3$ TB model including all of the interactions up to the third nearest neighbours and taking into account three centres integrals is considered here. The set of 20 TB parameters is reported in the work of Niquet *et al*. (2000). As can be seen from figure 1, the

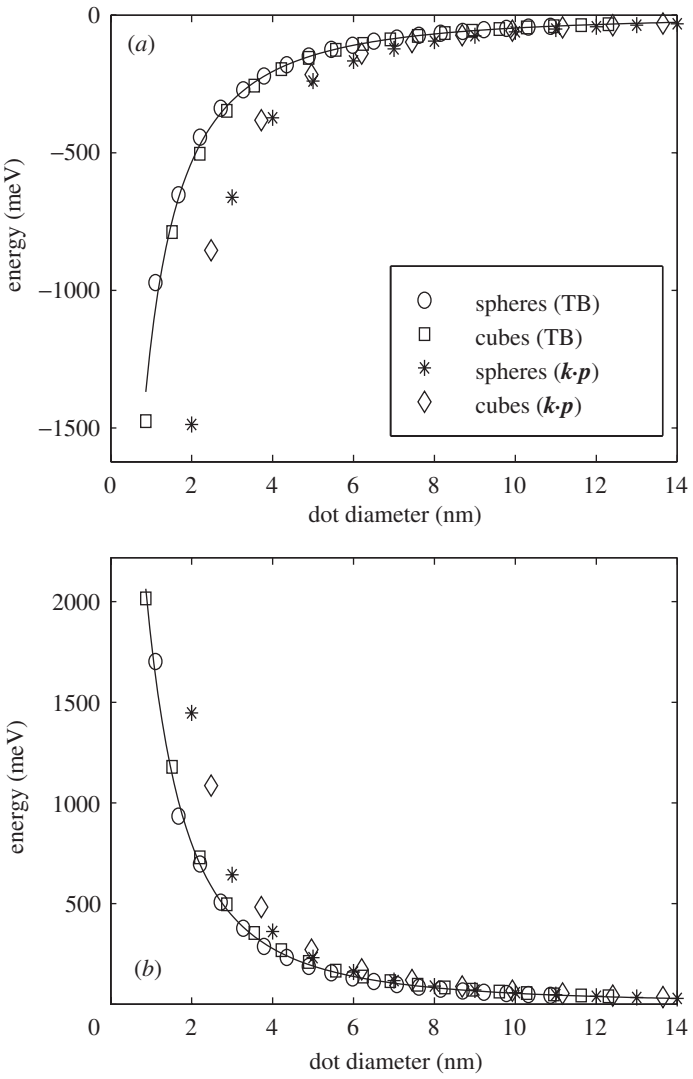


Figure 2. Energy of the ($a$) highest occupied state and ($b$) lowest unoccupied state for spherical and cubic ((100) faces) Si dots. Results are given versus effective dot diameter $d$ (see text), within TB and $\boldsymbol{k} \cdot \boldsymbol{p}$ or EMA models.

overall quality of the fit is excellent. This is, to our knowledge, the best description of the band structure obtained so far in Si with a $sp^3$ TB model.

To achieve better boundary conditions in nanostructures saturated by hydrogen atoms, the TB parameters describing the Si–H interactions have been fitted on the $SiH_4$ experimental excitonic gap (Rohlfing & Louie 1998) and charge transfer calculated within LDA. H atoms are described by their 1s orbital. Si–H parameters are also reported in the work of Niquet *et al.* (2000).

Results for spherical and cubic ((100) faces) Si dots are shown in figure 2. The energy of the highest occupied valence state ($\varepsilon_{\mathrm{v}}$) and lowest unoccupied conduction state ($\varepsilon_{\mathrm{c}}$) are plotted against the effective diameter $d$ of the dot, which is the diameter

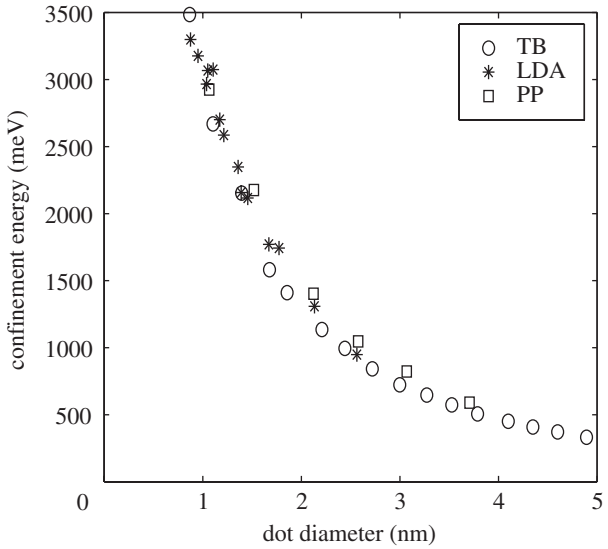


Figure 3. Comparison of confinement energy in spherical Si dots, between our TB model and PPs or LDA.

of the sphere with the same volume as the dot. The energy of these states can be fitted in the whole range with the expressions

$$\varepsilon_{\mathrm{v}}(d) = \frac{K_{\mathrm{v}}}{d^2 + a_{\mathrm{v}}d + b_{\mathrm{v}}} + \varepsilon_{\mathrm{v}}(\infty), \tag{1.1}$$

$$\varepsilon_{\mathrm{c}}(d) = \frac{K_{\mathrm{c}}}{d^2 + a_{\mathrm{c}}d + b_{\mathrm{c}}} + \varepsilon_{\mathrm{c}}(\infty), \tag{1.2}$$

where $K$, $a$ and $b$ are adjustable constants; $\varepsilon_{\mathrm{c}}(\infty) - \varepsilon_{\mathrm{v}}(\infty) = 1.143\,\mathrm{eV}$ is the bulk bandgap energy. This expression is more accurate than the widely used fit $K/d^{\alpha}$ when large range of dimensions are considered. It correctly behaves like $1/d^2$ in large structures, so that it can be considered valid over the whole range of sizes. The fits are reported in figure 2 (solid lines). The parameters are given in Niquet *et al.* (2000).

