Theory of one and two donors in silicon

A L Saraiva[1,2]**, A Baena**[2]**, M J Calderón**[3]** and Belita Koiller**[2]

[1] Department of Physics, University of Wisconsin-Madison, Madison, WI 53706, USA
[2] Instituto de Fisica, Universidade Federal do Rio de Janeiro, Caixa Postal 68528,
21941-972 Rio de Janeiro, Brazil
[3] Instituto de Ciencia de Materiales de Madrid, ICMM-CSIC, Cantoblanco, E-28049 Madrid, Spain

E-mail: also@if.ufrj.br

CrossMark

Abstract

We provide here a roadmap for modeling silicon nano-devices with one or two group V donors (D). We discuss systems containing one or two electrons, that is, D^0, D^-, D_2^+ and D_2^0 centers. The impact of different levels of approximation is discussed. The most accurate instances—for which we provide quantitative results—are within multivalley effective mass including the central cell correction and a configuration interaction account of the electron-electron correlations. We also derive insightful, yet less accurate, analytical approximations and discuss their validity and limitations—in particular, for a donor pair, we discuss the single orbital LCAO method, the Hückel approximation and the Hubbard model. Finally, we connect these results with recent experiments on devices with few dopants.

Keywords: donor, quantum electronics, donor molecule, effective mass theory, central cell correction, semiconductor

(Some figures may appear in colour only in the online journal)

1. Introduction

In a letter to Peierls in 1931, Pauli stated that semiconductors '*are a filthy mess*'[4]. Curiously, the impurities and defects to which Pauli refers led the way to the revolution which severely impacted upon industry and society over the last many decades, and is still under way. Development and advances in semiconductor-based devices started less than 20 years after Pauli's statement, as the transistor operated for the first time in 1947. This achievement prompted intensive research for a deeper understanding of semiconductors and the role of dopants. Currently, as 'Moore's Law' approaches its limit with devices reaching the atomic scale, quantum behavior of electrons and spins drive the advancement of the semiconductor research into the new fields of quantum electronics and spintronics [1–3].

Doping allows controlling the sign and density of carriers in a semiconductor; a flexibility not achievable in conductors or in insulators. For the applications where current carriers perform the needed operations, donor electrons are promoted to the nearby conduction band and similarly holes from acceptors operate in the valence band. In this case the concentration of donors and acceptors and their binding energies are the only relevant quantities defining the behavior of the doped material. A macroscopic concentration of dopants leads to macroscopically observable currents or voltages, and at this point the quantum behavior becomes irrelevant.

We revisit here the theoretical treatment of donors D in semiconductors from a single substitutional donor perspective, considering that its active electron remains bound to the core, and similarly for donor pairs D_2, a scenario where full quantum behavior prevails. Special attention must be paid to the electronic bound state wavefunctions.

Our approach incorporates the multivalley structure of the Si conduction band and a central cell correction within effective mass theory. Taking for simplicity a spherical envelope approximation, we show that central cell corrections can be included involving a single donor-dependent empirical parameter related to the range of the correction potential. This range is chosen such that the neutral donor ground state experimental energy is reproduced for each species. We verify that our central cell correction is transferable to negative ions and diatomic molecules correctly reproducing reported experimental bindings. Additional experimental

information on the single donor first excited states E and T_2 allows estimating these energies within our spherical model, providing another consistency check.

This paper is organized as follows. In section 2 we briefly review the effective mass theory (EMT) for shallow donors, highlighting its successes and limitations. In section 3, multivalley and anisotropy effects are analysed, details of the approach used are given, and the results for neutral donors are shown in section 4. Section 5 reports on a donor with two bound electrons, constituting a D$^-$ negatively charged donor. The donor pair is studied in sections 6 and 7 for the one and two electron states respectively. Section 8 discusses simplified models for the two donor problem and makes connections with experiments. Finally, summary and conclusions are presented in section 9. In order to keep the text reasonably self-contained, well established basic material and results are included with proper references provided.

2. Hydrogenic model for donors in semiconductors

The single donor description is one of the simplest and most successful implementations of the EMT. It is based on the usual effective-medium assumptions, namely that:

(1) The perturbation potential due to a substitutional donor varies slowly at the scale of the lattice constant of the host semiconductor;

(2) The envelope functions of the bound states extend over long enough distances such that it is composed by a very narrow distribution of Bloch wavevectors **k** around the band minimum.

Further assuming that the conduction band lower edge is non-degenerate and isotropic, one obtains the hydrogenic atom hamiltonian

$$\mathcal{H}_H = -\frac{\hbar^2 \nabla^2}{2m^*} - \frac{e^2}{4\pi\epsilon r} \qquad (1)$$

with the electron mass substituted by the effective mass m^* and the Coulomb potential screened by the static dielectric constant ϵ. This gives a hydrogenic atom of effective Bohr radius $a^* = a_0 \epsilon/m^*$ and ground state binding energy $E^* = E_H m^*/\epsilon^2$, with $a_0 = 0.053$ nm and $E_H = 13.6$ eV the respective values for Hydrogen in vacuum. The wavefunction for the ground state electron is then the product of the hydrogenic envelope-function by the band-edge Bloch function

$$\Psi = \frac{1}{\sqrt{\pi a^{*3}}} e^{-r/a^*} e^{i\mathbf{k}\cdot\mathbf{r}} u_\mathbf{k}(\mathbf{r}). \qquad (2)$$

The EMT provides a highly accurate description of shallow donors in single valley materials. It involves no fitting parameters, as the binding energy is dependent on the material only through ϵ and m^*, leading to a value of E^* that is independent of the donor species. This is in fact clearly demonstrated experimentally for example in GaAs, where deviations from the hydrogenic expression are hard to detect [5], as illustrated in figure 1. In contrast, the binding energy of donors in Si and Ge is strongly dependent on the donor species [4].
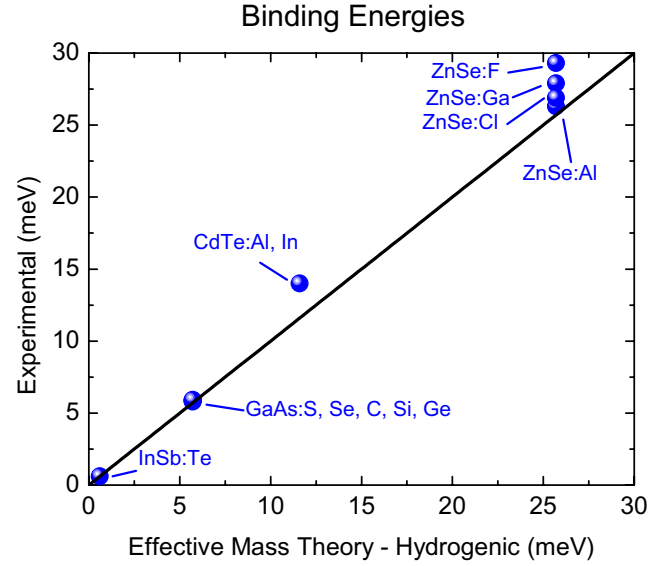

Figure 1. Comparison of the experimental donor binding energies for single valley materials (blue dots) with those calculated from the effective mass hydrogenic expression (black line). The agreement is very good mainly for the less bound donors. The ground state energies for donors in silicon (not shown here) are off the given scale and present a large variation ($\sim$50%) as a function of the donor species, see table 1. The same spread of values is found for germanium [4].

3. Multivalley effects

The case of multivalley semiconductors is not as simple. In a recent work [6], spatially resolved spectroscopy of isolated As donors in silicon was combined with multiband semi-empirical tight binding theory to study the valley interference. These results provide direct evidence for the rich valley structure predicted over 60 years ago by Kohn and Luttinger [7] (KL) and further outline the role of the environment on valley repopulation.

With the N minima (valleys) at finite wavevectors, the symmetry of the wavefunction is significantly lower, and highly anisotropic effective masses are not uncommon—in silicon the longitudinal and transverse effective masses are $m_L = 0.916 m_e$ and $m_T = 0.191 m_e$ at the $N = 6$ minima. The hydrogenic hamiltonian equation (1) may be adapted to account for the mass anisotropy, as discussed in section 3.1.

The mere reconciliation with the mass anisotropy still does not lead to results compatible with the experimental binding energies. The missing ingredient is the elusive valley-orbit coupling. In the absence of valley-orbit coupling, a wavefunction equivalent to equation (2) is obtained for each of the N valleys, leading to an N-fold degenerate ground state.

The $1/r$ Coulomb potential singularity at $r = 0$ is clearly incompatible with the assumptions discussed in section 2—it actually mediates a finite coupling among the N valleys. This accounts for over 30% of the binding energy in silicon donors, as will be discussed in section 3.2. We briefly present some of the most successful semi-empirical approaches to describe these departures from the hydrogenic model, namely the first order perturbation intervalley coupling (section 3.3) and the central cell correction (section 3.4).

3.1. Uncoupled anisotropic valleys

The first successful approach to treat donors in Si was presented by KL [7] (which also set a stronger formal basis for the EMT) and Kittel and Michel [8]. Taking initially the EMT assumption (2) to be valid—namely that only small deviations from the wavevector at the band minimum $\mathbf{k}_{\min}$ contribute to the wavefunction—means that any individual valley does not couple to the other five. In this way each valley is solved independently and the problem regains the hydrogenic simplicity.

