

# Improving the convergence of defect calculations in supercells: An *ab initio* study of the neutral silicon vacancy

M. I. J. Probert

*Department of Physics, University of York, Heslington, York YO10 5DD, United Kingdom*

M. C. Payne

*Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, United Kingdom*

(Received 17 June 2002; revised manuscript received 11 December 2002; published 13 February 2003)

We present a systematic methodology for the accurate calculation of defect structures in supercells, which we illustrate with a study of the neutral vacancy in silicon. This is a prototypical defect which has been studied extensively using *ab initio* methods, yet remarkably there is still no consensus about the energy or structure of this defect, or even whether the nearest-neighbor atoms relax inwards or outwards. In this paper, we show that the differences between previous calculations can be attributed to supercell convergence errors, and we demonstrate how to systematically reduce each such source of error. The various sources of scatter in previous theoretical studies are discussed and a different effect, that of supercell symmetry, is identified. It is shown that a consistent treatment of this effect is crucial in understanding the systematic effects of increasing the supercell size. This work therefore also presents the best converged *ab initio* study of the neutral silicon vacancy to date.

DOI: 10.1103/PhysRevB.67.075204

PACS number(s): 71.55.Cn, 61.72.Ji, 71.15.Dx, 71.20.Mq

## I. INTRODUCTION

There has been much interest over the past ten years in the calculation of the properties of defects in solids. Various theoretical techniques can be used, but two of the most common are *ab initio* electronic structure calculations of the defect in either a periodic supercell or a cluster. Both of these techniques have advantages and disadvantages. Periodic boundary conditions are a natural representation of a crystal, but the introduction of a defect into the supercell results in the calculation of a periodic array of defects and not an isolated defect. A cluster calculation might therefore seem more appropriate, but this introduces different problems due to finite-size effects and the possibility of interaction between the defect and the surface of the cluster. In this paper, we will focus exclusively on the periodic supercell technique.

There are various technical problems that must be overcome before a supercell calculation becomes an accurate representation of an isolated defect in bulk material. Many of these are already known, but not all are widely appreciated. It is the aim of this paper to synthesize “best practice” into a systematic approach to the study of defects in periodic supercells and show how best to overcome all these problems. As an example of this methodology, we shall consider the neutral vacancy in silicon, which is perhaps the simplest example of a point defect in a crystal. It is certainly one that has been often studied experimentally and theoretically. However, whilst the experimental picture is reasonably clear, there are some properties of the defect, such as its local volume, which cannot be extracted from experimental data and for which reliable theoretical values would be very much welcomed. Unfortunately, there is some confusion about the results of theoretical studies, which has prevented definitive statements being made about these properties. Recent work has achieved some degree of consensus, but there is still a lot of scatter in the results which is not well understood. The aim of this work is to show how a systematic approach to the

study of defects can explain the origin of the scatter in earlier works on the neutral silicon vacancy, although the approach can obviously be applied to the study of any defect. By reducing all the systematic errors to an acceptable level using this methodology, we therefore provide the most highly converged *ab initio* study of the neutral silicon vacancy to date.

We shall perform all our calculations using density-functional theory (DFT) (see Ref. 1 for a review), but the general features of our methodology will be applicable to any *ab initio* simulation technique that uses periodic supercells. In this particular case, we used plane-wave pseudopotential-based DFT, which has been shown in many previous studies to be a reliable technique for calculating many structural properties—typically agreeing with experiment to at least 1% accuracy (see Ref. 2 for a review of total-energy calculations using this technique).

More accurate treatments of the electronic structure exist, such as quantum Monte Carlo (QMC) (see Ref. 3 for a review), but these are much more expensive to apply and are not yet routinely used for such calculations. One reason for this is that whilst QMC is currently “state of the art” in terms of its accuracy in calculating the electronic structure, it is unable to calculate forces and is therefore not a practical methodology for structure determination. There is therefore still a role for DFT calculations to determine structural properties, as well as being a source of trial input wave functions to QMC calculations at the optimal structure.

Ultimately, the results obtained will be limited by the approximations inherent in the *ab initio* technique used. In the case of the neutral silicon vacancy calculation detailed herein, this will include effects due to the choice of pseudopotential, choice of exchange-correlation functional, and the neglect of zero-point motion and thermal effects. However, if these systematic effects are to be quantified, it is important that they are not obscured by random noise arising from other sources of error that can be removed, such as the various kinds of supercell convergence error that are discussed in

this paper. One simple way to estimate these effects would be to compare the DFT and QMC energies at the same defect structure (e.g., the DFT optimal structure). However, the supercell convergence errors discussed herein will still affect the QMC calculation, and so need to be understood and minimized before the comparison of two different *ab initio* methodologies becomes meaningful. The aim of this paper is to show how to systematically reduce the different supercell convergence errors independent of the *ab initio* methodology chosen.

This paper is structured as follows: in Sec. II, we will review what is already known about the vacancy in silicon, in Sec. III we will explain the key features of our methodology in some detail, illustrated with the neutral silicon vacancy calculation. We will report our results for the neutral silicon vacancy in Sec. IV, and will briefly summarize in Sec. V.

## II. REVIEW

The vacancy in silicon is a technologically important defect, as it is known to play an important role in both self-diffusion and impurity diffusion, and hence it is essential to have a detailed understanding of both the electronic and ionic structure of the defect. The vacancy also occurs in a variety of charge states, conventionally referred to as  $V^{2+}$ ,  $V^+$ ,  $V^0$ ,  $V^-$ , and  $V^{2-}$ . It is known that this system shows the *negative-U* effect, that is,  $V^{2+}$  spontaneously converts directly to  $V^0$ . For simplicity, this theoretical work will only focus on the neutral vacancy,  $V^0$ , although in some experimental techniques it is the charged vacancies that are actually studied.

### A. Experimental studies

The experimental studies have been reviewed by Watkins.<sup>4</sup> In summary, electron-paramagnetic-resonance studies can be used to give the symmetry and spatial distribution of the highest unpaired localized electron state. This has shown that the symmetry of the single neutral vacancy  $V^0$  is  $D_{2d}$ .<sup>5</sup> This is understood to be due to the four dangling bonds, created by the removal of a silicon atom from a perfect lattice, hybridizing with each other to form two new levels. These are the  $A_1$  singlet which lies deep in the bulk valence bands and the  $T_2$  triplet which lies in the energy gap. The neutral vacancy has only one of the gap states occupied, which results in a Jahn-Teller distortion, with the ionic relaxation lowering the  $T_d$ -point symmetry of the perfect lattice to that observed in the experiments.

Electron-nuclear double resonance has also been used to study the charged vacancies,<sup>6,7</sup> which in general have lower symmetry than the neutral vacancy considered here. Deep level transient spectroscopy has also been used to give information about the ionization levels associated with charge state changes.<sup>8-11</sup> Positron lifetime measurements have also given information about the defect volume associated with charge state changes.<sup>12,13</sup> However, none of these techniques gives information on the defect volume or formation energy of the  $V^0$  state. This has led to some confusion, with some

theoretical studies claiming an outwards relaxation of the atoms surrounding the vacancy, and others an inwards relaxation.

### B. Theoretical studies

There have been numerous theoretical studies of the silicon vacancy using different theoretical techniques. For example, Green's-function calculations<sup>14-16</sup> predicted an outwards relaxation of the vacancy, whilst more recent tight-binding<sup>17,18</sup> and *ab initio* studies<sup>19-25</sup> have proposed an inwards relaxation. Recent *ab initio* cluster calculations<sup>26</sup> have suggested an inwards-relaxed  $D_{2d}$  symmetry structure, whereas successive *ab initio* supercell calculations using different supercell sizes (from 32 to 216 atoms) and different special  $\mathbf{k}$ -point sampling techniques have yielded a broad spread of formation energies (from 2.6 eV to 4.6 eV) and symmetries (including  $D_{2d}$ ,  $C_{3v}$ ,  $C_{2v}$ , and  $T_{2d}$ ), including some outwards relaxations (although the majority favor inwards relaxation). In particular, Puska *et al.*<sup>25</sup> thoroughly reviewed the previous theoretical studies and also performed a sequence of *ab initio* supercell calculations using different system sizes and sampling techniques. They found a large spread in possible answers, which they attributed to the energy dispersion of the vacancy-induced deep levels, being therefore particularly sensitive to details of the Brillouin-zone sampling scheme used.