We now proceed to compare the TB model with other descriptions such as PP or LDA or with other TB models in the case of small spherical Si dots. Our results are first compared with those calculated with PP (Zunger & Wang 1996) in the range 1–4 nm. As shown in figure 3, they are in good agreement. Comparison with the *ab initio* LDA calculations (Delley & Steigmeier 1995) corrected for the bulk bandgap error of 0.65 eV shows that our results are extremely close down to the smallest clusters, except for some oscillations in the LDA that do not exist in TB. Finally, the agreement with other more complex TB models with overall band structure of similar quality, such as an $\mathrm{sp^3d^5s^*}$ model (Jancu *et al.* 1998) or a non-orthogonal TB model (Allen *et al.* 1986) is also excellent for small ($d_0 < 4$ nm) crystallites. Figure 3 thus provides a fairly coherent description of what will be called the eigenvalue gap $\varepsilon_{\mathrm{g}}^0 = \varepsilon_{\mathrm{c}}(d) - \varepsilon_{\mathrm{v}}(d)$ in the following.

### (*b*) *Quasi-particle and excitonic gaps*

For silicon, we have seen that the corrected LDA bandgaps $\varepsilon_g^0$ are in quite good agreement with the best TB or PP results. To find the excitonic gap, a common practice has been to subtract the screened direct electron–hole attraction from this value. However, the whole procedure is not clearly justified and conflicting points of view (Ögüt *et al*. 1997; Godby & White 1998; Franceschetti *et al*. 1999) have been expressed concerning its validity. To clarify this problem, we express the excitonic gap $\varepsilon_g^{\mathrm{exc}}$ as the difference between the quasi-particle gap $\varepsilon_g^{\mathrm{qp}}$ (the difference between the separate electron and hole quasi-particle energies) and $E_{\mathrm{coul}}$, the attractive interaction between these two quasi-particles,

$$\varepsilon_g^{\mathrm{exc}} = \varepsilon_g^{\mathrm{qp}} - E_{\mathrm{coul}} = \varepsilon_g^0 + \delta\Sigma - E_{\mathrm{coul}}, \tag{1.3}$$

where $\varepsilon_g^{\mathrm{qp}}$ is written as the sum of the independent particle value $\varepsilon_g^0$ (obtained from § 1 *a*) and a self-energy correction $\delta\Sigma$. The main result that we obtained (Delerue *et al*. 2000) is that there is strong cancellation between the two large quantities $\delta\Sigma - \delta\Sigma_{\mathrm{b}}$ (where $\delta\Sigma_{\mathrm{b}}$ is the bulk value) and $E_{\mathrm{coul}}$, such that $\varepsilon_g^{\mathrm{exc}} \approx \varepsilon_g^0 + \delta\Sigma_{\mathrm{b}}$. This justifies why the above-mentioned single-particle calculations should yield accurate results. We also show that $\delta\Sigma$ and $E_{\mathrm{coul}}$ are dominated to a large extent by classical electrostatic contributions.

To perform these calculations, we have proceeded in two steps, firstly calculating the separate electron and hole quasi-particle energies via the GW method (Hedin & Lundqvist 1969) and, secondly, determining the attractive Coulomb interaction between these quasi-particles by resolution of the Bethe–Salpeter equation. Similar work has already been successfully applied from an *ab initio* point of view to bulk semiconductors (Albrecht *et al*. 1997), $Na_4$ clusters (Onida *et al*. 1995) and small silicon clusters saturated by hydrogen atoms (up to $Si_{14}H_{20}$) (Rholfing & Louie 1998). However, the computation is very time consuming and cannot be extended to nanocrystals. This is why we have chosen a TB formulation which allows us to reduce the complex integrals occurring in the GW or Bethe–Salpeter equations to matrices whose size is given by the number of atoms in the system. Details about the calculations are given in Delerue *et al*. (2000).

In figure 4 we plot $\delta\Sigma - \delta\Sigma_{\mathrm{b}}$ calculated for nanocrystals containing up to 275 Si atoms from three approaches (Delerue *et al*. 2000) that give very similar results, especially for a crystallite radius $R > 0.6$ nm. In fact, the main contribution to $\delta\Sigma - \delta\Sigma_{\mathrm{b}}$ is actually a classical electrostatic effect (Lannoo *et al*. 1995): when one puts an extra electron (or hole) into a nanocrystal at site $\boldsymbol{r}$, the electronic relaxation (screening) induces charges at the surface and the extra particle interacts with this self-image charge distribution, leading to a self-polarization energy $E_{\mathrm{pol}}(\boldsymbol{r})$ as discussed by Lannoo *et al*. (1995). We see in figure 4 that the resulting self-energy $\Sigma_{\mathrm{pol}}$ for the separate addition of one electron plus a hole into the cluster gives the essential contribution to $\delta\Sigma - \delta\Sigma_{\mathrm{b}}$.