The band anisotropy leads to a specific effective mass hamiltonian for the different valleys. For instance, for the 2 valleys along z, it reads

$$\mathcal{H}_z = -\frac{\hbar^2}{2m_T}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) - \frac{\hbar^2}{2m_L}\frac{\partial^2}{\partial z^2} + V(r). \quad (3)$$

The ground state is found variationally, choosing a trial function for the $\mathbf{k}_z = 0.85\frac{2\pi}{a_{\text{Si}}}(0\,0\,1)$ valley (a_{Si} is the Si conventional lattice parameter) of the shape of a deformed 1s orbital[5] with two radii a and b

$$\phi_z(\mathbf{r}) = \frac{1}{\sqrt{\pi a^2 b}} e^{-\sqrt{(x^2+y^2)/a^2+z^2/b^2}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}). \quad (4)$$

For other valleys at $\{\mathbf{k}_\mu\}$, $\mu = \{x, -x, y, -y, z, -z\}$, the hamiltonians $\mathcal{H}_\mu$ and the wavefunctions ϕ_μ are immediately obtained exchanging z by x or y accordingly. The deformed orbital Bohr radii a and b are taken to minimize the expectation value of $\langle\phi_\mu|\mathcal{H}_\mu|\phi_\mu\rangle$. For the hydrogenic potential $V(r) = V_H(r) = -\mathrm{e}^2/4\pi\epsilon_{\text{Si}}r$, the ground state is sixfold degenerate, and one recovers the KL result with energy $E_{\text{KL}} = -31.2$ meV, and an anisotropy $b/a = 0.56$.

3.2. Valley-orbit coupling

The binding energy is significantly underestimated if the valley-orbit coupling is disregarded, as can be seen by comparison of E_{KL} to the observed values for shallow donors in Si in table 1. The valley-orbit coupling breaks the sixfold degeneracy, leading to a significant reduction of the ground state energy. Although the degeneracy reduction is evident by symmetry, a quantitative estimate of its value is challenging, since it depends strongly on the potential near the singularity, where departures from the screened point-charge behavior are expected. Even from first principles, the widely adopted pseudopotential approach provides very inaccurate estimates for the valley-orbit coupling and perturbative corrections are needed to obtain quantitative agreement with experimental data [9].

Even if the strength of this coupling is not known, group theory arguments give the appropriate superposition of valley states induced when the spherically symmetric potential is disrupted by the tetrahedral crystal field potential. In other words, the irreducible representation associated with the ground state of a spherical potential with six-fold

---

5 For the single donor problem it is convenient to take the donor position as the coordinate origin $\mathbf{r} = 0$.

valley degeneracy (the 1s manifold) becomes reducible in the presence of a local tetrahedral crystal field: the sixfold degenerate level splits into states that have the symmetry of the different irreducible representations of the T_d group. In this case, this leads to a singlet with A_1 symmetry, a triplet with T_2 symmetry and a doublet with E symmetry. Each of the six states in the 1s manifold correspond to a particular combination of the envelope-modulated Bloch functions from the six degenerate valleys ϕ_μ illustrated in figure 2. These states are given explicitly by [10].

$$\Psi_{A_1}(\mathbf{r}) = \frac{1}{\sqrt{6}}\big[\phi_x(\mathbf{r}) + \phi_{-x}(\mathbf{r}) + \phi_y(\mathbf{r}) + \phi_{-y}(\mathbf{r})$$
$$+ \phi_z(\mathbf{r}) + \phi_{-z}(\mathbf{r})\big]$$

$$\Psi_{T_2^x}(\mathbf{r}) = \frac{1}{\sqrt{2}}\big[\phi_x(\mathbf{r}) - \phi_{-x}(\mathbf{r})\big]$$

$$\Psi_{T_2^y}(\mathbf{r}) = \frac{1}{\sqrt{2}}\big[\phi_y(\mathbf{r}) - \phi_{-y}(\mathbf{r})\big]$$

$$\Psi_{T_2^z}(\mathbf{r}) = \frac{1}{\sqrt{2}}\big[\phi_z(\mathbf{r}) - \phi_{-z}(\mathbf{r})\big]$$

$$\Psi_{E^z}(\mathbf{r}) = \frac{1}{\sqrt{12}}\big[\phi_x(\mathbf{r}) + \phi_{-x}(\mathbf{r}) + \phi_y(\mathbf{r}) + \phi_{-y}(\mathbf{r})$$
$$- 2\phi_z(\mathbf{r}) - 2\phi_{-z}(\mathbf{r})\big]$$

$$\Psi_{E^{xy}}(\mathbf{r}) = \frac{1}{2}\big[\phi_x(\mathbf{r}) + \phi_{-x}(\mathbf{r}) - \phi_y(\mathbf{r}) - \phi_{-y}(\mathbf{r})\big]. \quad (5)$$

The direct calculation of the valley-orbit coupling is not trivial, though. We are unaware of any successful description of the complete spectrum of the 1s manifold from first principles. This difficulty is due to the incomplete knowledge regarding the impurity potential. The picture of a $1/r$ point charge potential screened by a dielectric constant ϵ_{Si} is realistic only at large distances from the impurity site. Close to the nucleus, the dielectric response becomes inhomogeneous. Furthermore, the valence core shells for the impurity are different from the Si atoms, presenting a strong departure from the excess proton picture. Atomic species differ from each other at the core region, leading to the strong donor species' dependence of the binding energy in silicon in contrast with other semiconductors. Early attempts of correcting the Coulomb potential adopting the *ab initio* dielectric function were modestly successful for the isovalent impurity P and were inaccurate for all other group V donors (for a complete discussion of early works, see [11]). Successful attempts involve empirical fitting ingredients [12–14]. Some examples used in the literature to describe this departure from the point charge picture are shown in figure 3. Two among the main strategies for describing the valley-orbit induced splitting are described next.

3.3. First order perturbation theory: the intervalley coupling

The EMT assumptions (1) and (2) are clearly incompatible with the Coulomb potential, which is singular at the donor site $r = 0$, implying that different valley states may couple. The single-valley KL description neglects this singularity, i.e. $\mathcal{H}$ defined in equation (3) does not include the singular part of the Coulomb potential V_H.
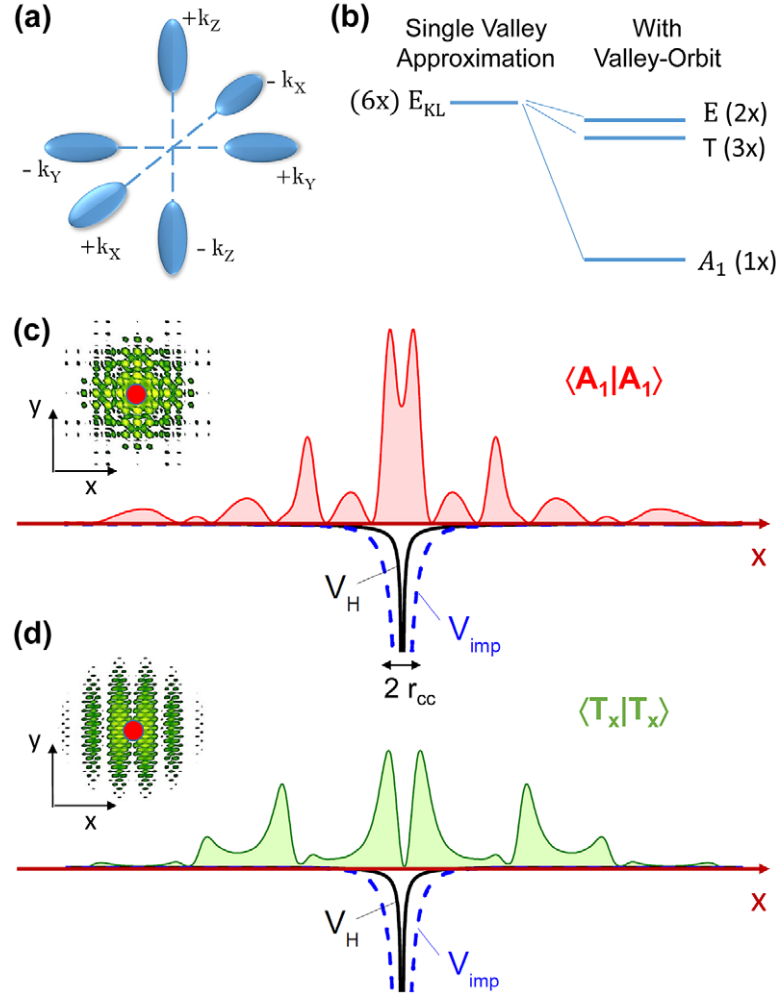
Figure 2. (*a*) Representation in k-space of the six constant energy surfaces around the degenerate conduction band minima in Si. (*b*) Splitting of the sixfold degenerate conduction band minimum by a tetrahedral crystal environment combined with the singular donor potential. Charge distribution for the (*c*) A_1 and (*d*) T_x states along the *x*-axis (large) and at the (0 0 1) Si crystal plane (inset). Note that A_1 has significant charge density at the donor site, while the T_x state has a node. The lower lines curving down give the plain screened Coulomb potential V_H (solid line) and the central-cell corrected potential V_{imp} (dashed line) which is more attractive for distances from the donor nucleus smaller than $\sim r_{cc}$. Due to the larger charge density of the A_1 state within r_{cc}, its energy is well below the T_2 and E states.

We may then include the intervalley coupling as a perturbation to the KL description. The total Hamiltonian would be $H = \mathcal{H} + H'$ with H' accounting for the singular part of the potential around the core region $r \to 0$, i.e. $V_H(r \to 0)$ plus additional corrections as discussed below. For simplicity in notation we are assuming an isotropic approximation (further discussed in section 4) which implies that $\mathcal{H}_z = \mathcal{H}_x = \mathcal{H}_y = \mathcal{H}$. Now, the single-valley approximation implies $\langle k_\mu | \mathcal{H} | k_\nu \rangle = E_{KL}\delta_{\mu,\nu}$, where $|k_\mu\rangle$ is the ket representing ϕ_μ. The valley-coupling part, H', is added and treated following standard degenerate perturbation theory, i.e. we solve the perturbed Hamiltonian $H = \mathcal{H} + H'$ restricted to the degenerate manifold basis $\{\phi_\mu\}$.