Three key quantities of interest are the following:

(1) The vacancy formation energy, which for a neutral vacancy in a supercell is defined as

$$E_V = E_{N-1} - \left(\frac{N-1}{N}\right)E_N, \quad (1)$$

where  $E_N$  is the total energy of the defect-free  $N$  atom supercell, etc.

(2) The symmetry of the defect.

(3) The volume of the defect (actually the tetrahedron formed by the positions of the four atoms  $\mathbf{r}_1, \dots, \mathbf{r}_4$  surrounding the vacancy),

$$V = \frac{1}{6} |(\mathbf{r}_4 - \mathbf{r}_1) \cdot (\mathbf{r}_2 - \mathbf{r}_1) \times (\mathbf{r}_3 - \mathbf{r}_1)|. \quad (2)$$

## III. METHOD

As an illustration of the methodology, we perform what we believe to be the best converged *ab initio* calculation of the neutral silicon vacancy yet undertaken. As discussed in Sec. II, this is not the first time such a calculation has been attempted. However, there has been a lot of scatter in the theoretical calculations, even within the same paper in some instances. We seek to explain the origin of this scatter, and in so doing, produce a definitive answer for the neutral silicon vacancy formation energy and the structure of the lattice relaxation around the vacancy. Our calculations are performed with the CASTEP [Ref. 27] code using the Perdew-Wang (PW91) [Ref. 28] generalized gradient approximation (GGA) for the exchange-correlation functional (which has been shown in many previous defect studies to result in very

accurate structures) with a plane-wave basis set. We use a Vanderbilt ultrasoft pseudopotential<sup>29</sup> for silicon which has been widely used and tested previously, e.g., it gives the cubic lattice constant as 5.440 Å which compares very favorably to the experimental value of 5.429 Å (+0.2%). For simplicity, we have therefore fixed the lattice constants at the value of the experimental lattice constant in all calculations.

As a measure of the reliability of the *ab initio* scheme used, we repeated certain calculations using the same CASTEP code but with three different exchange-correlation functionals—the PW91 GGA as mentioned above, and also the Ceperley-Alder<sup>30</sup> local-density approximation (LDA) and the Perdew-Burke-Ernzerhof (PBE) (Ref. 31) GGA. The same calculations were also repeated with an older norm-conserving pseudopotential<sup>32–34</sup> for silicon that has been part of the standard CASTEP distribution for many years and has been widely tested.

In the following sections, we will describe our methodology and illustrate it with the neutral silicon vacancy calculation for definiteness. Most of what follows can be applied to any supercell calculation, but where there are parts of the discussion which are specific to silicon, these will be clearly highlighted.

Note that it is an often overlooked fact that, as we shall be relaxing the atoms around the defect using forces derived from an *ab initio* calculation, we must ensure that the *ab initio* calculation is fully converged *before* we start to consider any atomic relaxation. That is, we must separately converge the electronic structure at fixed atomic positions, before we can have any confidence in the forces on the atoms being correct. Only then is it appropriate to attempt to converge the atomic relaxation around the defect.

### A. Basis-set size convergence

It is well known that the variational principle ensures that the total ground-state energy of a system will monotonically decrease as the size of the basis set is increased. With some basis sets, it is difficult to systematically improve the quality of the basis set, however, with a plane-wave basis set as used here (which is often used with supercell calculations) this is not a problem. We can associate an energy with each plane-wave basis function ( $|\phi_{\mathbf{g}}\rangle \sim e^{i\mathbf{g}\cdot\mathbf{r}}$ , where  $\mathbf{g}$  is a reciprocal-space lattice vector), and so by using all possible basis functions up to some maximum energy  $E_{cut}$  we may characterize the size of the plane-wave basis set used. Therefore, the basis set may be systematically improved by simply increasing  $E_{cut}$  with a corresponding decrease in the total energy of the system. It is a feature of plane-wave basis sets that typically very large basis-set sizes are required to achieve a reasonable tolerance for the convergence of the total energy of a system, e.g., 1 meV/atom. Therefore, the pseudopotential approximation<sup>29,32–34</sup> is invariably used which enables us to reduce the number of electrons in the problem, and also to reduce the size of the basis set used without affecting the accuracy of the treatment of the electrons outside the core of the atoms.

It is also well known that whilst the total energy of a given system might converge slowly with increasing basis-

set size, the energy difference between two similar systems is much more rapidly convergent due to the cancellation of systematic errors. This is often used to discriminate between two competing atomic arrangements with the same atoms in the same supercell. However, in this work, we shall consider not the energy difference, but the defect formation energy. Note that the variational principle does not apply to such formation energies, and so we are no longer guaranteed monotonic convergence, although this is often seen in practice. We therefore start the calculation by converging the unrelaxed defect formation energy as a function of basis-set size, for a reasonably small system. Everything else is kept fixed, e.g., supercell size, sampling of reciprocal space for the Brillouin-zone integration, pseudopotentials, etc. In the case of the silicon vacancy, we use the vacancy formation energy as defined in Eq. (1). This necessitates calculating the total energy of the vacancy-free system ( $N$  atoms), and the vacancy system ( $N-1$  atoms) which we shall perform with all atoms fixed at the perfect lattice coordinates. We shall then use the same cutoff energy (basis-set size) for all subsequent calculations unless otherwise noted.

For the neutral silicon vacancy, we compare the 16-silicon atom supercell with the 15-atom supercell with vacancy. All atoms are kept at the crystal positions with no relaxation. The Brillouin-zone integration is performed using a  $2 \times 2 \times 2$  Monkhorst-Pack (MP) grid. For this initial part of the calculation, we work with a vacancy formation energy convergence tolerance of 0.01 eV and it is readily shown that with the ultrasoft pseudopotential used, this corresponds to  $E_{cut} \sim 120$  eV.

### B. Brillouin-zone integration convergence

We perform the Brillouin-zone integration using the method of special  $\mathbf{k}$  points. Due to the localized nature of a defect in a (potentially large) supercell, it is important to have a fully converged integration here. This is the basis of the explanation for the difficulty of the calculation as given by Puska *et al.*<sup>25</sup> If we use the simplest Monkhorst-Pack sampling scheme,<sup>35</sup>

$$\mathbf{k} = \frac{2\pi}{a} \left( \frac{q_x}{2q_x^{max}}, \frac{q_y}{2q_y^{max}}, \frac{q_z}{2q_z^{max}} \right) \quad (3)$$

with

$$q_x = \begin{cases} 0, \pm 2, \dots, \pm(q_x^{max} - 1), & q_x^{max} \text{ odd} \\ \pm 1, \pm 3, \dots, \pm(q_x^{max} - 1), & q_x^{max} \text{ even,} \end{cases}$$

etc., then we may easily converge the defect formation energy, at fixed basis-set size and system size, as a function of the density of sample points in reciprocal space.

In order to maximize the separation of the defect from its periodic image, we choose supercells that have the same nearest-neighbor defect-defect distances and sample uniformly in each direction in reciprocal space, and so  $q_x^{max} = q_y^{max} = q_z^{max} = q$ . We may therefore systematically improve the convergence of the Brillouin-zone integration by simply increasing  $q$ . The basic number of special points in the grid is

TABLE I. Effect of increasing the Monkhorst-Pack grid parameter  $q$  on the number of symmetrized points in the grid ( $N_s$ ) and the (squared) radius of exact integration ( $R^2$ ) in units of reciprocal-lattice vectors, for three different supercell symmetries.