We have then calculated the excitonic gap $\varepsilon_g^{\mathrm{exc}}$ by solving the Bethe–Salpeter equation for the two-particle Green functions (Nozières 1964; Sham & Rice 1966; Strinati 1984). We follow the same procedure as in the work of Rholfing & Louie (1998). The corresponding results are given in figure 5. However, for reasons which will become clear later, we have preferred to plot $E_{\mathrm{coul}}$ against size, taken as the difference $\varepsilon_g^{\mathrm{qp}} - \varepsilon_g^{\mathrm{exc}}$. We also compare the computed $E_{\mathrm{coul}}$ with the result of the

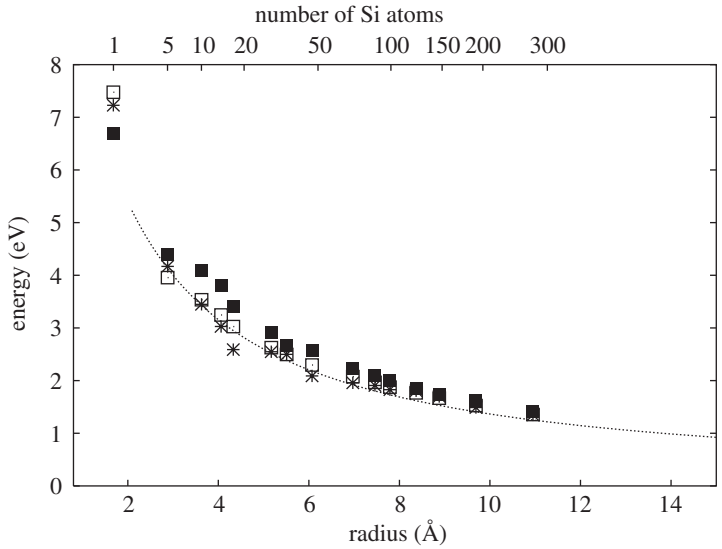


Figure 4. Variation against size of the self-energy correction $(\delta\Sigma - \delta\Sigma_{\mathrm{b}})$ in Si nanocrystals, calculated using the three different methods described in the work of Delerue *et al.* (2000). The continuous line denotes the classical electrostatic energy for the separate addition of an electron and a hole in a nanocrystal.

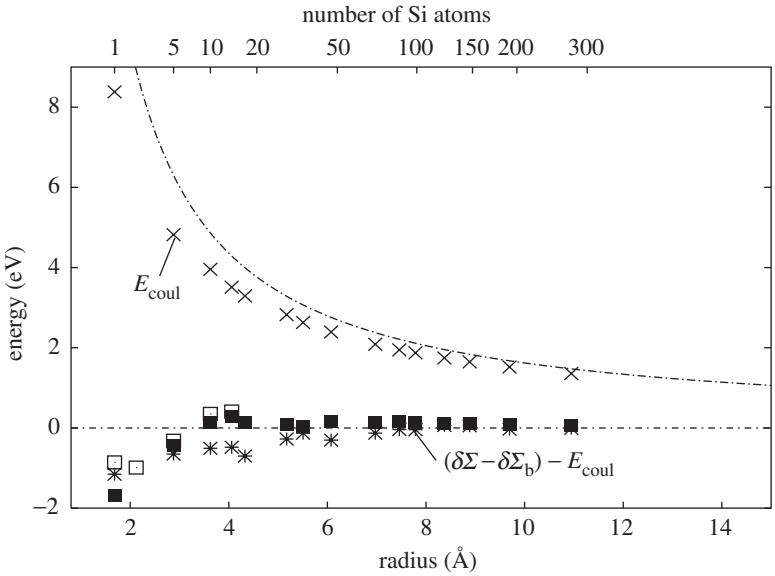


Figure 5. Graph of exciton Coulomb energy (crosses) against size in Si nanocrystals: curved line, full GW and Bethe–Salpeter calculation; straight line, classical electrostatics calculation with effective mass wave functions. The differences between the self-energy correction and the following: asterisks, present full TB calculations; closed squares, present second-LDA calculations; open squares, *ab initio* results of Rholfing & Louie (1998) are shown.

classical electrostatic argument (Lannoo *et al*. 1995), where the effective interaction for the electron and hole is the sum of two terms: a direct screened interaction plus the interaction of one particle with the polarization charge induced by the other.

In figure 5 we also plot the difference $(\delta\Sigma - \delta\Sigma_{\rm b}) - E_{\rm coul}$ for our two extreme values of $\delta\Sigma - \delta\Sigma_{\rm b}$, and we compare them with the same quantity obtained from the full *ab initio* calculation (Rholfing & Louie 1998) for $SiH_4$, $Si_5H_{12}$, $Si_{10}H_{16}$ and $Si_{14}H_{20}$. Our values fall in the same range as the *ab initio* values, especially those arising from the second LDA model. A striking feature displayed by figure 5 is that the quantities $E_{\rm coul}$ and $\delta\Sigma - \delta\Sigma_{\rm b}$, while being fairly large, compensate each other to a large degree and, for clusters with $R > 0.6$ nm, the two quantities are practically identical, so their contributions to the excitonic gap cancel each other. From equation (1.3) one thus obtains

$$\varepsilon_{\rm g}^{\rm exc} \approx \varepsilon_{\rm g}^{0} + \delta\Sigma_{\rm b}. \tag{1.4}$$

This means that $\varepsilon_{\rm g}^{\rm exc}$ is given directly by the single-particle gap $(\varepsilon_{\rm g}^{0})_{\rm TB}$ in full TB (where $\delta\Sigma_{\rm b} = 0$) and $(\varepsilon_{\rm g}^{0})_{\rm LDA} + 0.65$ eV in LDA calculations. This result not only justifies the use of single-particle calculations to find the excitonic gap but also explains the agreement between empirical and LDA results once these are shifted by the bulk correction 0.65 eV (Delley & Steigmeier 1995). Of course, the cancellation is not strictly exact, but for $R > 0.6$ nm it is verified to better than 0.2 eV in figure 5. One can also notice that equation (1.4) is likely to hold true to some extent for other semiconductor crystallites. We have checked that this is indeed the case for Ge and even for C, for which at $R = 0.8$ nm the deviation from prefect cancellation is 0.8 eV: a value that is still small compared with the gap value (12 eV in this case). Recent results based on the diffusion quantum Monte Carlo approach totally confirm the cancellation for small Si clusters (Porter *et al*. 2001).