It is easy to see by symmetry that $\langle k_\mu | H' | k_\mu \rangle$ is a negative constant (independent of μ) which we call $(-\delta)$, with this the matrix elements of the hamiltonian written in the $\{\phi_\mu\}$ basis may be obtained

$$\langle k_\mu | H | k_\mu \rangle = \langle k_\mu | \mathcal{H} | k_\mu \rangle + \langle k_\mu | H' | k_\mu \rangle = E_{KL} - \delta,$$

$$\langle k_\mu | H | k_{-\mu} \rangle = \langle k_\mu | H' | k_{-\mu} \rangle = \Delta_\parallel,$$

$$\langle k_\mu | H | k_\nu \rangle = \langle k_\mu | H' | k_\nu \rangle = \Delta_\perp, \text{ if } \nu \neq \pm\mu. \quad (6)$$

Direct calculation of the eigenvectors of the hamiltonian matrix yield a spectrum split in a singlet A_1 at energy $E_{KL} - \delta + \Delta_\parallel + 4\Delta_\perp$, a triplet T_2 at $E_{KL} - \delta - \Delta_\parallel$, and a doublet E at $E_{KL} - \delta + \Delta_\parallel - 2\Delta_\perp$.

A possible semi-empirical solution would be to treat the three matrix elements δ, $\Delta_\parallel$ and $\Delta_\perp$ as parameters chosen to reproduce the 1s manifold spectrum of each donor species. In [15, 16] the values $\Delta_\parallel = -1.52$ meV and $\Delta_\perp = -2.16$ meV are suggested for P donors in Si. This choice reproduces the relative energy splitting among the different 1s manifold levels. An overall negative shift of $\delta = 4.22$ meV for P is required in order to get the actual experimental energies. The small values of $\Delta_\parallel$, $\Delta_\perp$ and δ with respect to E_{KL} justify the perturbative treatment described here.

The observed experimental energies can also be understood from central cell corrections, which take into account the incomplete screening of the positive donor potential, as detailed in section 3.4. The adopted central cell correction involves a single, instead of three, empirical parameter characterizing each donor species.
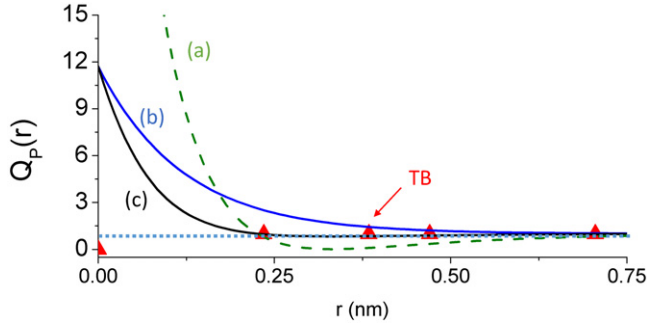
Figure 3. Comparison between different models adopted in the literature for the central cell correction. The figure gives $Q_P(r)$ such that $V_P(r) = -Q_P(r)V_H(r)$ calculated for P in Si. The different curves correspond to (a) [14], (b) $V_{imp}(r)$ as in equation (7), (c) [12], and the triangles give the approximation used in tight-binding [13]. The limit of a point charge $Q_P = 1$ is shown as a dotted horizontal line. Note that, for the tight-binding, $Q_P(r) = 1$ for all sites except the impurity site where V_{imp} is taken as a constant. In this case we indicate $Q_P(r) = 0$ in the figure, which eliminates the $1/r$ divergence at $r \to 0$, but a constant U_0 should be added.

3.4. Effective pseudopotential: the central cell correction

Another strategy to describe the donors beyond the single valley theory is to build a pseudopotential $V_{imp}(r)$ that effectively mimics the complex environment of the crystal and the screening due to the valence electrons of the donor. This approach received some attention lately [17, 18]. The suffix imp stands for the impurity type, so imp=P, As, Sb or Bi. Multivalley semiconductors are more strongly subject to this inner core potential than single valley semiconductors because the valley degree of freedom allows for the electron to concentrate closer to the nucleus. The rationale for that is simple—each wavefunction ϕ_μ spreads in k-space around $\mathbf{k}_\mu$ in a form determined by the Heisenberg uncertainty principle. The spread of the superposition of various ϕ_μ in k-space tends to be larger than on the non-degenerate case, so that the real space spread may be smaller without costing any kinetic energy. Thus, multivalley electronic states tend to concentrate around the nucleus, benefiting from its potential energy and therefore lowering its total energy. A consequence is that a larger portion of the ground state energy comes from the potential very close to the nucleus, which is precisely the region where the point charge model fails.

Formally the coupling between valleys, $\langle k_\mu | H' | k_\nu \rangle$, could be obtained directly from the perturbation due to the donor breaking the effective mass assumptions, modeled above by the term H'. The full donor potential, $V_{imp}(r)$, has been treated in the literature within several models, a few of which are given in figure 3. For example, a constant replacement for the screened Coulomb potential within a central cell region of radius R has been adopted in [19], namely, $V_{imp}(r < R) = V_0$, $V_{imp}(r > R) = V_H(r)$, with R taken as half the nearest neighbors' distance in Si and V_0 is a negative energy chosen to reproduce the experimental spectra. A conceptually similar scheme, widely adopted in tight-binding (TB) calculations, is presented in [20] for GaAs and in [13, 21] for Si. It consists of taking the usual screened Coulomb potential $\sim 1/r_i$ to correct the on-site energies at atomic positions $\mathbf{r}_i \neq 0$. For the

Table 1. Experimentally measured [23, 24] values E_{exp}^0 of the 1s manifold energy states for the neutral isolated donor are compared to the calculated values E_{calc}^0 obtained from the central cell corrected potential equation (7). The energy of the central cell corrected state A_1 is fitted to the corresponding experimental value by choosing an appropriate r_{cc}. The E and T_2 levels are assumed degenerate. The calculation of the levels E and T_2 involves $\tilde{E}$, the center of gravity of the sixfold 1s manifold, which is taken from experiment (see text). Note that the experimental and calculated values for the T_2 and E energies do not depend strongly on the donor species and are very similar to the single valley result E_{KL}, in contrast to the value of the ground state A_1 energy. For each donor, the effective Bohr radius corrected by the central cell a_{cc} is also given.

Donor	E_{exp}^0 (meV) [23, 24]	E_{calc}^0 (meV)	r_{cc} (nm)	a_{cc} (nm)
P^0	−45.58 (A_1)	E_{exp}^0	0.115	1.106
	−33.90 (T_2)	−33.12		
	−32.60 (E)	−33.12		
As^0	−53.77 (A_1)	E_{exp}^0	0.126	0.815
	−32.65 (T_2)	−32.15		
	−31.40 (E)	−32.15		
Sb^0	−42.71 (A_1)	E_{exp}^0	0.108	1.241
	−33.00 (T_2)	−32.04		
	−30.60 (E)	−32.04		
Bi^0	−70.88 (A_1)	E_{exp}^0	0.144	0.58
	−32.48 (T_2)	−30.56		
	−31.38 (E)	−30.56		

impurity site the Coulomb correction would diverge, which is incompatible with TB, so a finite correction U_0 is chosen to adjust the correct binding energies of each donor.

Realistic descriptions of the donor Coulomb potential should be consistent with the expected limits: $V_{imp}(r \to \infty) = -e^2/4\pi\epsilon_{Si}r = V_H(r)$ and $V_{imp}(r \to 0) = -e^2/4\pi\epsilon_0 r$. A convenient interpolation [22] is adopted here, see figure 3, namely

$$V_{imp}(r) = -\frac{e^2}{4\pi r}\left[\frac{1}{\epsilon_{Si}} + \left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon_{Si}}\right)e^{-r/r_{cc}}\right]$$
$$= V_H(r) + V_{cc}(r). \tag{7}$$

This involves a single donor-dependent parameter, r_{cc}, giving the range around the donor site where the central cell correction is effective, as opposed to the three quantities required for the description given in equation (6). Values for r_{cc} appropriate for P, As, Sb and Bi donors in Si are given in table 1. Note that while Bohr radii are of the order of a few nm, r_{cc} is a factor of 10 smaller. Therefore the deviation of the plain Coulomb potential from the central-cell corrected one is restricted to a small region around the donor, as illustrated in figure 2(c).

4. Neutral donor: D^0 center

The coupling matrix elements calculation may be simplified by assuming spherically symmetric envelopes. It is possible to conciliate the simpler hydrogenic description in equation (2) with the ground state energy E_{KL} obtained from equation (4) by choosing a value of the effective mass $m_{Si}^* = 0.3\, m_e$. Note that m_{Si}^* is very close to the geometric mean $(m_L \times m_T^2)^{1/3} = 0.32$. This approximation results, for the uncoupled valleys, into a sixfold degenerate ground state with ground state energy E_{KL}.

Now we take into account the multivalley structure of the Si conduction band and the central cell correction described in equation (7). The only parameter we are left with is r_{cc}, the central cell potential characteristic length, chosen such that the ground state A_1 experimental energy is reproduced for each of the donor species, as given in table 1.

The T_2 and E states are much less sensitive to the effect of V_{cc} since both have nodes at the impurity site. This is confirmed by their approximate degeneracy and energy very close to E_{KL} (see table 1). We take $\Delta_\parallel = \Delta_\perp = \Delta$, so that $E(A_1) = E_{KL} - \delta + 5\Delta$ and $E(T_2, E) = E_{KL} - \delta + \Delta$. This equality is a direct consequence of the spherical model. The mean value of the six levels is taken from experiments and identified with $\tilde{E} = E_{KL} - \delta$ which leads to $\delta = E_{KL} - \tilde{E}$. It then follows that $\Delta = [E_{KL} - E(A_1)]/5$ allowing to determine $E(T_2, E)$.

