Prediction of dopant ionization energies in silicon: The importance of strain

A. Rockett, University of Illinois at Urbana-Champaign
Duane D. Johnson, University of Illinois at Urbana-Champaign
S. V. Khare, University of Illinois at Urbana-Champaign
B. R. Tuttle
Prediction of dopant ionization energies in silicon: The importance of strain


1Department of Materials Science and Engineering, University of Illinois, 1304 West Green Street, Urbana, Illinois 61801, USA
2Department of Physics, School of Science Penn State Erie, The Behrend College, Erie, Pennsylvania 16563-0203, USA

Received 25 July 2003; published 30 December 2003

Based on a hydrogenic state and strain changes upon defect charging, we propose a simple, parameter-free model that agrees well with the observed group III and V monovalent-impurity ionization energies in Si, revealing the importance of such strain effects. Changes in lattice strain upon defect charging are obtained via superposition and elasticity theory using atomic relaxations from density functional theory.

DOI: 10.1103/PhysRevB.68.233208 PACS number(s): 71.55.Cn, 71.15.Mb

For electronic and optical semiconductor devices, a fundamental issue is the number of free carriers. This number is determined primarily by the density of impurity states and the energy separating them from the appropriate host band edge. A reliable and practical means to determine the relative position of impurity-specific defect states would be a major advance for device design. Save N, ionization energies in Si of substitutional group III and V impurities are known and shown in Fig. 1 versus dopant’s covalent radius. The smaller elements (B, Al, Ga, P, and As) produce states within ~40–60 meV of the appropriate Si band edge. However, acceptors become increasingly deep as their size increases. Electronically similar atoms, such as In and Sb or Tl and Bi, show dramatic differences in behavior. A full understanding, and quantitative prediction, of these trends is lacking.

Defect states arising from monovalent impurities can be estimated from dielectric screening arguments and effective-mass theory, yielding the impurity-independent hydrogenic-state model for the electron-hole binding energy.1,2 In contrast, impurity-dependent localized states can be treated using a model by Hjalmarson et al.,3 in which levels are mostly determined by the valence p-state energies of the impurity. This model has been extended and offers a means for understanding trends, but does not give quantitative predictions.

Figure 1 clearly suggests that strain is important. Yet covalent radii reflect neutral-atom strain and not strain (or size) changes upon dopant ionization. Some effects of strain have been discussed before. A phenomenological model based on hydrostatic lattice strains is useful for estimating segregation energies for dopants.3 Within the effective-mass theory of shallow impurities, central cell and strain effects have been discussed before. A phenomenological model based on screening arguments and effective-mass theory using a model by Hjalmarson et al.4 offers a means for calculating ionization energies for shallow dopants in semiconductors.

A simple explanation of Fig. 1 is that a large donor losing an electron to the host becomes smaller, decreasing strain energy and making ionization more favorable. Likewise a large acceptor gaining an electron should be even larger, increasing the strain energy. Here we propose a parameter-free model that incorporates these effects via elasticity theory using DFT atomic relaxations. Our calculated dopant levels agree well with those in Fig. 1 and show that they can have significant contributions from strain changes, even in levels over 150 meV deep.

Equation (1) can be anticipated by thermodynamics using \( \mu_d N = dU + PdV \) at \( T = 0 \) K. Here \( \mu \) is the chemical potential, \( dN \) is the change in free carriers, \( dU \) is the change in bonding energy, \( PdV \) is work due to hydrostatic pressure (\( P \)) change in the impurity volume (\( V \)) upon relaxation. The change \( \mu dN \) determines the ionization energy. From states \( I(dN = -1) \) and \( A(dN = -1) \), the shift is \( (dU_I - dU_A) + [(PdV)_I - (PdV)_A] \). The first and second bracketed terms are, respectively, the hydrogenic-state energy (ignoring \( \Delta_{\text{loc}} \)) and the change in strain upon dopant charging.