## 2. Single-particle tunnelling through semiconductor quantum dots

Coulomb charging effects are significant in nano-sized transistors and memories. It is thus of major importance to realize nanostructures in a well-controlled way, then to study and try to understand their electronic properties. A major step in this direction has been achieved recently by Banin *et al*. (1999), who reported tunnelling-spectroscopy experiments on InAs nanocrystals which reveal rich features due to the interplay between quantum confinement and single-charge effects. In this context we show that it is now possible to perform complete self-consistent electronic structure calculations which allow a detailed understanding of the experimental data.

We use an extension of the theory of Averin *et al*. (1991) and consider a standard double-barrier tunnel junction (van Houten *et al*. 1992). It consists of two metallic electrodes, E1 and E2, weakly coupled to a semiconductor quantum dot by two tunnel junctions, J1 and J2, with capacitances $C_1$ and $C_2$. The metallic electrodes E1 and E2 are characterized by their Fermi energies $\varepsilon_{\rm f}^{1} = \varepsilon_{\rm f} - eV$ and $\varepsilon_{\rm f}^{2} = \varepsilon_{\rm f}$, where $V$ is the bias voltage and $e$ is the electronic charge. The total energy of the quantum dot charged with $n$ electrons and $p$ holes with respect to the neutral state can be written in the orthodox model as (van Houten *et al*. 1992)

$$E(\{n_i\},\{p_i\},V) = \sum_i n_i\varepsilon_i^{\rm e} - \sum_i p_i\varepsilon_i^{\rm h} - \eta eVq + \tfrac{1}{2}Uq^2, \tag{2.1}$$

where $\varepsilon_i^{\mathrm{e}}$ and $\varepsilon_i^{\mathrm{h}}$ are the conduction and valence energy levels in the quantum dot, $n_i$ and $p_i$ are electron and hole occupation numbers ($n = \sum_i n_i$, $p = \sum_i p_i$), and $q = n - p$. In terms of the junction capacitances $C_1$ and $C_2$, the charging energy $U = e^2/(C_1 + C_2)$, and $\eta = C_1/(C_1 + C_2)$ is the part of the bias voltage $V$ that drops across J2 in the neutral quantum dot. Tunnelling of an electron onto the energy level $\varepsilon_i^{\mathrm{e}}$ occurs at transition energy

$$\begin{aligned}\varepsilon_i^{\mathrm{e}}(q, q+1) &= E(n_i = 1, \{p_j\}, V) - E(n_i = 0, \{p_j\}, V) \\ &= \varepsilon_i^{\mathrm{e}} - \eta eV + U(q + \tfrac{1}{2}),\end{aligned} \tag{2.2}$$

while tunnelling out occurs at $\varepsilon_i^{\mathrm{e}}(q-1, q)$. Symmetrically, hole tunnelling in (out) occurs at $\varepsilon_i^{\mathrm{h}}(q-1, q)$ ($\varepsilon_i^{\mathrm{h}}(q, q+1)$). The total tunnelling rates through junction J$\alpha$ ($\alpha = 1, 2$) can be expressed (Niquet *et al.* 2002) in terms of the tunnelling rates $\Gamma^{\alpha}$ onto levels $\varepsilon_i^{\mathrm{e}}$ and $\varepsilon_i^{\mathrm{h}}$, following the work of Averin *et al.* (1991). The difference is that we treat both electrons and holes at the same time and incorporate the electron–hole recombination rate $R(n, p)$ from the charge state $(n, p)$ to the charge state $(n-1, p-1)$ into the master equations. At $T \to 0$ K, the $I(V)$ curve looks like a staircase (Averin *et al.* 1991). It exhibits a step each time $\varepsilon_{\mathrm{f}}^1$ or $\varepsilon_{\mathrm{f}}^2$ crosses a transition energy $\varepsilon_i^{\mathrm{e}}(q, q+1)$ or $\varepsilon_i^{\mathrm{h}}(q-1, q)$. A new charge state then becomes available in the quantum dot (addition step), or a new channel $\varepsilon_i^{\mathrm{e}}$ or $\varepsilon_i^{\mathrm{h}}$ is opened for tunnelling to a given, already available charge state (excitation step).

This behaviour is apparent in the results of Banin *et al.* (1999), who reported tunnelling spectroscopy experiments on InAs nanocrystals. The differential conductance $G(V) = \mathrm{d}I(V)/\mathrm{d}V$ is shown in figure 6 for a nanocrystal 6.4 nm in diameter. The tip was retracted from the quantum dot so that $C_1/C_2$ is maximum and $\eta$ is close to unity. A zero-current gap is observed around $V = 0$, followed by a series of conductance peaks for $V < 0$ and $V > 0$.

The energy levels $\varepsilon_i^{\mathrm{e}}$ and $\varepsilon_i^{\mathrm{h}}$ of spherical InAs nanocrystals are calculated with an orthogonal $\mathrm{sp}^3$ TB model with second-nearest-neighbours interactions. The TB parameters are fitted to the bulk InAs band structure that was calculated with the local density approximation corrected from the bandgap problem, and to the experimental effective masses (Niquet *et al.* 2000). Spin–orbit coupling is included. The surface dangling bonds are saturated with pseudo-hydrogen atoms. We find that the lowest conduction level ($^1S_e$) is s-like, twofold degenerate, and the next level ($^1P_e$) is p-like, sixfold degenerate. The highest two valence levels $1_{\mathrm{VB}}$ and $2_{\mathrm{VB}}$ are found to be fourfold degenerate.