Combining spherically symmetric envelopes with the Bloch functions obtained from the expansion coefficients given in [25] yield the electronic densities for the A_1 and T_2^x bound donor states of P in Si shown in figures 2(c) and (d). Note that the charge distribution of the T_2^x state has a preferential alignment along the x direction and an exact node at the origin (*idem* to E states). Due to this node, any contact interaction (interaction potentials that are only finite at the immediate vicinity of the donor impurity) will have a reduced effect on the T_2 and E states—this justifies why these states are not strongly affected by the central cell correction as well as indicates that low hyperfine coupling is to be expected in these states. For this reason, the T_2 and E states energies are very similar to E_{KL}. The results are summarized in table 1.

5. Two electrons in one donor: D$^-$-center

In analogy with the hydrogen atom, singly ionized negative (H$^-$) or positive (H$^+$) states are also of practical and theoretical interest. The negative ion H$^-$ binding energy is defined as the energy required to remove one electron from the ion $E_B^{H^-} = E_{H^0} - E_{H^-}$. For hydrogen $E_B^{H^-} = 0.055 E_H$. The analogous D$^-$ center is specially important in Si nanoelectronics because its binding energy is very small and can be strongly affected by gates [26]. The presence of this bound state has been identified in transport experiments in single-atom transistors [27].

We may also regard D$^-$ centers as He-like systems [28], the main technical difficulty to obtain the spectrum consists in the evaluation of electron–electron repulsion terms. Variational wavefunctions with a large number of parameters, as used by Hylleraas [29] and others, are able to reproduce the experimental value of $E_B^{H^-}$.

In general, the ground state wavefunction is composed of an inner orbital with a Bohr radius similar to the electron bound to a neutral donor, and an outer orbital with a much larger radius, due to the screening of the Coulomb potential produced by the occupied inner orbital.

Within a single valley and isotropic mass approach for negatively ionized centers in Si, we can just rescale the exact results for H$^-$: $E_B^{D^-} = 0.055 E^* \sim 1.7$ meV, which gives a good estimate for phosphorus and arsenic

centers: $E_B^{P^-} = 1.7$ meV, $E_B^{As^-} = 2.05$ meV as measured in photoconductivity experiments [30].

The inclusion of the mass anisotropy and the valley degeneracy in Si would increase the binding energy of the D$^-$ centers even if valley-orbit interactions are not considered [31]. If no central cell corrections are included, each electron remains in a different valley. The lowest energy configuration is attained when the two electrons are located in perpendicular valleys (as $m_\parallel$ and $m_\perp$ are rotated in perpendicular valleys). Mass anisotropy then implies that the envelopes corresponding to the two orbitals do not overlap as much as they would for isotropic wavefunctions reducing the electron–electron repulsive interaction.

Valley-orbit interactions and central-cell corrections have been treated in the literature for the D$^-$ problem under different approximations [19, 32, 33]. Larsen [32] assumed that only the inner orbital electron is subject to valley-orbit coupling, behaving like a neutral donor electron. The isotropic condition $\Delta_\parallel = \Delta_\perp$ was also used. The effect of the valley-orbit coupling on the inner electron is to spread it in a symmetric (A_1) combination of valleys. If both electrons have the same valley composition, the binding energy decreases with respect to them being in perpendicular valleys. Oliveira and Falicov [19] went a step further by including the full 1s multiplet in the description of the bound electrons and a constant valley-orbit coupling different from zero only in a central-cell region. More recent calculations [33] included the valley degeneracy but not the valley-orbit splitting.

Here we treat the D$^-$ donors on the same level of approximation as the neutral donors, hence neglecting the mass anisotropy. We take the simplest possible form for the spin singlet ground state trial function (normalized):

$$\psi_{D^-} = \frac{1}{\pi\sqrt{2a_1^3 a_2^3}} \left(e^{-r_1/a_1} e^{-r_2/a_2} + e^{-r_2/a_1} e^{-r_1/a_2} \right), \quad (8)$$

and consider the central-cell corrected potential $V_{imp}(r)$, as defined in equation (7). The correction term V_{cc} is completely determined by the value of r_{cc} chosen from the corresponding neutral donor data (see table 1). Both a_1 and a_2 are calculated variationally to minimize the expectation value of the energy for the two-electrons hamiltonian: $H_{D^-} = K_1 + K_2 + V_{imp}(r_1) + V_{imp}(r_2) + e^2/\epsilon_{Si} r_{12}$ where K_i is the kinetic energy of electron i and $V_{imp}(r_i)$ is the central cell corrected potential. The last term is the electron–electron repulsion. The energies are calculated following the prescription for He atoms given in [28].

If central cell corrections were ignored, the variational wavefunction in equation (8) would lead to a binding energy $E_B^{D^-} = 0.027 E^* = 0.8$ meV, see [26]. The inclusion of the central cell correction allows for a significant improvement of the binding energy, see table 2. The Table also shows the value of the charging energy $U = E_{H^-} - 2E_{H^0}$ which can be measured in transport spectroscopy experiments in single atom transistors [26].

Note the very good agreement between calculated and experimental (when available) values in table 2, indicating the quantitative validity and transferability of the central cell correction given here and validating our predictions for Sb and Bi.

Table 2. Calculated and experimental energies and variational parameters for the negatively charged donors. All energies are given in meV and lengths in nm. E_B refers to the first ionization energy, namely, the energy required to ionize the less bound electron $E_B = E^0 - E^-$. The charging energy U is defined as $U = E^- - 2E^0$. The parameters a_1 and a_2 are the Bohr radii obtained variationally. The radius a_1 corresponds to the inner orbital, and is very similar to a_{cc} in table 1. The parameter a_2 is the radius of the outer orbital which is much less bound than the inner one. To the best of our knowledge, the experimental values for Sb and Bi are not known.

Donor	E_{calc}^-	E_{exp}^-	$E_{B(\text{calc})}$	$E_{B(\text{exp})}$ [30]	U_{calc}	U_{exp}	a_1	a_2
P$^-$	−47.10	−47.29	1.526	1.73	43.86	43.85	1.041	4.851
As$^-$	−56.44	−55.81	2.69	2.05	51.06	51.72	0.737	3.516
Sb$^-$	−44.00	N/A	1.29	N/A	41.44	N/A	1.177	5.349
Bi$^-$	−78.55	N/A	7.55	N/A	63.45	N/A	0.48	2.16

6. One electron and two donors: D$_2^+$ center

The problem of two donors sharing a single electron is analogous to the problem of an ionized H$_2^+$ molecule in vacuum. If it was identical it would be possible to solve this problem exactly through the transformation into spheroidal coordinates [34]. However, as discussed for single donors, the effective mass anisotropy and the valley orbit coupling in silicon complicate the spectrum. Many accounts of these effects are available in the literature (see [18] and references therein). Early approaches employed an altered form of the spheroidal coordinates to accommodate mass anisotropy to some extent [35].

We steer the analogy with the H$_2^+$ molecule along another direction, namely writing the eigenstates of this problem as bonding and anti-bonding molecular orbitals based on the single donor (atomic) problem, i.e. taking molecular wavefunctions as linear combination of atomic orbitals (LCAO). The two donors, referred to as A and B, located at positions $\mathbf{r_A}$ and $\mathbf{r_B}$ are separated by a vector $\mathbf{R} = \mathbf{r_B} - \mathbf{r_A}$. An obvious difference between the substitutional donor pair and the molecular problems is the meaning and range of possible values of $\mathbf{R}$. For donors, $\mathbf{R}$ is fixed at the sample fabrication and doping stage and coincides with a lattice vector, while in the free molecule situation, $R = |\mathbf{R}|$ is given by the minimization of the molecule energy.

The effective mass hamiltonian of a donor pair reads

$$H_{\text{DD}}(\mathbf{r}) = -\frac{\hbar^2 \nabla^2}{2m^*} + V_{\text{imp}}(r_A) + V_{\text{imp}}(r_B). \tag{9}$$

The matrix elements of this hamiltonian are calculated in the LCAO basis, i.e. in the basis set defined by the single donor wavefunctions $\{A_1, T_2, E\}$ given in equation (5), centered at $\mathbf{r_A}$ and $\mathbf{r_B}$. We expect this basis to give a fair account of the lowest electronic states. The hamiltonian is hence a 12×12 matrix, which for convenience we break into four 6×6 blocks,

$$H = \begin{bmatrix} H_{AA} & H_{AB} \\ H_{BA} & H_{BB} \end{bmatrix}. \tag{10}$$

Blocks H_{AA} and H_{BB} are not strictly diagonal since the problem of two donors has lower symmetry than a single donor. But a direct calculation shows that the off-diagonal terms in these 2 blocks are vanishingly small compared to other non-zero terms for interdonor distances larger than the lattice constant a_{Si}. We therefore take the diagonal blocks to be diagonal, which shows that the donor eigenstates basis

set is more convenient than the $\{\phi_\mu\}$ for the pair treatment. These blocks, therefore, represent the on-site energies while the tunnel coupling terms contribute to the blocks H_{AB} and H_{BA}. The diagonal blocks $[H_{AA}] = [H_{BB}]$ considering a basis ordered as in the sequence in equation (5):

$$[H_{ii}] = \begin{pmatrix} \varepsilon_{\text{CC}}^{\text{onsite}} & 0 & 0 & 0 & 0 & 0 \\ 0 & \varepsilon_{\text{sv}}^{\text{onsite}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \varepsilon_{\text{sv}}^{\text{onsite}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \varepsilon_{\text{sv}}^{\text{onsite}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \varepsilon_{\text{sv}}^{\text{onsite}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \varepsilon_{\text{sv}}^{\text{onsite}} \end{pmatrix}. \tag{11}$$

The onsite energy for the A_1 state $\varepsilon_{\text{CC}}^{\text{onsite}}$ is the ground state energy for the neutral donor including the central cell contribution from the A donor, corrected by a long range classical term $V_{\text{imp}}' = \langle \Psi_{A_1}(r_A)|V_{\text{H}}(r_B) + V_{\text{CC}}(r_B)|\Psi_{A_1}(r_A)\rangle$ to take into account the B donor potential. For T_2 and E, $\varepsilon_{\text{sv}}^{\text{onsite}}$ is the single valley E_{KL} corrected by $V_{\text{H}}' = \langle \Psi_{T/E}(r_A)|V_{\text{H}}(r_B)|\Psi_{T/E}(r_A)\rangle$. Following the argument in section 4, the central-cell correction on T_2 and E is neglected because these states have nodes at all Si sites.