Monkhorst-Pack	Simple cubic		Body-centered cubic		Face-centered cubic	
	$N_s$	$R^2$	$N_s$	$R^2$	$N_s$	$R^2$
1	1	1.0	1	0.75	1	0.5
2	1	4.0	2	3.0	2	4.0
3	4	9.0	4	6.75	4	4.5
4	4	16.0	6	12.0	10	16.0

then  $q^3$ . In order to minimize the number of special points at a given value of  $q$  we apply the symmetry operations of the supercell (*not* the point-group symmetry of the defect-free crystal lattice), and therefore work with a weighted set of symmetrized points. The reduction in the number of points depends on the symmetry of the supercell and the value of  $q$ .

However, this will in general lead to very slow convergence, with marked oscillations in both the total energy and the defect energy as  $q$  is increased. This gives rise to the popular belief that “odd  $q$  grids are less efficient than the corresponding even  $q+1$  grid.” However, this is a failure of the implementation of the grid, not the general Monkhorst-Pack method. This can be seen by calculating the radius of exact integration in reciprocal space for different values of  $q$  for some of the most common supercells—simple cubic (sc), body-centered cubic (bcc), and face-centered cubic (fcc) as seen in Table I. For an ideally efficient sampling scheme, the integration should be exact out to a radius given by  $R=q$ , but it can be seen from Table I that this is only achieved for all  $q$  for the simple-cubic supercell, and for even-valued  $q$  for the face-centered cubic supercell.

This flaw was overcome in the basic Monkhorst-Pack scheme by the possible inclusion of a rigid offset  $\mathbf{k}_0$  of the sampling grid from the origin of reciprocal space. This offset is often ignored, but is essential to achieve the full efficiency of the scheme. The use of the optimal offset for a given supercell symmetry and value of  $q$  removes the oscillations in the total energy and consequently accelerates the convergence of the Brillouin-zone integration. A comprehensive set of these optimal offsets was derived by Moreno and Soler<sup>36</sup>

(and independently by M.I.J.P.) but the significance of their work has not been widely appreciated. For completeness, the optimal offsets for the supercells considered in this work are shown in Table II.

We find that the use of a Monkhorst-Pack grid with the optimal offset is more widely applicable and technically superior to other schemes proposed, such as Ref. 37. Using this approach it is now possible to approach convergence of the Brillouin-zone integration in a consistent manner for any value of  $q$  required. Note, that for certain supercells, such as body-centered cubic, the use of offsets is beneficial for all values of  $q$ , whereas for others, such as face-centered cubic, it is only beneficial for odd-valued  $q$ .

When comparing the quality of the Brillouin-zone integration for two different sized or different shaped systems, it is not the value of  $q$  or the number of special points that should be compared, but rather the density of symmetry-unfolded special points in reciprocal space. Note that this comparison will be simplest for two different systems if in each case the sampling scheme is equally efficient (e.g.,  $R=q$ ). Therefore, we recommend the use of offsets at all stages in this methodology when calculating the convergence of the electronic structure.

Therefore, for the neutral silicon vacancy, we converge the Brillouin-zone sampling using the (16/15)-atom fcc supercells and  $E_{cut}=120$  eV. As shown in Table II, there is no advantage in using offsets with even- $q$  grids and fcc supercells but there is a difference for odd- $q$  grids. We therefore perform the odd- $q$  grid calculations twice, once with and once without offsets, and the results of increasing  $q$  on the vacancy formation energy, both with and without offsets, are shown in Fig. 1. There is clearly a dramatic improvement for  $q=1$  where using an offset shifts the sampling away from the  $\Gamma$  point, but it may not appear too dramatic for other values of  $q$  (although it may be hard to see from the figure, in fact the convergence of  $q=3$  with respect to  $q=4$  is significantly improved from  $-0.061$  eV to  $-0.005$  eV). Moreover, Table II suggests that the benefits of using offsets will be most marked with bcc supercells. Therefore, for illustration, we also repeat the Brillouin-zone convergence calculation with the (32/31)-atom bcc supercells and  $E_{cut}=120$  eV, again both with and without offsets. The results are shown as the insert to Fig. 1, and show a marked improvement in convergence with offsets. If we again apply a

TABLE II. Effect of optimal offset  $\mathbf{k}_0$  on maximizing the efficiency of the Brillouin-zone integration for three different supercell symmetries with increasing values of the Monkhorst-Pack grid parameter  $q$ . The number of symmetrized points in the grid ( $N_s$ ) and the (squared) radius of exact integration ( $R^2$ ) in units of reciprocal-lattice vectors is given for each optimal offset.

Monkhorst-Pack	Simple cubic			Body-centered cubic			Face-centered cubic		
	$\mathbf{k}_0$	$N_s$	$R^2$	$\mathbf{k}_0$	$N_s$	$R^2$	$\mathbf{k}_0$	$N_s$	$R^2$
1	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	1	4.0	$(0, \frac{1}{4}, \frac{1}{2})$	1	2.0	$(0, \frac{1}{2}, \frac{1}{2})$	1	1.0
2	$(\frac{1}{8}, \frac{1}{8}, \frac{1}{8})$	3	16.0	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	2	4.0	(0,0,0)	2	4.0
3	$(\frac{1}{4}, 0, \frac{1}{2})$	8	18.0	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	5	9.0	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	6	9.0
4	$(\frac{1}{16}, \frac{1}{16}, \frac{1}{16})$	20	64.0	$(\frac{1}{8}, \frac{1}{8}, \frac{1}{8})$	8	16.0	(0,0,0)	10	16.0

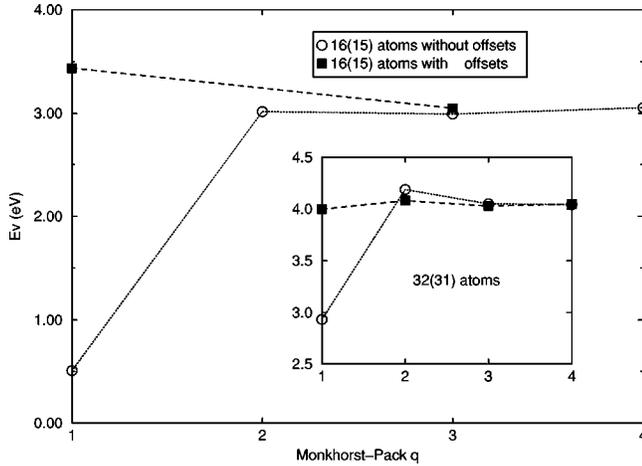


FIG. 1. Convergence of unrelaxed vacancy formation energy with respect to Brillouin-zone sampling, for (16/15)-atom system at  $E_{cut} = 120$  eV. The inset shows the corresponding convergence for the (32/31)-atom system.

vacancy formation energy convergence tolerance of 0.01 eV, it can be seen that this corresponds to a Brillouin-zone sampling density of  $\leq 0.033 \text{ \AA}^{-1}$ . This was therefore used as our sampling density in all subsequent calculations.

### C. Supercell finite-size convergence

Having fully converged the electronic structure calculation for the unrelaxed defect in a given size system, we can proceed to converge the effects of the finite-size supercell. This is the key difference between the supercell and the cluster approaches. With a cluster, we need to minimize the interaction between the defect and the surface of the cluster, but here, with a supercell, we need to minimize the interaction between the defect and its own periodic images. Hence the requirement to converge the supercell size. For insufficiently large supercells, there will be an appreciable overlap between the defect and its own images, resulting in an error in the overall charge density of the system, and hence the total energy and the forces on the atoms. The obvious solution to this is to repeat the defect formation energy calculation in different sized supercells, using an equivalent sized basis set (e.g., same plane-wave cutoff energy) and same Brillouin-zone sampling density.