To compare these results with the interpretation of Banin *et al.* (1999), we perform two types of calculation: firstly, we use the capacitative model of equation (2.1) with our calculated TB level structure and consider $U$ and $\eta$ as fitting parameters chosen to optimize the agreement with the position of the peaks in the $G(V)$ curve. Secondly, we make a full self-consistent treatment on a system with a realistic geometry described below. Here the ground-state energy $E_0(n, p, V)$ is self-consistently computed for a set of charge states $(n, p)$ and several voltages $V_i$. This is done in the Hartree approximation corrected from the unphysical self-interaction term (Niquet *et al.* 2002). The electrostatic potential inside the quantum dot is computed with a finite-difference method. The calculation is performed in the subspace spanned by the lowest $N_{\mathrm{VB}}$ valence states and $N_{\mathrm{CB}}$ conduction states of the neutral quantum dot calculated with the TB model. The transition energies are derived from equation

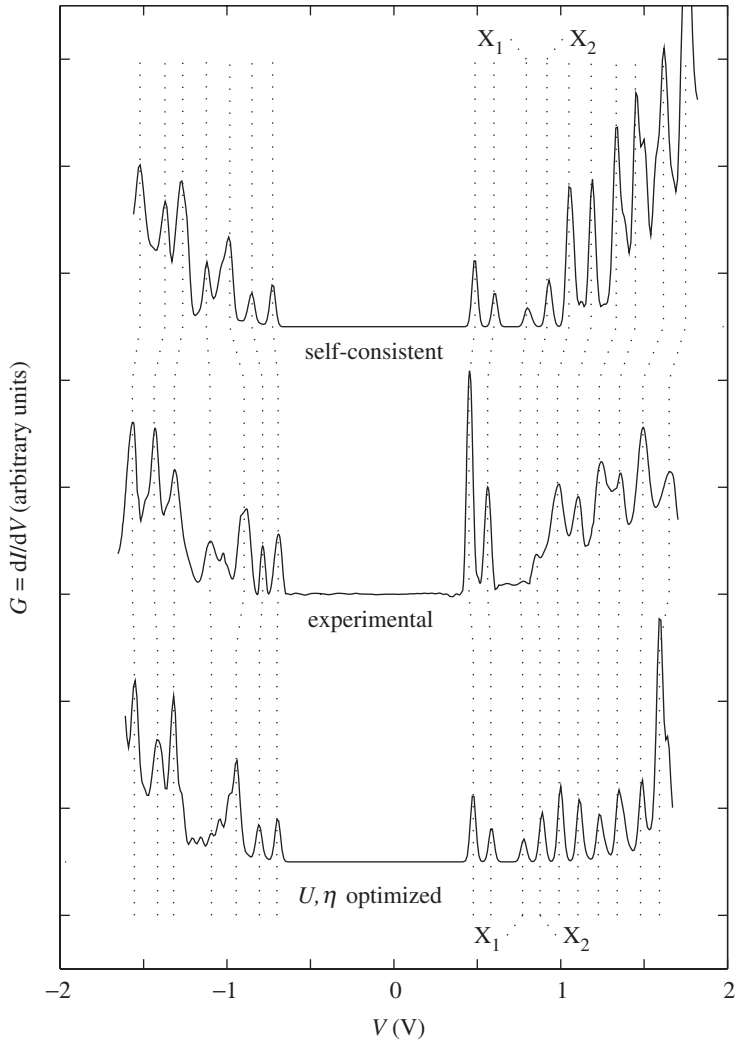


Figure 6. Comparison between calculated and experimental differential conductance $G(V)$ curves for a 6.4 nm diameter InAs nanocrystal. The optimized parameters for the capacitative model are $U = 100$ meV and $\eta = 0.9$.

(2.2). The ground-state energy is interpolated over the whole bias voltage range and excited configuration energies $E(\{n_i\}, \{p_i\}, V)$ are approximated from the spectrum of the self-consistent ground-state Hamiltonian and interpolated between successive $V_i$.

For the capacitative model, we consider a spherical InAs nanocrystal 6.4 nm in diameter. For the self-consistent calculation we obtain better results with a slightly flattened dot (Niquet *et al*. 2002). This does not significantly change the electronic structure of the quantum dot but changes only the details of the electrostatic potential. The dielectric constant of the quantum dot $\varepsilon_r = 13.6$ is calculated as in Lannoo *et al*. (1995). The quantum dot is linked to the gold substrate (E1) by a hexane dithiol layer 5 Å thick and is surrounded by a layer of molecular ligands ($\varepsilon_r = 2.6$)

5 Å thick. The radius of curvature of the Pt–Ir STM tip (E2) is $r = 2.5$ nm, and the tip–QD distance is $d = 5$ Å. Taking $N_{\mathrm{VB}} = 240$ and $N_{\mathrm{CB}} = 120$ ensures convergence of the self-consistent calculation. The zero-bias Fermi energy $\varepsilon_{\mathrm{f}} = 0.5$ eV is obtained from the position of the centre of the zero-current gap. Since an accurate calculation of the tunnelling rates $\Gamma^{\alpha}$ is not possible, we take them as adjustable parameters. A detailed analysis of the relative height of the low-lying current steps (Banin *et al*. 1999) around the zero-current gap suggests that $\Gamma^1 \simeq \Gamma^2$. Thus, we set $\Gamma^1 = \Gamma^2 = 6 \times 10^8\ \mathrm{s}^{-1}$ for electrons and $\Gamma^1 = \Gamma^2 = 3 \times 10^8\ \mathrm{s}^{-1}$ for holes to find the correct order of magnitude for the current. It is important to note that the position of the calculated conductance peaks does not depend on the tunnelling rates $\Gamma^{\alpha}$. Last, we set $R(n, p) = np/\tau$ and $\tau = 1$ ns, which is characteristic of direct-gap semiconductors.