The off-diagonal blocks are related by the hermiticity condition $H_{AB} = H_{BA}^\dagger$. Each term is a summation over integrals of the type

$$\langle \phi_\mu(\mathbf{r_A})|H_{\text{DD}}|\phi_\nu(\mathbf{r_B})\rangle$$
$$= \int F(r_A) e^{-i\mathbf{k}_\mu \cdot \mathbf{r_A}} u_\mu^*(\mathbf{r_A}) H_{\text{DD}} F(r_B) e^{i\mathbf{k}_\nu \cdot \mathbf{r_B}} u_\nu(\mathbf{r_B}) d^3\mathbf{r}. \tag{12}$$

These matrix elements can be straightforwardly calculated using the plane wave expansion of the Bloch functions (see [25]), and no further approximation is needed. Nevertheless, it is useful to simplify these integrals through some well tested approximations: (i) Neglecting the matrix elements for $\mu \neq \nu$, since rapidly oscillatory integrands are involved; (ii) Taking $u_\mu^* u_\mu \approx 1$, as suggested in [25]. Under these assumptions we obtain

$$\langle \tilde{\phi}_\mu^{A,i}|H_{\text{DD}}(\mathbf{r})|\tilde{\phi}_\nu^{B,j}\rangle = \delta_{\mu,\nu} e^{i\mathbf{k}_\mu \cdot \mathbf{R}} t_{\text{sv}}(R), \tag{13}$$

where t_{sv} is the single valley tunnel coupling

$$t_{\text{sv}}(R) = \int F(r_A) H_{\text{DD}} F(r_B) d^3\mathbf{r}. \tag{14}$$

The effect of the hopping blocks H_{AB} is analogous to the H$_2$ molecule—at distances much larger than a_{cc}, this block
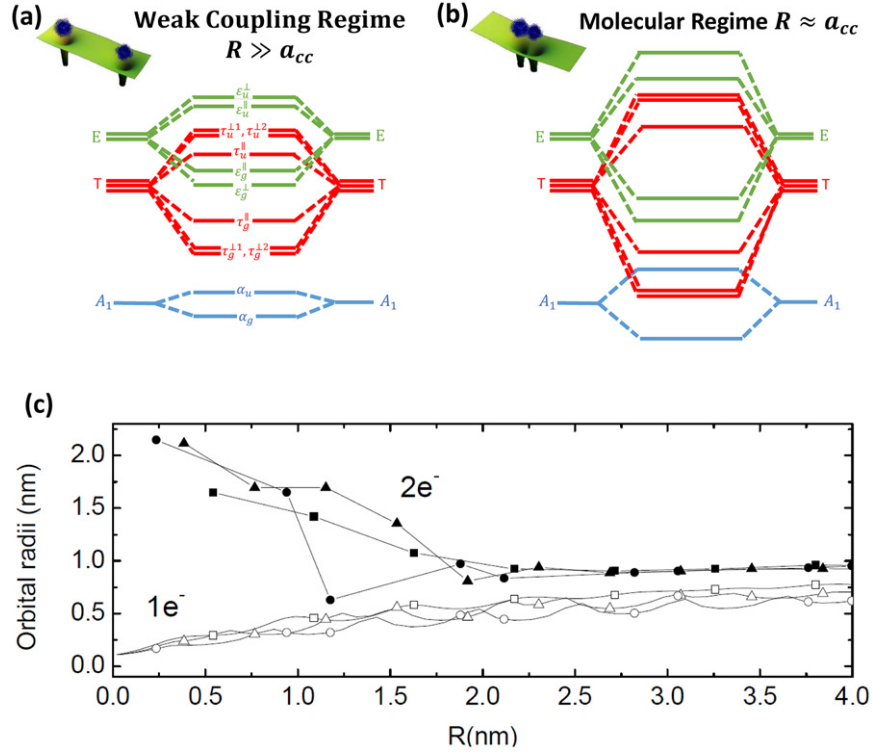
Figure 4. Schematic behavior of the molecular orbitals spectra from the 1s manifold with different valley compositions in the (*a*) weak coupling regime for a single electron and (*b*) molecular (strong coupling) regime for a single electron. (*c*) Variational radii calculated to minimize the energy of an As donor molecule oriented along the $\langle 1\,0\,0 \rangle$, $\langle 1\,1\,0 \rangle$, and $\langle 1\,1\,1 \rangle$ directions (squares, triangles and circles, respectively). The variational radii for the ionized $a_{\mathrm{As}_2^+}$ (empty symbols) and neutral $a_{\mathrm{As}_2^0}$ (solid symbols) molecule are shown.

is negligible and the states centered around sites A and B are degenerate, while for distances comparable to the one-atom wavefunction extension, *gerade* and *ungerade* combinations of the localized orbitals form, leading to split energies, see figures 4(*a*) and (*b*). At smaller interdonor distances, the *gerade–ungerade* splitting energy becomes comparable to the A_1-T valley-orbit splitting and a valley inversion of the first excited state is obtained. This result is discussed theoretically in [36] and confirmed experimentally in [37]. More recently, a complete analysis of this effect was presented in [18].

Unlike the H_2 molecule problem, it is possible for the antisymmetric state (referred to as antibonding in the context of H_2 in vacuum) to be the ground state here, since the hoppings are not necessarily real negative numbers due to the oscillatory phase $\exp(\mathrm{i}\mathbf{k}_\mu \cdot \mathbf{R})$. Still, we refer to the lower molecular orbital as *gerade* and the higher as *ungerade*.

The same oscillatory phase may also lift the degeneracies of the T_2 and E states. For instance, if the pair alignment is along the x direction, the states T_2^y and T_2^z are still equivalent, while the state T_2^x will have a symmetric-antisymmetric splitting that oscillates as a function of R. This effect is further enhanced in the presence of the effective mass anisotropy, as seen in [36]. The effective mass anisotropy also impacts the interdonor distance at which the valley inversion of the first excited state occurs.

It is known from the H_2 problem that the molecular orbital approximation gives accurate results only if the variational wavefunction radius is taken to minimize the expectation value of the complete hamiltonian including the two protons.

We do the same here, obtaining a variational radius $a_{\mathrm{D}_2^+}(\mathbf{R})$, which converges to the single atom orbit at large distances $a_{\mathrm{D}_2^+}(R \to \infty) = a_{\mathrm{cc}}$.

The resulting energies for the ground $[E_0(1\mathrm{e}^-)]$ and the first excited $[E_1(1\mathrm{e}^-)]$ states are plotted with symbols in figure 5(*a*). The squares, triangles and circles correspond to the three different molecule orientations considered $\langle 1\,0\,0 \rangle$, $\langle 1\,1\,0 \rangle$ and $\langle 1\,1\,1 \rangle$. The oscillations with R are due to the intervalley interference which produce the incommensurate oscillations on the single donor wavefunctions, illustrated in figures 2(*c*) and (*d*). This leads to subtle oscillations in the total energy, but significant oscillations in the energy difference $\Delta_{0-1} = E_1(1\mathrm{e}^-) - E_0(1\mathrm{e}^-)$ (figure 5(*b*)).

7. Two electrons in two donors: D_2^0 center

The quantum mechanical solution to the problem of two impurities with two interacting electrons is significantly less studied, even though it is the most commonly found configuration since it is neutral. The electron–electron correlations are hard to describe and add significant complexity to the picture in the previous section. The most systematic way to improve the accuracy of the two electron energies is to build a full Configuration Interaction (CI) wavefunction from the molecular orbitals discussed for the D_2^+ problem.