For the neutral silicon vacancy therefore, we considered all possible sc, bcc, and fcc supercells with between 2 and 256 atoms in the vacancy-free system. The actual value of  $q$  used and the corresponding sampling density are summarized in Table III. The unrelaxed vacancy formation energy, at full Brillouin-zone sampling convergence, for each different supercell is plotted in Fig. 2(a) as a function of the number of atoms in the corresponding vacancy-free system. This is a common way of presenting such information, yet this figure appears confusing, with no obvious trend apparent in the convergence of the vacancy formation energy with system size. However, separating the different points according to the supercell symmetry suggests that there may be a trend, but that this is not the best way to present such data. This is

TABLE III. List of all supercells considered with corresponding supercell symmetry. Also listed is the converged value of  $q$  (Monkhorst-Pack grid parameter) used in the calculation, the Monkhorst-Pack grid offset  $\mathbf{k}_0$  used, and the corresponding Brillouin-zone (BZ) sampling density.

$N$	Symmetry	$q$	$\mathbf{k}_0$	BZ density ( $\text{\AA}^{-1}$ )
2	fcc	8	(0,0,0)	0.040
8	sc	6	(0,0,0)	0.031
16	fcc	4	(0,0,0)	0.040
32	bcc	4	$(\frac{1}{8}, \frac{1}{8}, \frac{1}{8})$	0.033
54	fcc	3	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	0.036
64	sc	3	(0,0,0)	0.031
128	fcc	2	(0,0,0)	0.040
216	sc	2	(0,0,0)	0.031
250	fcc	2	(0,0,0)	0.032
256	bcc	2	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	0.033

because if we simply order the different possible supercells in terms of the total number of atoms (or equivalently, the defect-image distance), we will be misled as the *defect density* will be changing in a nonmonotonic manner. Instead, Fig. 2(b) plots the vacancy formation energy against the defect density, which clearly separates out the different supercell symmetries. This now eliminates the apparent scatter in Fig. 2(a) and instead three clear monotonic trends are evident, one for each supercell symmetry. These trends all appear to converge to the same value,  $\approx 4.40$  eV, in the limit of infinite supercell size (defect density is equal to 0) as would be expected. This therefore explains a common source of the scatter seen between and within the different theoretical studies of the silicon vacancy to date. This effect will obviously also apply to any other supercell defect study.

What then is the origin of the different rates of convergence of the defect formation energy for different symmetry supercells? A simple tight-binding model of nearest-neighbor interactions [with hopping matrix element  $\gamma(a)$ , where  $a$  is the separation of nearest neighbors] is given in many standard texts, e.g., Ref. 38. In this generic model, a band will be formed with a characteristic bandwidth of  $12\gamma$  for sc supercells, and  $16\gamma$  for bcc or fcc supercells with the same defect separation. This can be attributed to the effects of geometry as well as the different number of nearest neighbors in the different supercells. It might therefore be expected that sc supercells were to be preferred in general for defect calculations as they have the least defect-defect interaction (smallest bandwidth) at a given defect separation.

Indeed, some evidence for this is seen in Fig. 2(b), where it can be seen that sc supercells are converging at a faster rate than fcc ones. Indeed, it appears that the (64/63)-atom sc supercell gives a comparable representation of an isolated unrelaxed vacancy to the (250/249)-atom fcc supercell, which can be attributed in part to the number of nearest-neighbor defects. However, this model does not explain why (for the neutral silicon vacancy) the (32/31)-atom bcc supercell gives an even better representation (i.e., the energy is closer to the zero-density limit) than either the (64/63)-atom

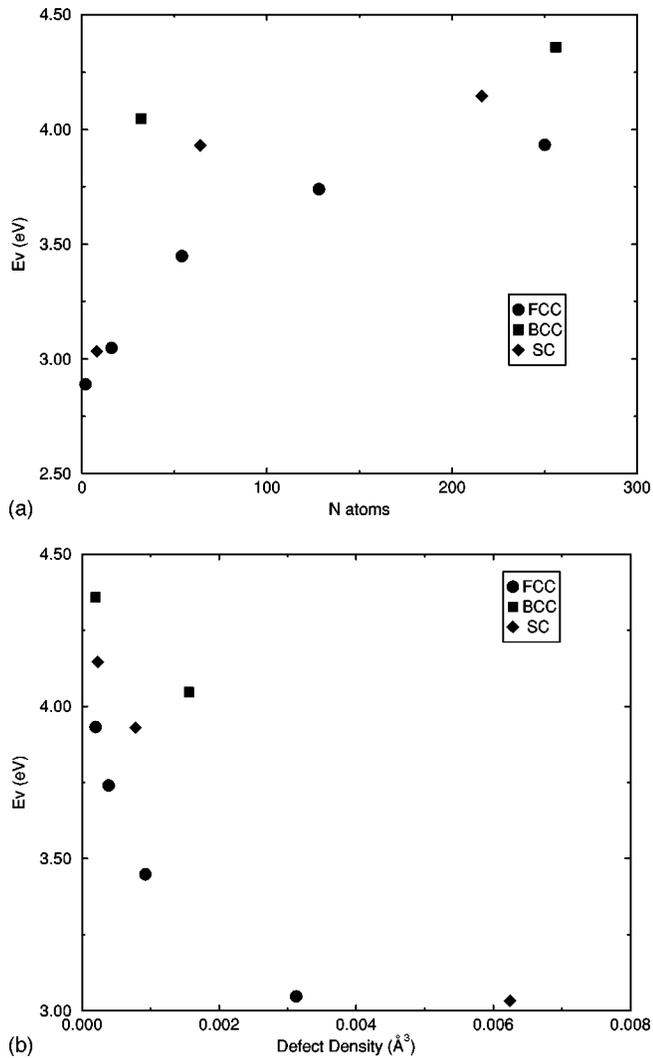


FIG. 2. Variation of unrelaxed defect formation energy at constant basis-set size and Brillouin-zone sampling. (a) shows the variation with system size, and (b) shows the variation with density. The different symmetry supercells are clearly separated in (b).

sc supercell or the (250/249)-atom fcc supercell, which implies that this simple tight-binding model is of limited usefulness.

There must therefore be a further consequence of the supercell symmetry that has not been considered so far. This is

that the defect-defect interaction will have the directionality of the supercell which may or may not be commensurate with the underlying crystal symmetry. This causes a perturbation in the electronic charge density that is another finite-size effect which must vanish in the limit of a sufficiently large supercell. A rigorous analysis of this would involve calculating the density-density response function (see, e.g., Ref. 39). However, the effect becomes apparent if we simply plot the charge-density difference between the unrelaxed vacancy and vacancy-free systems.

For the case of the neutral silicon vacancy, such plots are shown for the (216/215)-atom sc supercell in Fig. 3, the (250/249)-atom fcc supercell in Fig. 4, and the (256/255)-atom bcc supercells in Fig. 5. In the sc supercell we see that there is a localized charge-density difference around the vacancy, and then a longer-ranged component which spans the supercell that is clearly aligned with the  $\langle 100 \rangle$  directions. Similarly, in the fcc supercell the long-ranged component is along the  $\langle 110 \rangle$  directions and in the bcc supercell it is along the  $\langle 111 \rangle$  directions. The normal silicon-silicon bonds are in  $\langle 111 \rangle$  directions which then explains why the bcc supercell is superior for silicon defects—the spurious charge movements caused by the finite supercell size effect are commensurate with the underlying charge density of the system and hence make little difference to the total energy. This is not the case in the fcc and sc supercells where it can be seen that there have been spurious charge movements in the interstitial regions where the charge density is naturally lower, which therefore has a more significant effect. This clearly shows that it is not sufficient to simply increase the size of the system to get a “better” answer, which contributes to the confusion in some earlier studies of the silicon vacancy. The directionality effect of the supercell symmetry will also apply in general to any other defect system, although the detailed considerations will, of course, vary.