We obtain a simple but accurate treatment of strain from linear elasticity theory based on atomic relaxations from DFT calculations, which incorporate local atomic-level effects. Eshelby\(^\text{12}\) derived the strain energy of a misfit inclusion in an elastic solid as \( E_{\text{strain}} = C \pi r_{\text{imp}}^2 \), where \( r_{\text{imp}} \) is the radius of the impurity and \( \epsilon \) is the central strain due to the insertion of that inclusion into the host. Although here we present only the result for spherical strain fields, one can calculate the strain energy of an elliptical inclusion.\(^\text{12}\) Here

\[ \Delta = (\Delta_{\text{loc}} + \Delta_{\text{hyd}}) + \Delta_{\text{strain}}, \]

where \( \Delta_{\text{loc}} \) is a localized state energy, \( \Delta_{\text{hyd}} \) is the hydrogenic-state energy \([-0.032 \text{ eV for donors and 0.053 eV for acceptors in Si (Ref. 1)} \] and \( \Delta_{\text{strain}} \) results from a change in strain energy upon dopant charging.

\( \Delta_{\text{strain}} \) is the central strain due to the insertion of the inclusion into the host. Although here we present only the result for spherical strain fields, one can calculate the strain energy of an elliptical inclusion.\(^\text{12}\) Here
The bond strain is included indirectly via use of superposition and DFT strains, strain changes due to localization of defect states are included from separate calculations. Therefore, applying linear elasticity generally works down to atomic scales.\textsuperscript{14,15} To the dopant level should be correct for dilatational strain, as band edge versus covalent radii is straightforward and quite accurate. The 64-atom impurity-Si bond lengths, if used, would introduce large errors in strain ($\sim 2 \times$overestimate for Bi). Yet an extrapolation from smaller cells to dopant charge state, particularly if the dilute-limit strain is correctly obtained.

Periodic boundary conditions impose severe restrictions on the strain field around a defect, by producing, e.g., symmetry-induced, zero-strain nodes. The first node occurs between the second and third NN for our cells, reducing the magnitude of the strain field throughout the cell and significantly affecting values outside of the nearest neighbor, at least until cells are very large.

Our Eq. (2) (derived from Eshelby’s formula) is particularly useful as it determines the energy of an infinite elastic strain field from a given central (NN bond) strain, as long as periodicity does not affect the central strain. Because periodicity (and impurity density) does affect central strain, an analysis of the bond lengths as a function of cell size is needed. In Fig. 2 we show NN bond lengths versus 1/$N$ for fixed-symmetry cells for a neutral acceptor (TI). (Similar plots can be made for all impurities and charge states.) While bond lengths from cells with differing symmetry for the same impurity are very different, the $N=\infty$ extrapolated values from the two different fixed-symmetry cells agree well. Therefore, the $N=\infty$ value is independent of cell symmetry and provides the impurity-limit bond length for all configurations, i.e., (un)doped Si in $I$, $A$, and neutral charge states. Clearly, the 64-atom impurity-Si bond lengths, if used, were chosen so that the total energy for each cell varied less than 50 meV. As the results depend on relative energies, we expect the systematic numerical convergence to be 10 meV or better.

All (un)doped calculations were done in $N$-atom simple-cubic (sc) or face-centered cubic (fcc) cubic cells with periodic boundary conditions for $N=8$, 16, 54, 64, 128, and 216 atoms. Impurity atoms studied were Al, Ga, In, TI, As, Sb, and Bi. Total energies and atomic positions were obtained for pure and doped Si for each charge state and cell size, both for relaxed and unrelaxed (i.e., at bulk Si sites) states. Relaxed atomic positions were found when the total energy changed by less than 1 meV for a force tolerance of about 10 meV/nm.
independent of cell symmetry.