Now the suitability of truncating our basis set at the 1s manifold becomes clear. This manifold contains 6 atomic orbitals (one for each valley configuration), leading to 12 molecular orbitals and 144 two-electron Slater determinants.
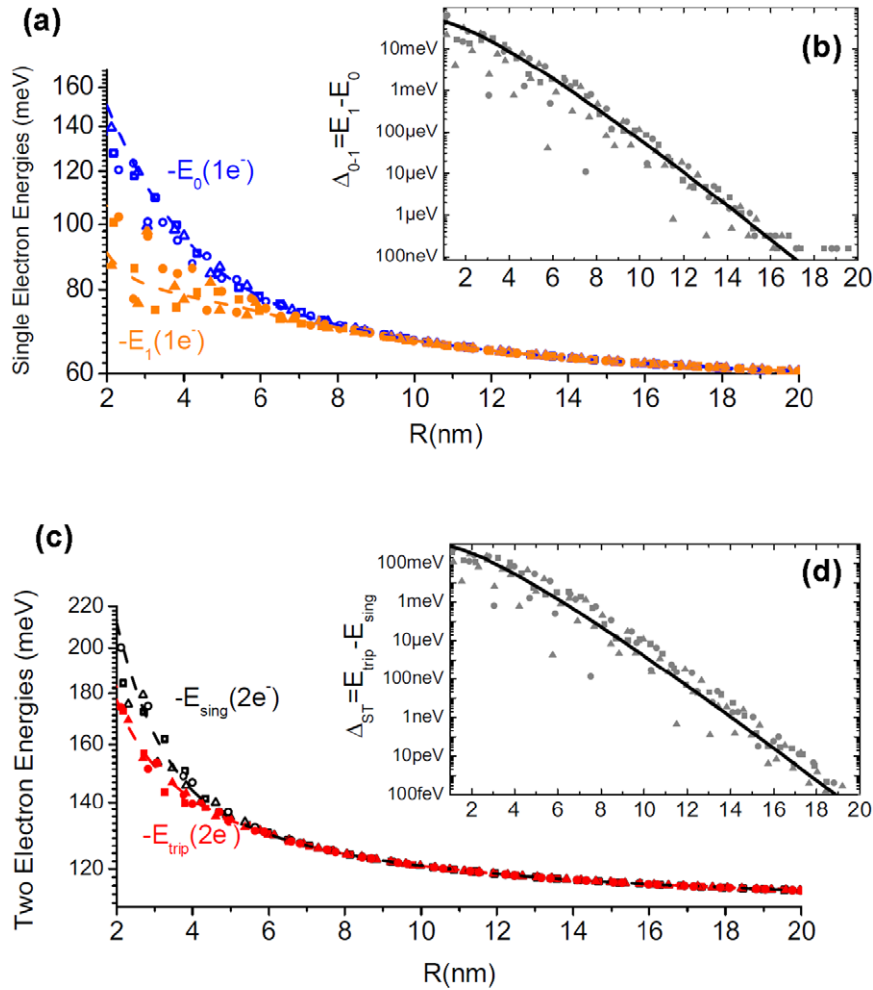
Figure 5. Ground and first excited state energies for one and two electrons in an As molecule. (*a*) The ground state is the bonding α_g orbital and the first excited is the antibonding α_u in the weak coupling regime ($R \gg a_{cc}$) or the bonding τ_g in the strongly interacting regime ($R \ll a_{cc}$), see figure 4. The dashed lines are the LCAO results using single-valley central cell corrected 1s wavefunctions, and taking the Hückel approximation for the off-diagonal integrals. (*b*) The energy separation between ground and first excited states Δ_{0-1}, which coincides with the tunnel coupling in the weak coupling regime. Squares, triangles and circles correspond to the $\langle 1\,0\,0 \rangle$, $\langle 1\,1\,0 \rangle$ and $\langle 1\,1\,1 \rangle$ molecule orientations respectively. (*c*) Two electron energies. The ground state is a spin singlet composed by two electrons in the bonding α_g orbital. The first excited state is a spin triplet, with one electron in α_g and the second electron in either α_u or τ_g (in the weakly or strongly interacting regimes, respectively). The dashed lines show the Hubbard approximation. (*d*) The singlet–triplet separation Δ_{ST}, which coincides with the spin–spin Heisenberg exchange coupling in the case of two electrons.

This way, the CI matrix for this problem would contain 20 736 elements, each consisting of a few Coulomb integrals—six-dimensional improper numerical integrals with an integrand diverging at all points where the electron–electron distance is null $r_{12} = 0$. Even a small increment of the atomic basis set increases the numerical demand of this problem significantly.

Still the problem stated above is too hard to be solved without any algebraic maneuvers and approximations. The first step is to identify elements that are null by symmetry. In the case of two electrons, one can always write down the wavefunction as a product of orbital and spinorial parts. Since the total spin operator and its projection both commute with the hamiltonian, we rewrite the molecular orbital basis set to be composed of eigenstates of these operators (singlets and triplets). This constitutes the so-called SACI (Spin Adapted

Configuration Interaction) method [38], decomposing the CI matrix into two uncoupled blocks: one singlet 78×78 block and one triplet 66×66 block. This cuts the computational effort by roughly a factor of two.

These dense singlet and triplet matrices are still too challenging if all the Coulomb integrals are calculated numerically. A very robust quantum chemistry approximation to overcome this problem is to expand the Slater-type orbitals into a series of N gaussians with given radii. This approach converges quickly with N, and for the calculations performed here for interdonor distances up to $R = 20$ nm, the energies were well converged with $N = 3$. For larger distances the number of gaussians increases due to the ill described tail decay [39]. The results for the numerical calculation of the singlet and triplet energies are presented in figure 5(*c*) while the Bohr radii $a_{D_2^0}$ resulting from the variational minimization are in

figure 4(c). The singlet triplet separation Δ_{ST}, in figure 5(d) reveals the same oscillations due to valley interference [14, 15] as the single electron excitation energy Δ_{0-1}.

The CI method is considered the standard model of modern quantum chemistry. In principle, arbitrarily accurate results may be obtained if a large enough single particle basis set is adopted. In practice, though, a truncated basis set could lead to largely incorrect results. For instance, in section 5 we adopted a simple two electron wavefunction, equation (8), which describes with a single orbital the energy of the D^- center with great accuracy. If instead a regular CI basis set is constructed from the multiorbital LCAO method, the variational energy obtained is much higher and the result is qualitatively wrong—an unbound D^- state is obtained, in contrast to the observed bound D^-.

8. Discussion

8.1. Simplified models and effective Hamiltonians

In certain studies, the description of donor states in Si based on microscopic model hamiltonians as presented in the previous sections may not be necessary or useful. Instead, simple models may be more adequate. In this section we investigate possibilities to map the problem into simplified models.

8.1.1. D_2^+: single orbital LCAO and the Hückel approximation.
Inclusion of higher excited orbitals improves the accuracy of the calculations, but in the simplest picture a single 1s orbital is considered, so that only two molecular-orbitals are obtained. Another simplifying assumption is to disregard the effect of valley physics. This approximation disregards the oscillations in the tunnel coupling. Instead, the tunnel coupling becomes a real monotonic function of the interdonor distance, and the molecular states are the symmetric and antisymmetric combinations of the atomic orbitals

$$|\alpha_g\rangle = \left(|A_1^A\rangle + |A_1^B\rangle\right)/\sqrt{2}, \quad (15)$$

$$|\alpha_u\rangle = \left(|A_1^A\rangle - |A_1^B\rangle\right)/\sqrt{2}. \quad (16)$$

The energies of these two states are $E = E_{onsite} \pm |t|$, where E_{onsite} is the onsite energy of the atomic orbital and t is the tunnel coupling, or hopping parameter.

The onsite energy is readily obtained. We rewrite the hamiltonian as $H = H_A + V_B$, so that H_A is identical to the single donor case, and $\langle A_1^A|H_A|A_1^A\rangle = E_0$ is simply the single electron in a single donor energy, known with great accuracy from experiments, hence we take $E_0 = E_{exp}$. The second term $\langle A_1^A|V_B|A_1^A\rangle$ could in principle be calculated exactly, but since our wavefunction description is not particularly accurate, especially for $R \approx a_{cc}$, we take a simplified point charge interaction, such that $\langle A_1^A|V_B|A_1^A\rangle \approx 125.76\,\text{meV nm}/R$.

The tunnel coupling could also be explicitly calculated, but a simpler result is obtained using the Hückel approximation. The argument behind this approximation is that the atomic orbital is an exact eigenstate of the single donor part H_A, and the second center potential V_B may be treated approximately

as a small correction over the H_A. In our case, we go further and take $\langle A_1^B|V_B|A_1^A\rangle = 0$, leading to $t = \langle A_1^B|H|A_1^A\rangle \approx \langle A_1^B|H_A|A_1^A\rangle = E_0 S(R, a_{cc})$, where the overlap function is $S(R, a_{cc}) = \langle A_1^B|A_1^A\rangle = e^{-R/a_{cc}}(1 + R/a_{cc} + R^2/3a_{cc}^2)$.

A comparison between these approximations and the results for the complete numerical calculation, shown in figures 5(a) and (b), reveals that this approximation gives the order of magnitude of the tunnel coupling (estimated as the separation between the ground and first excited states Δ_{01} in our complete multiorbital LCAO method of section 6), as well as the ground and first excited energies to an accuracy of approximately 10–15%, as seen in figure 5. The spread of the tunnel coupling in the numerical results is due to valley interference, which is disregarded here.

8.1.2. D_2^0: Hubbard model.
The energy for two electrons may be calculated with similar arguments. The electron–electron repulsion suppresses the probability of double occupation of the same donor. We take into account the correction to the singlet ground state due to the virtual occupation of this excited state, but discard any contributions from the triplet with both electrons at the same site since it has a much higher onsite energy. This is one of the ingredients of the so-called Hubbard approximation.

This way, the triplet energy is the easiest to calculate, consisting only of the binding energies of the two electrons to each respective donor and the classical attraction/repulsion across donors/electrons. Hence, $E_{trip}(2e^-) = \varepsilon_{(1,1)} = 2E_0 - e^2/4\pi\epsilon_{Si}R$.

The singlets are obtained diagonalizing the hamiltonian written in the basis $\{|s^A(\mathbf{r}_1)s^A(\mathbf{r}_2)\rangle, (|s^A(\mathbf{r}_1)s^B(\mathbf{r}_2)\rangle + |s^B(\mathbf{r}_1)s^A(\mathbf{r}_2)\rangle)/\sqrt{2}, |s^B(\mathbf{r}_1)s^B(\mathbf{r}_2)\rangle\}$, which reads

$$[H]_{singlets} = \begin{pmatrix} \varepsilon_{(2,0)} + U & \sqrt{2}t & 0 \\ \sqrt{2}t & \varepsilon_{(1,1)} & \sqrt{2}t \\ 0 & \sqrt{2}t & \varepsilon_{(0,2)} + U \end{pmatrix}. \quad (17)$$

We take the coupling between the (1,1) and (2,0)/(0,2) states to be determined by the single particle hopping t, which is calculated within the Hückel approximation discussed previously. This approximation is not necessarily accurate, since the size of the wavefunction of (2,0)/(0,2) is most likely comparable to the size of the doubly occupied single donor ion D^-. The most important parameter for the Hubbard model is the onsite charging energy U, which we take to be the experimental charging energy of the D^- single dopant discussed in section 5 (with exception of Sb and Bi, for which we adopt the results calculated in section 5).