The calculated $G(V)$ curves are compared with the experimental one in figure 6. They are broadened with a Gaussian of width $\sigma = 15$ meV. The optimized parameters for the capacitative model are $U = 100$ meV and $\eta = 0.9$ ($C_1 = 1.44$ aF, $C_2 = 0.16$ aF). As shown in figure 6 the agreement with experiment is extremely good with practically a one to one correspondence between the calculated and experimental peaks over a range of 3.5 V. The negative-bias-voltages side is clearly improved in the self-consistent calculation.

We can now proceed to a more detailed analysis of the $G(V)$ curves. For $V > 0$, the first group of two peaks is assigned to the tunnelling of electrons filling the ${}^1S_e$ level (Banin *et al*. 1999), and the next group of six peaks to the tunnelling of electrons through the ${}^1P_e$ level and tunnelling of holes. There are also two excitation peaks $\mathrm{X}_1$ and $\mathrm{X}_2$ in figure 6 (tunnelling through the ${}^1P_e$ level in the charge states $n = 0$ and $n = 1$) that are hardly visible on the experimental $G(V)$ curve but disappear at higher $\Gamma^2/\Gamma^1$ ratios. For $V < 0$, the first two peaks can be unambiguously assigned to the tunnelling of holes filling the highest valence level $1_{\mathrm{VB}}$. However, the next group of peaks is a very intricate structure involving single-hole charging peaks and tunnelling of electrons through the ${}^1S_e$ level. This disagrees with the interpretation of Banin *et al*. (1999), based on single-hole transitions, and means that the multiplicity of the peaks is not the same from one size to another as has been found experimentally. Finally, the strong increase in the current below $-1.25$ V is mainly related to the tunnelling of electrons through the ${}^1P_e$ level.

## 3. Concept of dielectric constant for a single quantum dot

It is of major interest to know when the use of macroscopic quantities such as the dielectric constant breaks down for quantum confined systems as a function of the reduction in size. Some calculations have been performed recently for spherical dots (Lannoo *et al*. 1995; Wang & Zunger 1994*b*; Tsu & Babic 1994), tending to show that one can still define a size-dependent inner dielectric constant $\epsilon_{\mathrm{in}}(R)$ but with a reduced value compared with the bulk one. However, these calculations were lacking generality and were not conclusive about the origin of this reduction with smaller size. It is generally accepted that such a reduction can result from two effects: firstly, the opening of the gap due to confinement which should reduce the first-order polarizabilities and, secondly, a boundary effect which exists even if the space-dependent susceptibility keeps its bulk value inside the dot but is zero outside. The reason most often invoked is the former, but we show in the following that it is not the case. A

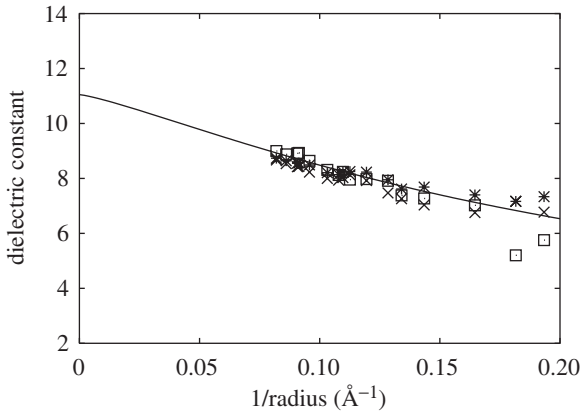


Figure 7. Dielectric constant $\epsilon_{\mathrm{in}}$ obtained by the fit of the donor potential (squares) and by the calculation of the screened electric field after application of a constant external field using two methods: interpolation of the potential (crosses) and calculation of the induced charges (asterisks). The line corresponds to $\epsilon_s$ of Wang & Zunger (1994*b*).

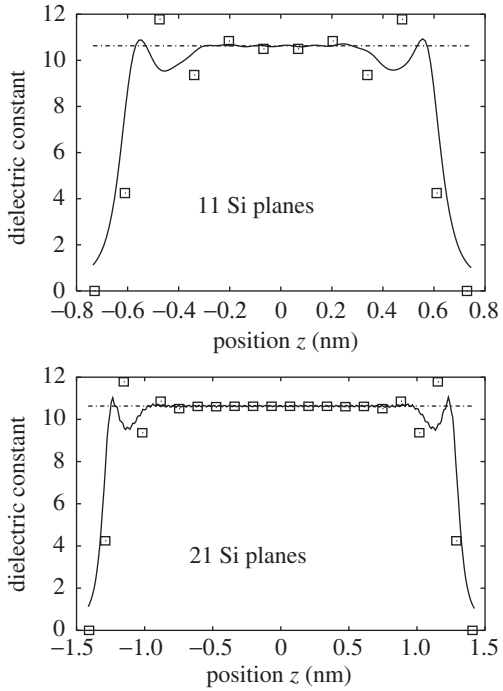


Figure 8. Calculated $E_{\mathrm{b}}/E$ for a constant applied electric field $E_{\mathrm{b}}$ for different layer thicknesses. The constant dotted line is the bulk value ($\epsilon(q=0)=10.62$). The continuous curve is the value calculated with the bulk $\epsilon(q)$.

second conclusion is that locally the concept of a dielectric constant remains valid down to very small sizes.

To describe macroscopic screening in real space, one must define macroscopic fields as averages of the microscopic fields over a suitable volume. This procedure leaves

Maxwell's equations valid for the macroscopic quantities. For bulk crystals the natural volume over which to perform this average is the unit cell. One can keep this procedure within a nanocrystal and, in principle, still get a set of equations relating the macroscopic fields. We shall here examine this question in a tight binding framework (details will be given in a future publication) which automatically provides a weighted average of the potentials over the atomic cells.