From Eq. \(\Delta E_i = 0.064 \text{ eV and } \Delta E_A = -0.066 \text{ eV} \) so \( \Delta E_i - \Delta E_A = 0.129 \text{ eV} \) for Tl. The "strain-hydrogenic" model [Eq. (3)] predicts \( \Delta E_T = 0.182 \text{ eV} \). The Tl "A" state strain energy is negative, reflecting its smaller size in this charge state. Generally the calculated strain energies for \( I \) and \( A \) states of the doped cells relative to the \( N \) state have opposite signs (Table I). This is not surprising as a bond in which an electron is added to a bonding state should expand, while a bond with an electron removed from a bonding state should contract. The converse applies to antibonding states. Acceptor ionization consists of removing an electron from a bonding state of the matrix and placing it in the bonding state of the impurity.

The levels for the other dopants may be similarly calculated and are shown in Fig. 3 versus LDA bond strains, along with the experimental data. Equation (3) yields agreement to observed trends and values and provides an intuitive understanding. Our results suggest that many impurity levels, while far deeper than a normal hydrogenic state (even 150 meV from a band edge), can have a significant contribution from a strain change upon defect charging. Thus, it is not surprising that it would be difficult to predict impurity states accurately based on a purely localized-state or purely hydrogenic model. An exact or improved DFT calculation could include all three competing effects concomitantly, but this remains to be done quantitatively.

Only Tl and Bi show any significant discrepancy from experiment, and they exhibit the largest strain effect. (Bi has the largest error bars\(^{19}\). For the largest dopants, the discrepancy is due either to (1) an underestimate of the local impurity-specific strain effect (e.g., by not properly handling band-overlap error via extrapolation) and/or, more likely, (2) there is an emergence of a localized atomic-state behavior. Notably, item 2 is more important for acceptors as the valence (conduction) band extremum is due primarily to \( p \) states (antibonding \( s \) states) such that acceptor-Si strain results in a more localized state near the valence band. The Tl and Bi results in Fig. 3 are in accord with this.

**Conclusions.** We presented a strain-hydrogenic model to calculate impurity ionization energies in semiconductors. Our model predicts, without parameters and with reasonable agreement to experiment, a range of dilute-concentration monovalent-impurity ionization energies in silicon.

**Level calculation.** With DFT bond strains for each charge state (Table I and Fig. 2), we may obtain dopant levels. Consider the acceptor Tl. With neutral Si as reference and values from Table I, we find \( r_{N_2}^{\text{Tl}} = 0.2521 \text{ nm} \), \( r_{A_2}^{\text{Tl}} = 0.2510 \text{ nm} \), \( r_{I_2}^{\text{Tl}} = 0.2534 \text{ nm} \), and \( r_{A_2}^{\text{Tl}} = 0.2336 \text{ nm} \), and the misfits \( \Delta r_N = 0.0198 \text{ nm} \), \( \Delta r_A = 0.0174 \text{ nm} \), and \( \Delta r_N = 0.0187 \text{ nm} \). From Eq. (2), we obtain \( \Delta E_i = 0.064 \text{ eV} \) and \( \Delta E_A = -0.066 \text{ eV} \) so \( \Delta E_i - \Delta E_A = 0.129 \text{ eV} \) for Tl. The "strain-hydrogenic" model [Eq. (3)] predicts \( \Delta E_T = 0.182 \text{ eV} \). The Tl "A" state strain energy is negative, reflecting its smaller size in this charge state. Generally the calculated strain energies for \( I \) and \( A \) states of the doped cells relative to the \( N \) state have opposite signs (Table I). This is not surprising as a bond in which an electron is added to a bonding state should shrink, while a bond with an electron removed from a bonding state should expand. The converse applies to antibonding states. Acceptor ionization consists of removing an electron from a bonding state of the matrix and placing it in the bonding state of the impurity.

The levels for the other dopants may be similarly calculated and are shown in Fig. 3 versus LDA bond strains, along with the experimental data. Equation (3) yields agreement to observed trends and values and provides an intuitive understanding. Our results suggest that many impurity levels, while far deeper than a normal hydrogenic state (even 150 meV from a band edge), can have a significant contribution from a strain change upon defect charging. Thus, it is not surprising that it would be difficult to predict impurity states accurately based on a purely localized-state or purely hydrogenic model. An exact or improved DFT calculation could include all three competing effects concomitantly, but this remains to be done quantitatively.