By symmetry we have $\varepsilon_{(2,0)} = \varepsilon_{(0,2)}$ in the absence of detuning fields. The electrostatic arrangement of the (1,1) and the (2,0)/(0,2) charge configurations is significantly different, and therefore the onsite energy of these two states is not the same. On the other hand, the large same-site charging energy U leads to strong Coulomb blockade, and the occupation of the (2,0) and (0,2) states is only virtual. We argue that the error imposed by the approximation $\varepsilon_{(1,1)} = \varepsilon_{(2,0)}$ has a very small impact in the ground state energy (composed almost exclusively by the (1,1) configuration) as well as the singlet–triplet separation Δ_{ST}. These assumptions are confirmed *a*

Table 3. Explicit expressions (see section 8.1) for the energies of the ionized donor molecule $E(1e^-)$ and the neutral donor molecule $E(2e^-)$ as a function of the distance R between the donors and the Bohr radius calculated for the neutral donors a_{cc}. Input parameters for the numerical estimates are taken from tables 1 and 2. The quantity 125.76 is given in units of meV nm (see text).

$$E(1e^-) = E_0 - \frac{125.76}{R} - t(R, a_{cc})$$

$$E(2e^-) = 2E_0 - \frac{125.76}{R} - J(R, a_{cc})$$

$$t(R, a_{cc}) = E_0 e^{-R/a_{cc}} \left(1 + \frac{R}{a_{cc}} + \frac{R^2}{3a_{cc}^2}\right)$$

$$J(R, a_{cc}) = 2\frac{t(R, a_{cc})^2}{U}$$

Donor	E_0 (meV)	a_{cc} (nm)	U (meV)
P	−45.58	1.11	43.0
As	−53.77	0.82	51.7
Sb	−42.77	1.24	41.4
Bi	−70.88	0.58	63.45

posteriori, see the comparison between the numerical and the approximated values in figures 5(*c*) and (*d*).

8.2. Implications for experiments

These results indicate possible directions for the design and characterization of one and two donor quantum devices. They may also bring new insights in specific aspects of larger systems such as donor clusters [40], impurity chains [41] and δ-doped systems [42]. Recent experimental results confirm the adequacy and validity of the present theory [37, 43].

8.2.1. Transport through coupled donors.
One of the most reliable techniques for probing an isolated set of impurities is quantum transport spectroscopy [27, 43–47]. From these measurements the energy differences between occupation numbers (the chemical potential) are accessed directly.

From the expressions in table 3, one can see that the energy of one and two electron states contain the same long distance electrostatic term. Therefore, the difference between these two energies—i.e. the first ionization energy—indicates there is quantum interaction between the two donors. If the tunnel coupling between the donors is not comparable to the binding energy, then the first ionization energy is essentially the same one would obtain for a single donor, i.e. $E(2e^-) - E(1e^-) \approx E_0$. We refer to this as the *weak coupling regime* [47]. This is the most convenient regime for electrical control of the charge distribution and spin–spin interactions necessary for quantum computation with electron spins.

The opposite regime consists of small distances—referred to as the *molecular regime* [43]—in which an electronic cloud is shared between the sites A and B with no classically forbidden region between the two donors. The first ionization energy departs strongly from the single donor binding energy, in such a way that the identification of donor pairs in this regime is easier. The charging energy also provides an important independent confirmation of the donor interaction, since it is larger than that of a single donor D$^-$ state. This regime may be adopted for few dopant single electron transistors, leading to an enhanced charging energy and therefore to a more stable operation at room temperature.

Our formalism also provides the wavefunctions and valley compositions of the donor states, so that not only the spectrum, but also the conductance through these pairs could be analysed. For this to be accomplished, a model of the nanodevice at hand is necessary, since the overlap between these donor wavefunctions and the wavefunctions at the leads plays a central role in determining the conductance of these systems. Moreover, depending on the design, the electrons at the leads may be valley polarized (at least partially), leading to an angular dependence of the transport through two electron states and excited one electron states—as observed in [37]. A deeper analysis of the information that may be extracted from the conductance through pairs of donors is out of the scope of this paper.

So far we have only discussed transport through an isolated pair, which is typically probed by some nanodevice such as a FET based tunnel barrier [43]. The interactions between donors will also impact macroscopic ensembles of donors, as discussed in [48]. The impact of formation of donor molecules or clusters is clearly illustrated by comparison of our results in the large and short interdonor distance regimes. Molecules and clusters of donors form traps which limit the transport of electrons due to stronger binding energies. With some discretion, it might be possible to extend these results to describe the electronic structure of such ensembles of donors, which at low densities are predominantly isolated from each other in the extremely dilute limit (well described by the weak coupling regime here), but at higher densities—also called semidilute—form strongly interacting pairs (molecular regime), or even larger clusters in the intermediate doping limit. This intermediate doping cluster limit is an important problem [49] that could potentially be solved within the formalism presented here, while it would be very hard to solve it from a fully atomistic point of view.

8.2.2. Choice of impurity for quantum devices.
Often quantum technologies are proposed without a profound scrutiny of which group V impurity best serves the task at hand from the electronic structure point of view. For example in quantum computers, 1 and 2-qubit gates and measurements usually involve electron shuttling among donors or between donors and an interface, exchange pulses, all controlled by electric fields [50]. Different donors respond differently to electric fields, as expected from table 1.

Our non-perturbative approach provides energies, potentials and wavefunctions tailored for each chemical species of the group V substitutional donors in Si. In figure 6 we show (*a*) the tunnel coupling within single orbital LCAO theory and (*b*) the Heisenberg exchange coupling calculated within the Hubbard model. Both differ significantly from the results obtained using the KL wavefunction. The KL wavefunction overestimates the real tunnel and exchange coupling since it does not account for the effect of the valley-orbit coupling in shrinking the envelope function.

The simplified expressions in section 8.1, on the other hand, involve some approximations that overestimate and some that underestimate the tunnel and exchange couplings. For instance, taking a spherical effective mass of $0.3m_e$ leads to
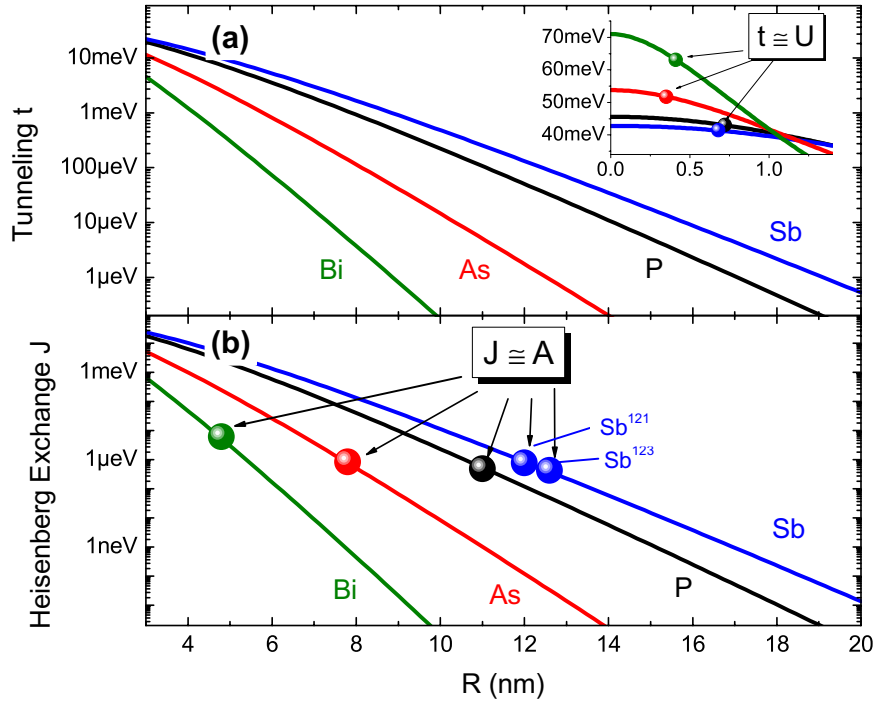
Figure 6. Quantum interactions for different impurity species. (*a*) The tunnel rate is plotted according to the single orbital LCAO simplified expression in table 3. The data points for which $t = U$ are marked (see inset), adopting the values of U listed in table 3. (*b*) The exchange coupling is obtained from the Hubbard model, with the expression also shown in table 3. The marked values are the data points for which $J = A$, where the hyperfine coupling A is taken from [52]. The values of the nuclear spin I and the hyperfine coupling for P, As, Sb^{121}, Sb^{123} and Bi are, respectively, $I = 1/2$, $3/2$, $5/2$, $7/2$ and $9/2$; and $A = 117.5$ MHz, 198.3 MHz, 186.8 MHz, 101.5 MHz and 1.4754 GHz.

a smaller wavefunction overlap compared to the transverse effective mass. On the other hand, disregarding valley interference and the wavefunction size dependence on the interdonor distance R leads to an increase in the overlap. We expect the net effect of these approximations to have a lesser negative impact than the approximations in the KL theory, besides providing a path to differentiate and compare impurity chemical species. We highlight particularly important values of the tunnel and the exchange coupling.