Let us first come back to the case of a hydrogenic donor impurity at the centre of a crystallite (Lannoo *et al.* 1995). The bare potential energy of the donor electron at distance $r = |\boldsymbol{r}|$ from the nucleus of charge $+e$ is $V_{\mathrm{b}}(\boldsymbol{r}) = -e^2/r$. In TB this has only diagonal matrix elements $(V_{\mathrm{b}})_j$ on each site. Screening reduces this bare potential on site $j$ to a value $V_j$. Fig. 1 of Lannoo *et al.* (1995) shows that the ratio $[V]_j/[V_{\mathrm{b}}]_j$, where $j$ runs over the Si atoms, can be fitted by a straight line when plotted against the distance from the centre. This result has a straightforward interpretation in a classical macroscopic picture, where the screened potential energy is given by

$$V(\boldsymbol{r}) = -e^2\left[\frac{1}{\epsilon_{\mathrm{in}} r} + \frac{1}{R'}\left(1 - \frac{1}{\epsilon_{\mathrm{in}}}\right)\right], \tag{3.1}$$

$R'$ being the effective radius of the macroscopic dielectric sphere corresponding to the average position of the surface polarization charge $e(1 - 1/\epsilon_{\mathrm{in}})$. In principle, $R'$ should differ slightly from $R$, the radius of the sphere containing the silicon atoms. We consider $\epsilon_{\mathrm{in}}$ and $R'$ to be adjustable parameters. Figure 7 shows the values of $\epsilon_{\mathrm{in}}$ deduced in this way. We find that $R' \approx R + 0.4$ Å over the full range of crystallite sizes. $R' > R$ is consistent with the presence of polarizable Si–H bonds at the surface, which participate in the screening. To get some feeling about the accuracy of such a definition of the effective dielectric constant we have estimated $\epsilon_{\mathrm{in}}$ using other approaches. A quite natural approach is to calculate the dielectric response of the nanocrystals to a constant external field $E_{\mathrm{b}}$. Thus, the bare potential $V_{\mathrm{b}}$ is a linear function of the position. We then calculate the screened electric field $E$ on each atom using two methods. Firstly, we calculate the induced charges $[Q]$ on the atoms ($[Q] = [P]([V]-[V_{\mathrm{b}}])$, where $[P]$ is the polarizability matrix (Delerue *et al.* 1997), and we deduce the electric field on each atom induced by these point charges (excluding the charge on the atom itself); secondly, we estimate the electric field on each atom using a linear interpolation of the potential $V$ from the knowledge of the potential on the atom itself and on its four nearest neighbours. In both cases the electric field $E$ is close to constant, with some scattering around the average value. We calculate the average, $\bar{E}$, of the screened electric field over all the silicon atoms. We then deduce the effective dielectric constant using the macroscopic relation

$$\bar{E} = \frac{3}{2 + \epsilon_{\mathrm{in}}} E_{\mathrm{b}}. \tag{3.2}$$

Figure 7 shows that the values of $\epsilon_{\mathrm{in}}$ calculated in this way agree with the previous ones and with those of Wang & Zunger (1994*b*) performed in a pseudo-potential framework. This agreement gives us some confidence concerning the accuracy of the results.

It now remains to understand the origin of the variation of the effective dielectric constant with size of nanostructures. For that purpose, we consider the case of quantum wells, which are simpler systems to analyse because the confinement takes place in one direction only. We treat the case of $N$ silicon (100) planes to which we apply

a constant electric field $E_b$ perpendicular to the surface. Figure 8 shows that the concept of a constant local dielectric constant remains valid except in the vicinity of the surface. This value remains practically equal to the bulk one. Furthermore, the response remains extremely close to that obtained by applying $E_b$ to a layer of the bulk material with the same thickness and by using the bulk dielectric response $\epsilon(q)$.

All these results show that the concept of a local macroscopic dielectric constant remains valid even within very small systems. This dielectric constant keeps its bulk value except in the vicinity of the surface, where it decreases. This is consistent with the general fact that boundary conditions affect the electron density only over few Thomas–Fermi wavelengths. The average dielectric constant, on the other hand, is a combination of the two and exhibits a decrease with decreasing size of the nanostructure due to the surface contribution.

## References


Albrecht, S., Onida, G. & Reining, L. 1997 *Ab initio* calculation of the quasiparticle spectrum and excitonic effects in $Li_2O$. *Phys. Rev.* B **55**, 10 278.

Allen, P. B., Broughton, J. Q. & MacMahan, A. K. 1986 Transferable nonorthogonal tight-binding parameters for silicon. *Phys. Rev.* B **34**, 859.

Averin, D. V., Korotkov, A. N. & Likharev, K. K. 1991 Theory of single-electron charging of quantum wells and dots. *Phys. Rev.* B **44**, 6199.

Banin, U., Cao, Y., Katz, D. & Millo, O. 1999 Identification of atomic-like electronic states in indium arsenide nanocrystal quantum dots. *Nature* **400**, 542.

Delerue, C., Allan, G. & Lannoo, M. 1993 Theoretical aspects of the luminescence of porous silicon. *Phys. Rev.* B **48**, 11 024.

Delerue, C., Allan, G. & Lannoo, M. 1997 Calculations of the electron-energy-loss spectra of silicon nanostructures and porous silicon. *Phys. Rev.* B **56**, 15 306.

Delerue, C., Lannoo, M. & Allan, G. 2000 Excitonic and quasiparticle gaps in Si nanocrystals. *Phys. Rev. Lett.* **84**, 2457.

Delley, B. & Steigmeier, F. 1995 Size dependence of band gaps in silicon nanostructures. *Appl. Phys. Lett.* **67**, 2370.