Only Tl and Bi show any significant discrepancy from experiment, and they exhibit the largest strain effect. (Bi has the largest error bars\(^{19}\). For the largest dopants, the discrepancy is due either to (1) an underestimate of the local impurity-specific strain effect (e.g., by not properly handling band-overlap error via extrapolation) and/or, more likely, (2) there is an emergence of a localized atomic-state behavior. Notably, item 2 is more important for acceptors as the valence (conduction) band extremum is due primarily to \( p \) states (antibonding \( s \) states) such that acceptor-Si strain results in a more localized state near the valence band. The Tl and Bi results in Fig. 3 are in accord with this.

**Conclusions.** We presented a strain-hydrogenic model to calculate impurity ionization energies in semiconductors. Our model predicts, without parameters and with reasonable agreement to experiment, a range of dilute-concentration monovalent-impurity ionization energies in silicon.

---

**TABLE I.** DFT nearest-neighbor distances \( r_i \) (nm) for dopants in relaxed cells with \( N = 216 \) (top) and extrapolated \( N \to \infty \) (bottom) for charge states \( i = N, I \), and \( A \). The neutral host (reference) Si-Si value is \( r_{N_2}^{\text{host}} = 0.2336 \text{ nm} \).

<table>
<thead>
<tr>
<th></th>
<th>As-Si</th>
<th>Sb-Si</th>
<th>Bi-Si</th>
<th>Al-Si</th>
<th>Ga-Si</th>
<th>In-Si</th>
<th>Tl-Si</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>0.2406</td>
<td>0.2545</td>
<td>0.2617</td>
<td>0.2397</td>
<td>0.2364</td>
<td>0.2504</td>
<td>0.2521</td>
</tr>
<tr>
<td></td>
<td>0.24064</td>
<td>0.25454</td>
<td>0.2617</td>
<td>0.23968</td>
<td>0.23634</td>
<td>0.25044</td>
<td>0.25225</td>
</tr>
<tr>
<td>( I )</td>
<td>0.2403</td>
<td>0.2542</td>
<td>0.2609</td>
<td>0.2401</td>
<td>0.2367</td>
<td>0.2511</td>
<td>0.2530</td>
</tr>
<tr>
<td></td>
<td>0.24039</td>
<td>0.25441</td>
<td>0.26116</td>
<td>0.24019</td>
<td>0.23669</td>
<td>0.25148</td>
<td>0.25337</td>
</tr>
<tr>
<td>( A )</td>
<td>0.2408</td>
<td>0.2547</td>
<td>0.2621</td>
<td>0.2392</td>
<td>0.2362</td>
<td>0.2496</td>
<td>0.2510</td>
</tr>
<tr>
<td></td>
<td>0.24086</td>
<td>0.25463</td>
<td>0.26198</td>
<td>0.23916</td>
<td>0.23615</td>
<td>0.24958</td>
<td>0.25102</td>
</tr>
</tbody>
</table>
model is simple to apply and yields important insight into a highly practical and long-standing problem. We find that strain change upon ionization is an important, and occasionally dominant, contribution to dopant levels. It is responsible for even small deviations from hydrogenic values found for donors in Si. Local DFT calculations from finite, fixed-symmetry cells were shown to provide accurate impurity-host bond lengths and strains when obtained via a simple extrapolation to infinite-cell size. The strain method presented here has the potential for describing the ionization energy for a variety of systems including compound semiconductors, extended defects and doped quantum dots.

Support at the University of Illinois at Urbana-Champaign was from the U.S. Department of Energy, Division of Materials Sciences (DEFG02-91ER45439) through the Frederick Seitz Materials Research Laboratory, a NSF FRG (DMR-9976550) and ITR (DMR-0121695), an IBM SURS grant and the SGI-Origin2000 at the National Center for Supercomputing Applications. We thank G. Kresse for the Tl pseudopotential.


19 The error bar is $\delta CG_{r}N \sqrt{[(\Delta r_{x})^{2}+(\Delta r_{y})^{2}]}$. 