Unfortunately, there are only two bcc supercells in the range 2–256 atoms (32 and 256 as in Table III), which would therefore seem to limit our ability to make judgments about the efficacy of bcc supercells for silicon defects. As a further test, the calculation of the unrelaxed neutral silicon vacancy was then repeated for the next bcc supercell, which corresponds to 864 atoms with a defect density of  $0.000\,058\ \text{\AA}^{-3}$ . Again, the same basis-set cutoff, Brillouin-zone sampling density and offset were used. The corresponding unrelaxed

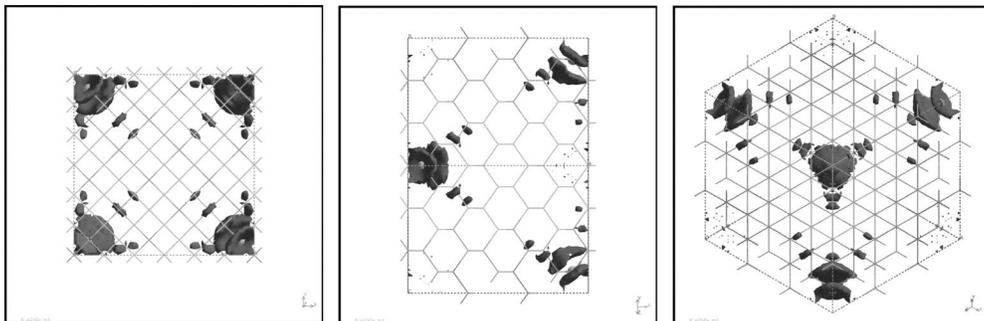


FIG. 3. (Color online) Charge-density difference isosurface at  $\rho = 0.002\ \text{eV}/\text{\AA}^3$  between the unrelaxed 216- and 215-atom sc supercells. Leftmost figure is viewed along the  $[001]$  direction, central figure is along the  $[011]$  direction, and rightmost figure is along the  $[111]$  direction.

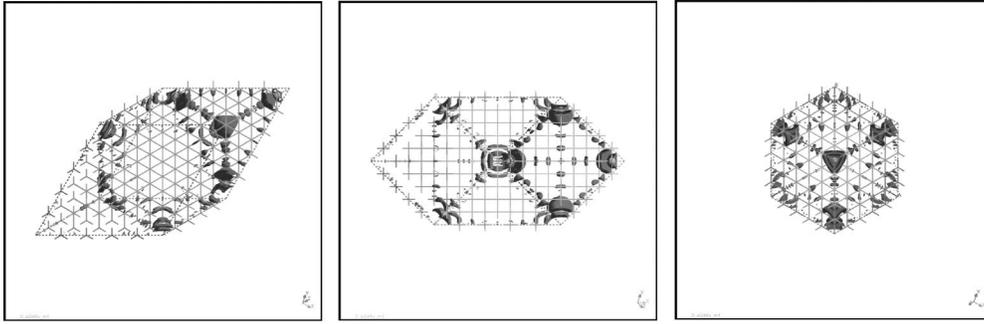


FIG. 4. (Color online) Charge density difference iso-surface at  $\rho=0.002 \text{ eV/\AA}^3$  between the unrelaxed 250- and 249-atom fcc supercells. Leftmost figure is viewed along the [001] direction, central figure is along the [011] direction, and rightmost figure is along the [111] direction.

defect formation energy was 4.358 eV. This confirmed the prediction about the infinite supercell size limit, and shows that the 256-atom bcc supercell has converged (to better than 0.002 eV) the electronic structure of the unrelaxed neutral silicon vacancy with respect to finite supercell size.

In the case of charged defects, the effects of the finite-size supercell will be even more marked due to the long-ranged nature of the Coulomb interaction. Specialized energy correction schemes have been introduced [e.g., Makov-Payne<sup>40</sup>], which accelerate the convergence of the total energy with increasing supercell size.

#### D. Hellmann-Feynman forces convergence

Having finally fully converged all the necessary factors in the unrelaxed defect formation energy, we can now be confident that we have an accurate representation of the ground state electronic wave function. We may now use the Hellmann-Feynman theorem to calculate the forces on the atoms and hence start to relax the defect. However, it must be noted that we converged the basis-set size using an energy difference calculation. The variational principle assures us that the ground-state energy is correct to second-order errors in the ground-state wave function, but the forces will only be correct to first-order errors. Also, as noted previously, an energy difference will converge more rapidly than the total energy.

The advantage of using the defect formation energy criterion in the early stages of this methodology is that it pro-

duces a smaller basis set which makes the other (unrelaxed structure) convergence calculations rapid. This can produce significant savings, as to be sure of convergence it is often necessary to go to one size of calculation beyond that at which convergence first appears, as, for example, when converging the finite-size supercell effect in the neutral silicon vacancy where an 864-atom bcc supercell was evaluated.

In order to produce accurate forces therefore, it is necessary to converge the basis-set size with respect to the forces and so we choose the rms force of the unrelaxed defect structure as a simple scalar parameter to converge. This additional convergence is especially important for defect calculations, as it is often found that the energy surface around a defect is very flat, and so particularly prone to errors in the forces due to the use of underconverged basis sets. This sort of effect can be easily detected by monitoring the direction of the forces on each atom surrounding the defect as the basis-set size is increased. Any tendency for this direction to change significantly is a clear warning that there are serious systematic errors in the forces due to basis-set incompleteness.

An example calculation for the case of the neutral silicon vacancy in the (32/31)-atom bcc supercell is shown in Fig. 6, where it can be seen that whilst from an energy calculation it appears that  $E_{cut}=120 \text{ eV}$  and  $q=2$  is reasonably converged, this is not sufficient for the forces. Applying a criterion that the rms force must be converged to  $0.005 \text{ eV/\AA}$  (which is often used as the convergence tolerance in high quality *ab initio* structural relaxations), we see that  $E_{cut}$

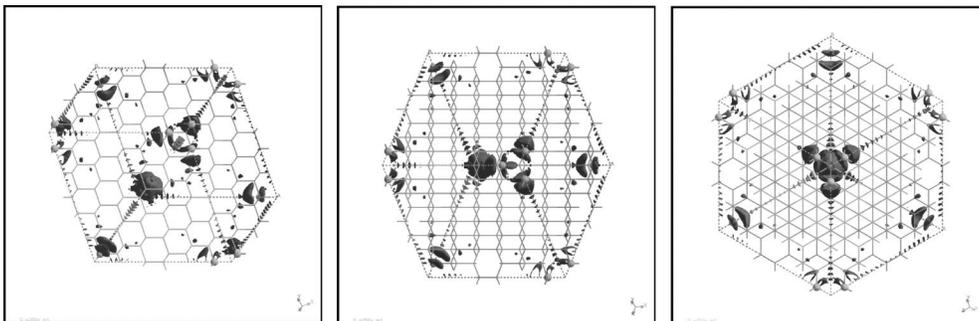


FIG. 5. (Color online) Charge-density difference isosurface at  $\rho=0.002 \text{ eV/\AA}^3$  between the unrelaxed 256- and 255-atom bcc supercells. Leftmost figure is viewed along the [001] direction, central figure is along the [011] direction, and rightmost figure is along the [111] direction.

