

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

A. Rockett,¹ D. D. Johnson,¹ S. V. Khare,¹ and B. R. Tuttle²

¹*Department of Materials Science and Engineering, University of Illinois, 1304 West Green Street, Urbana, Illinois 61801, USA*

²*Department 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 known¹ 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 extended⁴ 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.⁵ Within the effective-mass theory of shallow impurities, central cell and strain effects have been included using empirical^{6,7} and *ab initio*^{6,8} methods. The latter *ab initio* method is formally rigorous but technically involved and, to our knowledge, it has only been applied to acceptors in germanium.^{8,9} Given the present situation, there is clearly a need for a physically sound and technically accessible method 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 us-

ing 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.

Strain-hydrogenic model. As strain changes with charge state of the matrix and impurity atom, it is essential to consider the relevant charge states. For (un)doped systems, the total energies E_A with one extra electron (affinity *A* state) and E_I with one missing electron (ionization *I* state) are determined relative to the neutral (*N*) state. In an exact DFT, the energy gap (dopant level) of a undoped (doped) semiconductor is $\Delta^{u(d)} = E_I^{u(d)} - E_A^{u(d)}$,¹⁰ with contributions from Kohn-Sham eigenvalues (Δ_{KS}) and exchange-correlation (Δ_{xc}) effects. Δ^u for pure Si using a local-density approximation (LDA) is too small by 50%, where error in Δ_{xc} gives $\sim 80\%$ of that discrepancy.^{10,11} Although LDA dopant levels are in error, we note that LDA bond-strains are accurate.

To model the effect of strain on defect states, we propose that the Δ^d is (dropping the superscript)

$$\Delta = (\Delta_{loc} + \Delta_{hyd}) + \Delta_{strain}, \quad (1)$$

where Δ_{loc} is a localized state energy, Δ_{hyd} is the *hydrogenic-state* energy [0.032 eV for donors and 0.053 eV for acceptors in Si (Ref. 1)], and Δ_{strain} results from a change in strain energy upon defect ionization.

Equation (1) can be anticipated by thermodynamics using $\mu dN = dU + PdV$ at $T = 0$ K. Here μ 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 μdN 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 Δ_{loc} directly) 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. Eshelby¹² derived the strain energy of a misfit inclusion in an elastic solid as $E_{strain} = CGr_{imp}^3 \epsilon^2$, where r_{imp} is the radius of the impurity and ϵ 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.^{12,15} Here

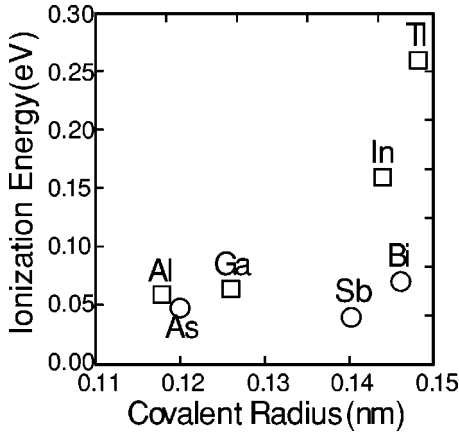


FIG. 1. Experimental data (Ref. 1) for levels (eV) of monovalent dopants from groups IIIa and Va relative to the appropriate Si band edge versus covalent radii (nm).

$C = [8\pi(1 + \nu)]/[3(1 - \nu)]$, ν is Poisson's ratio, and G is the shear modulus of the host. For Si, ν is 0.17 and G is 79.6 GPa (or 497 eV/nm³),¹³ giving $CG = 5869$ eV/nm³. Assuming the bond strain is $\epsilon = (\Delta r_i)/r_i$ for misfits $\Delta r_i \equiv r_i - r_{\text{host}}$ for the dopant radii r_i in charge states $i = I, A$, or N , the strain energy is $E_{\text{strain}} = CGr_i(\Delta r_i)^2$. Hence, the change in strain energy upon ionization (I or A), $\Delta E_i = E_i - E_N$, is

$$\Delta E_i = CGr_N[(\Delta r_i)^2 - (\Delta r_N)^2], \quad (2)$$

where a term of order $(r_i - r_N)^2(\Delta r_i)^2$ was ignored. Here E_i (E_N) is the strain energy associated with the dopant radius r_i (r_N) for ionized (neutral) case. The simple estimate from Eq. (2) for the impurity-limit bond-strain contribution to the dopant level should be correct for dilatational strain, as linear elasticity generally works down to atomic scales.^{14,15}

In applying Eq. (2), as in an exact DFT $E_I - E_A$ analysis, we require a comparison of strain energy changes for I and A states from separate calculations. Therefore, applying linear superposition, we obtain a classical estimate from $(\Delta E_I - \Delta E_A)_{\text{Eshelby}}$, and the predicted impurity level from our *strain-hydrogenic* model is

$$\Delta = \Delta_{\text{hyd}} + (\Delta E_I - \Delta E_A)_{\text{Eshelby}}. \quad (3)$$

So levels shift in energy in proportion to the stress from the hydrostatic strain changes due to dopant charging. Note that strain changes due to localization of defect states are included indirectly via use of superposition and DFT strains, while any direct Δ_{loc} effect on ionization energy in Eq. (1) is ignored.