Franceschetti, A., Wang, L. W. & Zunger, A. 1999 Comment on quantum confinement and optical gaps in Si nanocrystals. *Phys. Rev. Lett.* **83**, 1269.

Godby, R. W. & White, I. D. 1998 Density-relaxation part of the self-energy. *Phys. Rev. Lett.* **80**, 3161.

Hedin, L. & Lundqvist, S. 1969 *Solid state physics* (ed. H. Ehrenreich, S. Seitz & D. Turnbull), vol. 23, p. 1. Academic.

Hybertsen, M. S. & Louie, S. G. 1985 First-principles theory of quasiparticles: calculation of band gaps in semiconductors and insulators. *Phys. Rev. Lett.* **55**, 1418.

Jancu, J. M., Scholz, R., Beltram, F. & Bassani, F. 1998 Empirical $sp^3d^5s^*$ tight-binding calculation for cubic semiconductors: general method and material parameters. *Phys. Rev.* B **57**, 6493.

Lannoo, L., Delerue, C. & Allan, G. 1995 Screening in semiconductor nanocrystallites and its consequences for porous silicon. *Phys. Rev. Lett.* **74**, 3415.

Niquet, Y. M., Delerue, C., Allan, G. & Lannoo, M. 2000 Method for tight-binding parametrization: application to silicon nanostructures. *Phys. Rev.* B **62**, 5109.

Niquet, Y. M., Delerue, C., Allan, G. & Lannoo, M. 2002 *Phys. Rev.* B **65**, 165334.

Nozières, P. 1964 *Theory of interacting Fermi systems*. New York: Benjamin.

Ögüt, S., Chelikowsky, J. R. & Louie, S. G. 1997 Quantum confinement and optical gaps in Si nanocrystals. *Phys. Rev. Lett.* **79**, 1770.

Onida, G., Reining, L., Godby, R. W., DelSole, R. & Andreoni, W. 1995 *Ab initio* calculations of the quasiparticle and absorption spectra of clusters: the sodium tetramer. *Phys. Rev. Lett.* **75**, 818.

Porter, A. R., Towler, M. D. & Needs, R. J. 2001 Excitons in small hydrogenated Si clusters. *Phys. Rev.* B **64**, 035320.

Proot, J. P., Delerue, C. & Allan, G. 1992 Electronic structure and optical properties of silicon crystallites: application to porous silicon. *Appl. Phys. Lett.* **61**, 1948.

Ren, S. Y. 1997 Quantum confinement of edge states in Si crystallites. *Phys. Rev.* B **55**, 4665.

Rohlfing, M. & Louie, S. G. 1998 Excitonic effects and the optical absorption spectrum of hydrogenated Si clusters. *Phys. Rev. Lett.* **80**, 3320.

Sham, L. J. & Rice, T. M. 1966 Many-particle derivation of the effective-mass equation for the Wannier exciton. *Phys. Rev.* **144**, 708.

Slater, J. C. & Koster, G. F. 1954 Simplified LCAO method for the periodic potential problem. *Phys. Rev.* **94**, 1498.

Strinati, G. 1984 Effects of dynamical screening on resonances at inner-shell thresholds in semiconductors. *Phys. Rev.* B **29**, 5718.

Takagahara, T. & Takeda, K. 1992 Theory of the quantum confinement effect on excitons in quantum dots of indirect-gap material. *Phys. Rev.* B **46**, 15 578.

Tsu, R. & Babic, D. 1994 Doping of a quantum dot. *Appl. Phys. Lett.* **64**, 1806.

van Houten, H., Beenakker, C. W. J. & Staring, A. A. M. 1992 *Single charge tunneling: Coulomb blockade phenomena in nanostructures* (ed. H. Grabert & M. H. Devoret). New York: Plenum Press.

Wang, L. W. & Zunger, A. 1994*a* Electronic structure pseudopotential calculations of large (approx. 1000 atoms) Si quantum dots. *J. Chem. Phys.* **100**, 2394.

Wang, L. W. & Zunger, A. 1994*b* Dielectric constants of silicon quantum dots. *Phys. Rev. Lett.* **73**, 1039.

Zunger, A. & Wang, L. W. 1996 Theory of silicon nanostructures. *Appl. Surf. Sci.* **102**, 350.

## *Discussion*

A. M. Stoneham (*Department of Physics and Astronomy, University College London, UK*). I think you have used highly symmetric, non-polar clusters, in which the excited states are all delocalized. But Gavartin (see Gavartin & Stoneham 2003) showed lower symmetries, perhaps with electric fields, which could lead to localized excited states. How many of your results will carry over to these cases?

M. Lannoo. In the case of localized states induced by internal electric fields, the effect of the confinement is not the same. The calculations presented here do not take these effects into account, but they can be straightforwardly included in a Hartree-like approximation for example.

B. K. Ridley (*Optoelectronic Materials and Devices Laboratory, University of Essex, UK*). What exactly were the orbitals? Using $sp^3$ orbitals with basis $|s\rangle$, $|p_x\rangle$, $|p_y\rangle$, $|p_z\rangle$ turns out to mean that $x$, $y$, $z$ do not commute (Foreman 2002). Also, how does the interaction with phonons affect the energies?

M. Lannoo. We use a basis with one s and three p orbitals for each silicon atom and one s orbital for each hydrogen atom. This is a standard tight binding approach. Phonons have a minor effect on the confinement energies, but they can drastically influence the optical transition probabilities, in particular in the case of indirect gap materials.

*Additional references*

Gavartin, J. L. & Stoneham, A. M. 2003 Quantum dots as dynamical systems. *Phil. Trans. R. Soc. Lond.* A **360**, 275–290.

Foreman, B. F. 2002 *Phys. Rev.* B **66**, 165212.