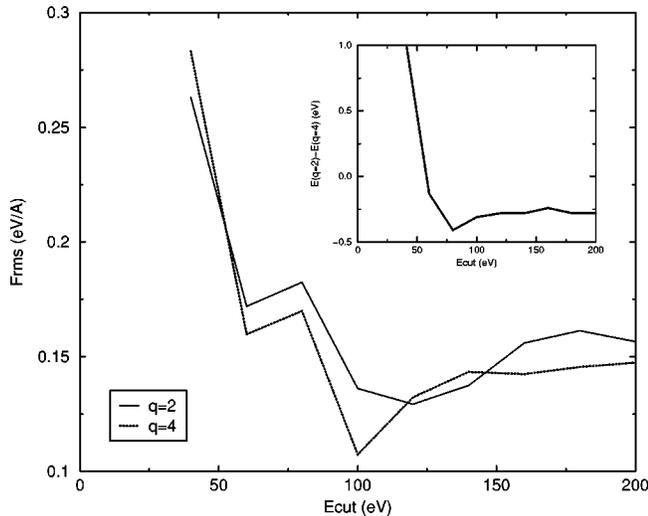


FIG. 6. Variation of rms force in 31-atom silicon cell with respect to basis-set size at two different Brillouin-zone sampling densities, corresponding to different values of the Monkhorst-Pack grid parameter  $q$ . The inset figure shows the corresponding difference in the total energy for the two different values of  $q$ . This clearly shows that whilst it might appear that the total energy is adequately converged at  $E_{cut} = 120$  eV and  $q = 2$ , this is not sufficient for the forces. In all subsequent relaxation calculations, a Brillouin-zone sampling density equivalent to that corresponding to  $q = 4$  in this calculation, and  $E_{cut} = 160$  eV was used.

=160 eV must be used, and that a Brillouin-zone sampling of  $q = 4$  (corresponding to a density of  $0.033 \text{ \AA}^{-1}$ ) must be used.

### E. Atomic relaxation convergence

Finally, we are now ready to relax the atomic structure around the defect, using the forces derived from the systematically converged *ab initio* calculation. We move the atoms according to some minimization algorithm, and stop when we have simultaneously satisfied the various relaxation criteria to prescribed tolerances: For example convergence of the total energy, the rms force, and the rms displacement of the atoms between successive iterations. If we had been simultaneously optimizing the lattice parameters using *ab initio* stresses, then it would be appropriate to also check for convergence of the stress on the supercell.

Note that we are often starting the atomic relaxation from a state of relatively high symmetry. It may therefore be necessary to perturb each atom by a small amount from the symmetry sites at the start of the calculation in order to ensure that symmetry breaking is possible in the relaxation process. Also, because of the possibility of local minima in the structure minimization, the calculation should be restarted several times from different initial arrangements of atoms around the defect (e.g., random symmetry-breaking displacements, directed relaxation inwards, directed relaxation outwards, etc.) in order to be sure that the minimized structure found is indeed the global minimum.

Note that proving that any particular minimum found is indeed the global minimum is a difficult matter. More ad-

vanced techniques, such as simulated annealing<sup>41</sup> or *ab initio* molecular dynamics,<sup>42</sup> are better adapted to exploring the energy surface but at much increased computational cost. In practice, if several independent starting configurations all converge to the same answer, then that is usually sufficient to have a reasonable amount of confidence that the structure found is (a close approximation to) the global minimum.

There is also a popular belief that it is more efficient to relax a structure using a small basis set to get an approximate structure and then to increase the size of the basis set until there is no further change, than to use a sufficiently large basis set throughout. The neutral vacancy in silicon is a counterexample to that belief. If the defect structure is relaxed using too small a basis set (e.g.,  $E_{cut} = 120$  eV), then the systematic errors in the forces cause the vacancy to relax *outwards*. This outwards relaxation is remarkably robust with respect to different perturbations of the surrounding atoms prior to starting the relaxation, including gross inwards and outwards distortions, and the final state is also locally stable with respect to subsequent increases in the basis-set size proving that it is a local minimum. However, if the vacancy is relaxed using a larger basis set ( $E_{cut} \geq 160$  eV) at all times, then the resulting relaxation is *inwards* which illustrates the importance of monitoring the direction of the forces on the unrelaxed atoms surrounding the defect as the basis-set size is increased, as suggested above. This inwards relaxation is also robust with respect to a range of different starting configurations, and the final minimized structure is *lower* in energy than the outwards-relaxed structure at the same basis-set size.

This might appear confusing at first, as the local potential-energy surface around each atom should be quadratic, as silicon at low temperatures is a harmonic crystal to a good approximation—hence the equilibrium geometry ought to be reasonably insensitive to the detail of the calculation. However, this result implies that changing the basis-set size (i.e., reducing the systematic errors in the forces) causes a significant change in the gradient of the potential-energy surface around the unrelaxed defect, i.e., the forces as seen in Fig. 6. So it is actually the boundaries of the different basins of attraction for the relaxation minimizer which are being moved.

This therefore explains another common source of the scatter seen between the different theoretical studies of the neutral silicon vacancy to date, and shows that the only way to relax the defect reliably is to use the larger basis-set size at all stages in the relaxation to reduce the systematic errors in the forces, and to check that the resulting configuration found is the global and not just a local minimum.

When comparing the energetics of two different structures, it should be borne in mind that experiments are usually conducted at finite temperature, whereas energy minimization strategies usually correspond to zero temperature. This means that a true comparison should be based upon free-energies and not just total energies. The entropy difference or the free energy difference between the structures can be obtained by various techniques, such as thermodynamic integration using constrained molecular dynamics (e.g., see Ref. 43 for details). Another complication that arises with finite

temperature is that the presence of nearby local minima will produce significant temperature dependencies in many physical properties, whereupon it then becomes important to know the location of these other minima and the saddle points separating them from the global minimum.

### F. Defect structure convergence

However, even at this stage, there is still one more convergence criterion to meet. The atomic relaxation around the defect may be quite long ranged, and the pattern of relaxations must be contained within the supercell. That is, if we consider successive shells of atoms around the defect (i.e., all those atoms at a common distance from the defect in the unrelaxed structure), then there should be negligible relaxation for atoms beyond a certain distance from the defect, and certainly before the largest shell allowed by the periodic boundary conditions (i.e., half the defect-image separation).

One way to provide an upper bound on the relaxation energy is to perform the atomic relaxation calculation in stages, that is, in the first calculation to only relax those atoms in the first shell around the defect, and then in successive calculations to increase the number of shells allowed to relax, up to the largest allowed shell. Each successive calculation will then provide an improved estimate of the relaxation energy, and allow a simple determination as to whether or not the relaxation has been properly contained within the finite-size supercell. This approach is known as “relaxation under a constant strain field” and is useful for calculating an upper bound on the relaxation energy in a small system, but has the disadvantage that it might result in the system being trained into a local minima which is not the global minimum. Therefore, the best approach is to calculate the relaxation

energy without any constraints—in which case, if the supercell is large enough, there will be negligible relaxation of the largest allowed shells. It is standard practice for relaxation calculations to be repeated for different random perturbations of the atom coordinates, to ensure that the same minimum structure is reached each time.

As a cross-check that the supercell is large enough, and that spurious symmetry effects (e.g., force cancellation in certain directions) have not caused a misleading conclusion, the strain on the supercell should be evaluated, and the volume allowed to relax as appropriate. However, in a fixed volume calculation, there will often be a uniform breathing-mode expansion or contraction of the further-out shells as the underlying lattice accommodates the local relaxation around the defect. This effect will tend to increase the apparent size of the relaxation and cause a volume relaxation that may not be warranted. Therefore, to assess convergence of the defect structure, we consider the relative displacement of successive shells of atoms between the relaxed and unrelaxed defect systems in a fixed volume calculation (as in Ref. 17) and check that this is converged (to some appropriate tolerance) *before* the largest allowed shell allowed by the periodic boundary conditions (i.e., half the defect-image separation).