A modest impurity concentration generates impurity bands that are separated by the Mott gap due to the onsite electron–electron repulsion U. When the tunnel coupling—which determines the impurity band width—becomes comparable to U in figure 6(*a*), the system becomes a metal at half-filling. This is the Mott metal-insulator transition. The detailed interdonor distance at which this transition occurs for each impurity is hard to predict—strong electronic screening is expected at the phase transition. We highlight the distance at which the unscreened repulsion U matches the tunnel coupling t.

Another characteristic interdonor distance is that at which the exchange coupling becomes comparable to the nuclear hyperfine coupling. At these distances, the interaction between neighboring nuclei mediated through its electrons becomes optimally feasible [51]. This distance is very different amongst different species, as observed in figure 6(*b*). Moreover, this value is markedly different from the hydrogenic estimate adopted originally by Kane [50], which would set this distance at $R = 23$ nm independently of the chemical species.

On the other hand, the possibility of shuttling a single electron across the donor pair through a time dependent electric field (in analogy with the quantum operation of electrostatic quantum dots) is most efficient at $R > 20$ nm, for which the tunnel rate is in the range $t/h = 10$–100 MHz.

9. Summary and conclusions

Current progress in single or few dopants characterization [2] requires higher degree of sophistication where not only the nature (donor or acceptor) but actually identifying the atomic species of each impurity should be accessible to experiment. It is thus desirable to develop simple and reliable theories for such systems. We have presented a comprehensive study of donors in Si, which includes a central cell corrected potential obtained from consistency requirements with experiment. A single species-dependent parameter r_{cc} reproduces data for neutral P, As, Sb and Bi levels within the 1s sixfold lower energy manifold in Si. The parameter r_{cc} characterizes the range of the correction potential, and suffices to differentiate each species and respective potential.

We verify by experimental comparison (when available, otherwise by plausibility arguments) that the proposed central cell corrected potential for a given species in Si (D^0) is transferrable to other contexts of this species in Si, such as different charge states (D$^-$) and/or other atomic arrangements, e.g. donor pairs (D$_2^+$ and D$_2^0$). All approximations involved in different levels of the study are exposed and the range of validity or limitations are explicit. More realistic studies

exist in the literature, at the cost of treating only specific centers. For example, while we assume isotropic envelopes, a recent study [36] preserves the Si band anisotropy effects in an analysis restricted to D_2^+ (one-electron donor pair ions), with which our results are in fair agreement.

Among the calculated properties, the singlet–triplet difference Δ_{ST} (see figure 5(d)) in the neutral donor pair deserves special attention for donor-based qubits applications. This parameter may be identified with the exchange coupling between the electrons, which is the basic two-qubit entanglement mechanism in the first Si-based quantum processing proposal by Kane [50]. We obtain the spatial range of the wavefunctions, quantified by a_{cc}, significantly contracted relative to previous estimates based on the screened point-charge potential. One implication is that sizeable exchange coupling between electrons bound to a donor pair requires interdonor distances significantly smaller than previously expected and attempted. This is probably one of the reasons why exchange-coupled pairs were only reported very recently for As [43] and P [37] in Si. In both cases the models detailed here were instrumental to understand the reported results.

Acknowledgments

References

[1] Awschalom D D, Bassett L C, Dzurak A S, Hu E L, Petta J R 2013 *Science* **339** 1174
[2] Zwanenburg F A, Dzurak A S, Morello A, Simmons M Y, Hollenberg L C L, Klimeck G, Rogge S, Coppersmith S N and Eriksson M A 2013 *Rev. Mod. Phys.* **85** 961
[3] Morley G W 2015 Towards spintronic quantum technologies with dopants in silicon *Electron Paramagnetic Resonance* vol 24 (London: Royal Society of Chemistry)
[4] Ramdas A K and Rodriguez S 1981 *Rep. Prog. Phys.* **44** 1297
[5] Fetterman H R, Larsen D M, Stillman G E, Tannenwald P E and Waldmant J 1971 *Phys. Rev. Lett.* **26** 975
[6] Salfi J, Mol J A, Rahman R, Klimeck G, Simmons M Y, Hollenberg L C L and Rogge S 2014 *Nat. Mater.* **13** 605
[7] Luttinger J M and Kohn W 1955 *Phys. Rev.* **97** 869
[8] Kittel C and Mitchel A H 1954 *Phys. Rev.* **96** 1488
[9] Yamamoto T, Uda T, Yamasaki T and Ohno T 2009 *Phys. Lett.* A **373** 3989
[10] Kohn W 1957 *Shallow Impurity States in Silicon and Germanium* (*Solid State Physics* vol 5) (New York: Academic) pp 257–320
[11] Pantelides S T and Sah C T 1974 *Phys. Rev. B* **10** 621
[12] Nara H and Morita A 1966 *J. Phys. Soc. Japan* **21** 1852
[13] Martins A S, Boykin T B, Klimeck G and Koiller B 2005 *Phys. Rev. B* **72** 193204
[14] Wellard L C and Hollenberg L C L 2005 *Phys. Rev. B* **72** 085202
[15] Koiller B, Hu X and Das Sarma S 2002 *Phys. Rev. B* **66** 11520
[16] Baena A, Saraiva A L, Koiller B and Calderón M J 2012 *Phys. Rev. B* **86** 035317
[17] Pica G, Lovett B W, Bhatt R N and Lyon S A 2014 *Phys. Rev. B* **89** 235306
[18] Gamble J K, Jacobson N T, Nielsen E, Baczewski A D, Moussa J E, Montaño I and Muller R P 2014 arXiv:1408.3159
[19] Oliveira L E and Falicov L M 1986 *Phys. Rev. B* **33** 6990
[20] Menchero J G, Capaz R B, Koiller B and Chacham H 1999 *Phys. Rev. B* **59** 2722
[21] Martins A S, Menchero J G, Capaz R B and Koiller B 2002 *Phys. Rev. B* **65** 24205
[22] Wang J S-Y and Kittel C 1973 *Phys. Rev. B* **7** 713
[23] Grimmeiss H G, Janzen E and Larsson K 1982 *Phys. Rev. B* **25** 2627
[24] Pajot B 2010 *Optical Absorption of Impurities and Defects in Semiconducting Crystals* (Heidelberg: Springer)
[25] Saraiva A L, Calderón M J, Capaz R B, Hu X D, Das Sarma S and Koiller 2011 *Phys. Rev. B* **84** 155320
[26] Calderón M J, Verduijn J, Lansbergen G P, Tettamanzi G C, Rogge S and Koiller B 2010 *Phys. Rev. B* **82** 075317
[27] Sellier H, Lansbergen G P, Caro J, Rogge S, Collaert N, Ferain I, Jurczak M and Biesemans S 2006 *Phys. Rev. Lett.* **97** 206805
[28] Bethe H A and Salpeter E E 1957 *Quantum Mechanics of One and Two Electron Atoms* (New York: Academic)
[29] Hylleraas E A 1929 *Z. Phys.* **54** 347
[30] Narita S, Shinbashi T and Kobayashi M 1982 *J. Phys. Soc. Japan* **51** 2186
[31] Kamimura H 1979 *J. Non-Cryst. Solids* **32** 187
[32] Larsen D M 1981 *Phys. Rev. B* **23** 5521
[33] Inoue J, Nakamura J and Natori A 2008 *Phys. Rev. B* **77** 125213
[34] Slater J C 1963 *Quantum Theory of Molecules and Solids* vol 1 (New York: McGraw-Hill)
[35] Miller A and Abrahams E 1960 *Phys. Rev. B* **120** 745
[36] Klymenko M V and Remacle F 2014 *J. Phys.: Condens. Matter* **26** 065302
[37] Dehollain J P, Muhonen J T, Tan K Y, Saraiva A, Jamieson D N, Dzurak A S and Morello A 2014 *Phys. Rev. Lett.* **112** 236801
[38] Wang J and Hines C 2009 *J. Phys.: Conf. Ser.* **185** 012053
[39] Saraiva A L, Calderón M J and Koiller B 2007 *Phys. Rev. B* **76** 233302
[40] Weber B, Tan Y H M, Mahapatra S, Watson T F, Ryu H, Rahman R, Hollenberg L C L, Klimeck G and Simmons M Y 2014 *Nat. Nanotechnol.* **9** 430
[41] Prati E, Hori H, Guagliardo F, Ferrari G and Shinada T 2012 *Nat. Nanotechnol.* **7** 443
[42] Shamim S, Mahapatra S, Scappucci G, Klesse W M, Simmons M Y and Ghosh A 2014 *Phys. Rev. Lett.* **112** 236602
[43] Gonzalez-Zalba M F, Saraiva A, Calderón M J, Heiss D, Koiller B and Ferguson A J 2014 *Nano Lett.* **14** 5672
[44] Gonzalez-Zalba M F, Heiss D and Ferguson A J 2012 *New J. Phys.* **14** 023050

[45] Lansbergen G P, Rahman R, Wellard C J, Woo I, Caro I J, Collaert N, Biesemans S, Klimeck G, Hollenberg L C L and Rogge S 2008 *Nat. Phys.* **4** 656

[46] Pierre M, Wacquez R, Jehl X, Sanquer M, Vinet M and Cueto O 2009 *Nat. Nanotechnol.* **5** 133

[47] Roche B, Riwar R-P, Voisin B, Dupont-Ferrier E, Wacquez R, Vinet M, Sanquer M, Splettstoesser J and Jehl X 2013 *Nat. Commun.* **4** 1581

[48] Thomas G A, Capizzi M, DeRosa F, Bhatt R N and Rice T M 1981 *Phys. Rev.* B **23** 5472

[49] Voyles P M, Muller D A, Grazul J L, Citrin P H and Gossmann H-J L 2002 *Nature* **416** 826

[50] Kane B E 1998 *Nature* **393** 133

[51] Kalra R, Laucht A, Hill C and Morello A 2014 *Phys. Rev.* X **4** 021044

[52] Feher G 1959 *Phys. Rev.* **114** 1219