Computational details. We employ VASP (Vienna ab initio simulation package¹⁶) based on LDA.¹⁰ VASP evaluates forces directly and treats core electrons by ultrasoft Vanderbilt-type pseudopotentials,¹⁷ as provided by Kresse and Hafner.¹⁸ The Si pseudopotential has an outer core radius of 2.48 a.u. (≈ 0.131 nm). Wave functions are expanded in a plane-wave basis using a 150 eV kinetic-energy cutoff. Integration over the Brillouin zone is performed using a symmetry-reduced k -point grid whose size varied depending on the symmetry and size of the supercell. The k -point grid and other param-

eters 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, Tl, 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.

Dilute-impurity-limit strains. One intrinsic strength of DFT is the accuracy of its structural parameters. The LDA Si nearest-neighbor (NN) distance is 0.2336 nm at $T = 0$ K, while the accepted experimental value is 0.2351 nm at 120 K,¹ indicating an absolute error of 0.0015 nm. LDA bond lengths are systematically low, consistent with these values. However, the changes in bond lengths (relative values) are expected to be about an order of magnitude more accurate than the absolute lengths. Analyzing fixed-cell results, we estimate relative changes in bond lengths are determined to an accuracy of $\delta = \pm 0.0002$ nm upon changes in cell charge, doping, and relaxation (providing error bars,¹⁹ see below). We therefore expect LDA to provide an accuracy sufficient to describe variations of lattice strains due to changes in 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 (Tl). (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, would introduce large errors in strain ($\sim 2 \times$ overestimate for Bi). Yet an extrapolation from smaller cells to $N = \infty$ values is straightforward and quite accurate. The 216-atom and $N = \infty$ NN bond length for all dopants (and charge states) studied are given in Table I. Bond lengths for use in Eq. (2) are taken from the $N = \infty$ values.

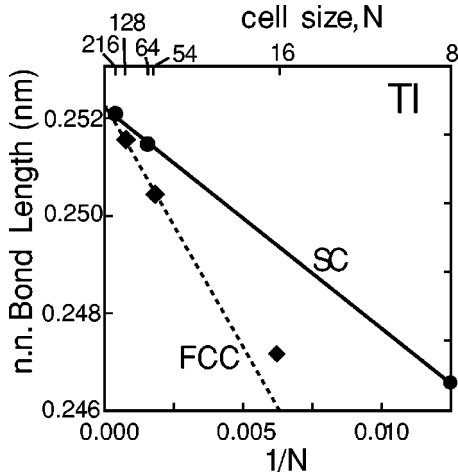


FIG. 2. Impurity-Si bond lengths (nm) for neutral TI vs $1/N$. Fixed-symmetry results are for sc ($N=8, 64,$ and 216 : circles) and fcc ($N=16, 54,$ and 128 : diamonds) cells. The $N=\infty$ values is independent of cell symmetry.

Level calculation. With DFT bond strains for each charge state (Table I and Fig. 2), we may obtain dopant levels. Consider the acceptor TI. With neutral Si as reference and values from Table I, we find $r_N^{\text{TI}}=0.2521$ nm, $r_A^{\text{TI}}=0.2510$ nm, $r_I^{\text{TI}}=0.2534$ nm, and $r_N^{\text{Si}}=0.2336$ nm, and the misfits $\Delta r_I=0.0198$ nm, $\Delta r_A=0.0174$ nm, and $\Delta r_N=0.0187$ nm. From Eq. (2), we obtain $\Delta E_I=0.064$ eV and $\Delta E_A=-0.066$ eV so $(\Delta E_I-\Delta E_A)_{\text{Eshelby}}=0.129$ eV for TI. The “strain-hydrogenic” model [Eq. (3)] predicts $\Delta_{\text{TI}}=0.182$ eV. The TI “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,

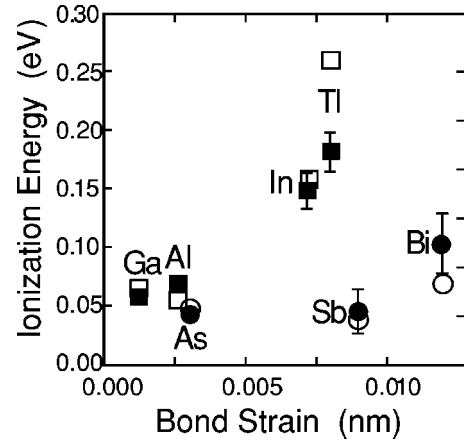


FIG. 3. Equation (3) ionization energies (eV) vs DFT neutral impurity-Si bond strains. Theoretical results (solid points) are compared to experimental data (open points) for acceptors (squares) and donors (circles). Strain errors (see text) are shown as bars if bigger than points.