If it is found that the relaxation is not contained within this largest allowed shell, then the supercell must be increased in size and the above procedure repeated until this is no longer the case. Only then can it be claimed that the calculation is representative of an isolated defect. Of course, this might result in supercell sizes that are impracticable with current computer resources. It has long been recognized that the best way to improve supercell calculations is to use a larger supercell, and for a long time the largest supercell practical for studying the neutral silicon vacancy was sus-

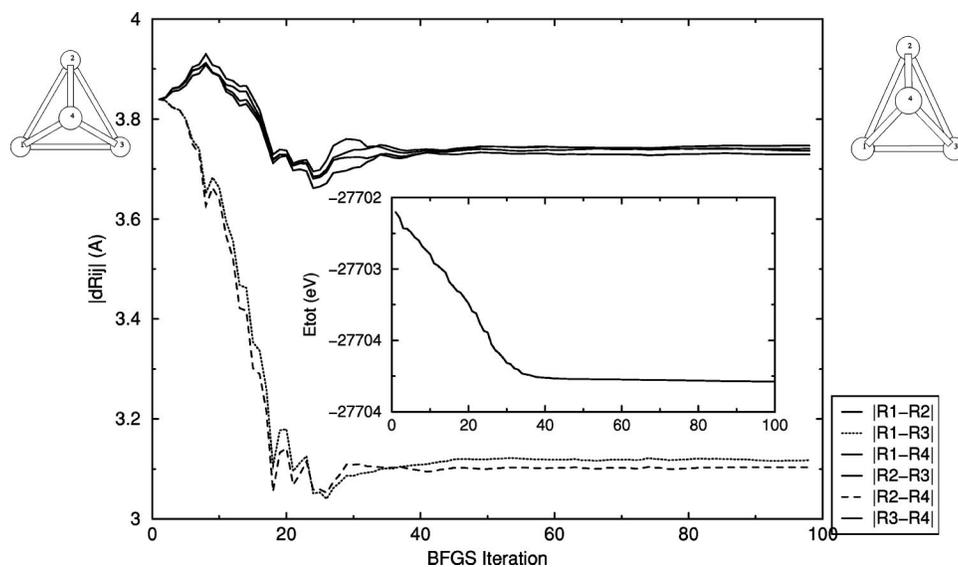


FIG. 7. Relaxation of vacancy using a quasi-Newton minimizer. The six distances between the four silicon atoms surrounding the vacancy are shown. This clearly shows the change in symmetry around the defect, with the initial and final states of the first shell of atoms around the vacancy shown. The atoms are numbered as in the sketches. In the initial state, all bond lengths are equal and the defect has  $T_d$ -point symmetry, whereas in the final relaxed state of the first shell of atoms, there are four equal, longer bond lengths, and two equal, shorter bond lengths, which therefore corresponds to  $D_{2d}$ -point symmetry. Also shown in the inset is the convergence of the total energy of the system as the relaxation proceeds. The relaxation lowers the energy of the system by 1.186 eV.

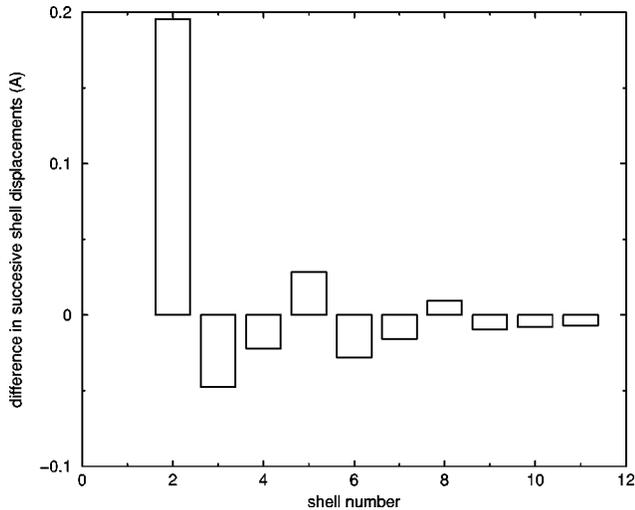


FIG. 8. Convergence of the ionic relaxation of successive shells of atoms across the supercell.

pected to be too small. However, one of the conclusions of this study is that the best way to improve a calculation is not just to increase the supercell size, but to do so in an appropriate manner bearing in mind the interaction of the supercell symmetry with the defect.

For the neutral silicon vacancy we have therefore established the necessary parameters to achieve an accurate energy surface, and so only now do we relax the vacancy in the 255-atom bcc supercell, without using any symmetrization of the electronic parameters (wave function, charge density, forces, etc.) at any stage. This is necessary to ensure that any symmetry in the relaxed structure is spontaneous and not imposed from the initial conditions. Tight convergence tolerances are imposed, namely, that at convergence the rms force be less than  $0.001 \text{ eV}/\text{\AA}$ , the rms displacement be less than  $0.0001 \text{ \AA}$  per iteration, and that the energy difference per iteration be less than  $0.00001 \text{ eV/atom}$ . The results of such a calculation are shown in Fig. 7. This calculation is also repeated for different random perturbations of the atoms in the first shell surrounding the vacancy, to ensure that the same minimum structure is reached each time. To test that the atomic relaxation is contained within the supercell, we calculate the relative displacements of the successive shells of atoms surrounding the vacancy as shown in Fig. 8. From this we can see that shells 9–11 are essentially unchanged (where shell 12 is the halfway point in the supercell), and so we conclude that the ionic relaxation is fully contained within the finite size of the supercell.

#### IV. RESULTS

We now summarize our results for the neutral silicon vacancy.

It was found that when converging the electronic structure of the unrelaxed vacancy, bcc supercells gave superior finite-size supercell convergence, and a (256/255)-atom bcc supercell was required to get satisfactory convergence which was confirmed against an (864/863)-atom calculation. Remarkably, the (32/31)-atom bcc supercell gave an unrelaxed va-

TABLE IV. The unrelaxed vacancy formation energy for the (32/31)-atom bcc supercell, with different exchange-correlation functionals and pseudopotentials.

Scheme	$E_v$ unrelaxed (eV)		
	LDA	PW91	PBE
Ultrasoft, MP $q=2$	4.068	4.106	4.113
Ultrasoft, MP $q=4$	3.995	4.018	4.025
Norm conserving, MP $q=4$	4.016	4.040	4.051

cancy formation energy that was closer to the infinite supercell size limit than that of the (250/249)-fcc supercell calculation. This is attributable to the interaction of supercell symmetry and the symmetry of the underlying silicon lattice.

As a measure of the reliability of the *ab initio* scheme used, we repeated the calculation of the unrelaxed vacancy formation energy in the (32/31)-atom bcc supercell using different exchange-correlation functionals and different pseudopotentials. The results are summarized in Table IV. It has been found many times before that there is a general tendency for LDA-DFT calculations to overbind, and GGA-DFT calculations to underbind. Hence, we conclude that a worst-case error estimate for our *ab initio* scheme is  $\pm 0.02 \text{ eV}$ , but a more likely error estimate is  $\pm 0.01 \text{ eV}$ . It can also be seen that the systematic convergence studies as presented above, such as the Brillouin-zone sampling, can make a more significant change than changing the exchange-correlation functional at a given set of parameters (e.g., going from  $q=2$  to  $q=4$  reduces the formation energy by  $\sim 0.08 \text{ eV}$ ). Of course, as noted in Sec. I, a more thorough comparison would be between DFT and QMC calculations, but at present there are no available QMC data to compare against.

The 255-atom bcc supercell was then used to relax the defect structure, and it was demonstrated that this relaxation was fully contained within the supercell. This relaxation reduced the total energy of the system by  $1.186 \text{ eV}$  and from the observed bond lengths of the four atoms in the first shell surrounding the vacancy, we can see that the final relaxed structure has spontaneously achieved the  $D_{2d}$ -point symmetry, with a final volume [as given by Eq. (2)] that is reduced from the unrelaxed vacancy by  $-27\%$ . The relaxed defect formation energy is therefore estimated as  $3.17 \pm 0.01 \text{ eV}$  (where the error estimate is that due to the *ab initio* scheme used—the convergence error estimate is an order of magnitude smaller). The final parameters used in the calculation and the final result for the structure of the defect are summarized in Table V.

#### V. CONCLUSIONS

We have presented a systematic methodology for the accurate calculation of defect structures in supercells. Various potential pitfalls have been highlighted, and it has been demonstrated how to systematically reduce each source of error in the various convergence parameters, to better than the inherent accuracy of the *ab initio* method used.

As an example of the methodology, the single neutral va-

TABLE V. Final parameters for the fully converged calculation of the neutral silicon vacancy.

Quantity	Value
Number of atoms	256
Symmetry of supercell	bcc
Basis-set size	160 eV
Brillouin-zone sampling density	$0.033 \text{ \AA}^{-1}$
Vacancy formation energy (unrelaxed)	4.36 eV
Vacancy formation energy (relaxed)	3.17 eV
Symmetry of defect (unrelaxed)	$T_d$
Symmetry of defect (relaxed)	$D_{2d}$
Volume of defect (unrelaxed)	$6.671 \text{ \AA}^3$
Volume of defect (relaxed)	$4.874 \text{ \AA}^3$

cancy in silicon has been treated. This has been extensively studied in the past, but with many different answers presented in the literature. The various sources of scatter in previous results have been discussed, such as problems with too small a basis-set size leading to a spurious outwards relaxation as seen in the earlier studies, and problems caused

by underconvergence of the Brillouin-zone sampling leading to inaccurate forces in some more recent studies. The use of offset grids has been shown to be very useful in accelerating the convergence of the Brillouin zone sampling. A different effect, that of supercell symmetry, has been identified, and a consistent treatment of this has been shown to be crucial in understanding the systematic effects of increasing the supercell size. This has resulted in great difficulty in the past with identifying the convergence trends with increasing supercell size, and it is shown herein that the best systematic way to treat this effect is to consider the defect density for each different supercell symmetry separately. Therefore it is believed that this work presents the best converged calculation of the silicon vacancy to date.

### ACKNOWLEDGMENTS

Financial support for M.I.J.P. was provided by the Lloyds of London Tercentenary Foundation. The calculations were performed on the Silicon Graphics Origin 2000 at CSAR, and on the Hitachi SR2201 at the University of Cambridge HPCF.

- <sup>1</sup>R.O. Jones and O. Gunnarsson, *Rev. Mod. Phys.* **61**, 689 (1989).
- <sup>2</sup>M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos, *Rev. Mod. Phys.* **64**, 1045 (1992).
- <sup>3</sup>W.M.C. Foulkes, L. Mitas, R.J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- <sup>4</sup>G. Watkins, in *Deep Centres in Semiconductors*, edited by S. Pantiledes (Gordon and Breach, New York, 1986), p. 147.
- <sup>5</sup>G. Watkins, in *Defects and Their Structure in Non-metallic Solids*, edited by H.B. and H.A.E. (Plenum, New York, 1976), p. 203.
- <sup>6</sup>M. Sprenger, S.H. Muller, and C.A.J. Ammerlaan, *Physica B & C* **116**, 224 (1983).
- <sup>7</sup>M. Sprenger, S.H. Muller, E.G. Sieverts, and C.A.J. Ammerlaan, *Phys. Rev. B* **35**, 1566 (1987).
- <sup>8</sup>G.A. Samara, *Phys. Rev. B* **36**, 4841 (1987).
- <sup>9</sup>G.A. Samara, *Phys. Rev. B* **37**, 8523 (1988).
- <sup>10</sup>G.A. Samara, *Phys. Rev. B* **39**, 11 001 (1989).
- <sup>11</sup>G.A. Samara, *Phys. Rev. B* **39**, 12 764 (1989).
- <sup>12</sup>J. Makinen, C. Corbel, P. Hautojarvi, P. Moser, and F. Pierre, *Phys. Rev. B* **39**, 10 162 (1989).
- <sup>13</sup>J. Makinen, P. Hautojarvi, and C. Corbel, *J. Phys.: Condens. Matter* **4**, 5137 (1992).
- <sup>14</sup>M. Scheffler, *Festkoerperprobleme* **22**, 115 (1982).
- <sup>15</sup>O. Gunnarsson, O. Jepsen, and O.K. Andersen, *Phys. Rev. B* **27**, 7144 (1983).
- <sup>16</sup>M. Scheffler, J.P. Vigneron, and G. Bachelet, *Phys. Rev. B* **31**, 6541 (1985).
- <sup>17</sup>C.Z. Wang, C.T. Chan, and K.M. Ho, *Phys. Rev. Lett.* **66**, 189 (1991).
- <sup>18</sup>L.J. Munro and D.J. Wales, *Phys. Rev. B* **59**, 3969 (1999).
- <sup>19</sup>A. Antonelli and J. Bernholc, *Phys. Rev. B* **40**, 10 643 (1989).
- <sup>20</sup>O. Sugino and A. Oshiyama, *Phys. Rev. Lett.* **68**, 1858 (1992).
- <sup>21</sup>P.E. Blochl, E. Smargiassi, R. Car, D.B. Laks, W. Andreoni, and S.T. Pantelides, *Phys. Rev. Lett.* **70**, 2435 (1993).
- <sup>22</sup>M. Ramamoorthy and S.T. Pantelides, *Phys. Rev. Lett.* **76**, 4753 (1996).
- <sup>23</sup>O. Pankratov, H.C. Huang, T. Diazdelarubia, and C. Mailhot, *Phys. Rev. B* **56**, 13 172 (1997).
- <sup>24</sup>A. Zywietz, J. Furthmuller, and F. Bechstedt, *Mater. Sci. Forum* **258-2**, 653 (1997).
- <sup>25</sup>M.J. Puska, S. Poykko, M. Pesola, and R.M. Nieminen, *Phys. Rev. B* **58**, 1318 (1998).
- <sup>26</sup>S. Ogut, H. Kim, and J.R. Chelikowsky, *Phys. Rev. B* **56**, 11 353 (1997).
- <sup>27</sup>Molecular Simulations Inc., CASTEP4.2 Academic Version, licensed under the UKCP-MSI agreement, 1999.
- <sup>28</sup>J.P. Perdew, in *Electronic Structure of Solids '91*, edited by P. Zeishe and H. Eschrig (Akademie Verlag, Berlin, 1991).
- <sup>29</sup>D. Vanderbilt, *Phys. Rev. B* **41**, 7892 (1990).
- <sup>30</sup>D.M. Ceperley and B.J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).
- <sup>31</sup>J.P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- <sup>32</sup>G.B. Bachelet, D.R. Hamann, and M. Schluter, *Phys. Rev. B* **26**, 4199 (1982).
- <sup>33</sup>D.R. Hamann, *Phys. Rev. B* **40**, 2980 (1989).
- <sup>34</sup>L. Kleinman and D.M. Bylander, *Phys. Rev. Lett.* **48**, 1425 (1982).
- <sup>35</sup>H.J. Monkhorst and J.D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- <sup>36</sup>J. Moreno and J.M. Soler, *Phys. Rev. B* **45**, 13 891 (1992).
- <sup>37</sup>G. Makov, R. Shah, and M.C. Payne, *Phys. Rev. B* **53**, 15 513 (1996).
- <sup>38</sup>N.W. Ashcroft and N.D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976), Chap. 10.
- <sup>39</sup>M. Springborg, *Methods of Electronic-Structure Calculations* (Wiley, Chichester, England, 2000).
- <sup>40</sup>G. Makov and M.C. Payne, *Phys. Rev. B* **51**, 4014 (1995).
- <sup>41</sup>S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi, *Science* **220**, 671 (1983).
- <sup>42</sup>R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).
- <sup>43</sup>D. Frenkel and B. Smit, *Understanding Molecular Simulations* (Academic, San Diego, 1996), Chap. 7.