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 TI and Bi show any significant discrepancy from experiment, and they exhibit the largest strain effect. (Bi has the largest error bars¹⁹). 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 TI 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. The

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

	As-Si	Sb-Si	Bi-Si	Al-Si	Ga-Si	In-Si	Tl-Si
N	0.2406	0.2545	0.2617	0.2397	0.2364	0.2504	0.2521
	0.24064	0.25454	0.2617	0.23968	0.23634	0.25044	0.25225
I	0.2403	0.2542	0.2609	0.2401	0.2367	0.2511	0.2530
	0.24039	0.25441	0.26116	0.24019	0.23669	0.25148	0.25337
A	0.2408	0.2547	0.2621	0.2392	0.2362	0.2496	0.2510
	0.24086	0.25463	0.26198	0.23916	0.23615	0.24958	0.25102

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 TI pseudopotential.

-
- ¹E.g., C. Kittel, *Introduction to Solid State Physics*, 5th ed. (Wiley, New York, 1976), pp. 233, 234; for TI, W. Scott and J.L. Schmit, *Appl. Phys. Lett.* **33**, 294 (1978); S.M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981), p. 21; A.K. Ramdas and S. Rodriguez, *Rep. Prog. Phys.* **44**, 1297 (1981).
- ²J.M. Luttinger and W. Kohn, *Phys. Rev.* **97**, 869 (1955).
- ³H. Hjalmarson *et al.*, *Phys. Rev. Lett.* **44**, 810 (1980).
- ⁴R.W. Jansen and O.F. Sankey, *Phys. Rev. B* **33**, 3994 (1986).
- ⁵C.W. Myles and W.G. Li, *J. Phys. Chem. Solids* **61**, 1855 (2000); W.G. Li and C.W. Myles, *Phys. Rev. B* **47**, 4281 (1993).
- ⁶S.T. Pantelides, *Rev. Mod. Phys.* **50**, 797 (1978).
- ⁷A.H. Edwards and W.B. Fowler, *Phys. Rev. B* **16**, 3613 (1977); N.O. Lipari, A. Baldereschi, and M.L.W. Thewalt, *Solid State Commun.* **33**, 277 (1980); G.T. Einevoll and Y.C. Chang, *Phys. Rev. B* **40**, 9683 (1989); D. Sasireka and E. Palaniyandi, *ibid.* **46**, 2047 (1993).
- ⁸N. Binggeli and A. Baldereschi, *Phys. Rev. B* **45**, 5944 (1992).
- ⁹H. Fritzsche, *Phys. Rev.* **125**, 1552 (1962).
- ¹⁰E.g., R.M. Dreizler and E.K.U. Gross, *Density Functional Theory* (Springer-Verlag, Berlin, 1990), p. 147.
- ¹¹M. Hybertsen and S. Louie, *Phys. Rev. B* **34**, 5390 (1986).
- ¹²J.D. Eshelby, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press, New York, 1956), Vol. 3.
- ¹³J.J. Wortman and R.A. Evans, *J. Appl. Phys.* **36**, 153 (1965).
- ¹⁴E.g., T.A. Arias and J.D. Joannopoulos, *Phys. Rev. Lett.* **73**, 680 (1994).
- ¹⁵R. Phillips, *Crystals, Defects and Microstructures: Modeling Across Length Scales* (Cambridge University Press, Cambridge, 2001).
- ¹⁶G. Kresse and J. Hafner, *Phys. Rev. B* **47**, R558 (1993); G. Kresse, Ph.D. thesis, Technische Universität, Wien, 1993; G. Kresse and J. Furthmüller, *Comput. Mater. Sci.* **6**, 15 (1996); *Phys. Rev. B* **54**, 11 169 (1996).
- ¹⁷D. Vanderbilt, *Phys. Rev. B* **41**, 7892 (1990).
- ¹⁸G. Kresse and J. Hafner, *J. Phys. (Paris)* **6**, 8245 (1994).
- ¹⁹The error bar is $\delta CGr_N \sqrt{8[(\Delta r_N)^2 + (\Delta r_i)^2]}$.