PHYSICAL REVIEW B VOLUME 53, NUMBER 23 15 JUNE 1996-I

Periodic boundary conditions in *ab initio* calculations. II. Brillouin-zone sampling for aperiodic systems

G. Makov
Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, United Kingdom
and Institut für Energieverfahrenstechnik, Forschungszentrum Jülich GmbH (KFA), Jülich D-52425, Germany

R. Shah and M. C. Payne
Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, United Kingdom

(Received 30 May 1995; revised manuscript received 6 November 1995)

Brillouin-zone sampling in total-energy calculations of aperiodic systems using periodic boundary conditions is considered. Although the energies converge to the exact result in the limit of large supercells for any **k**-point sampling scheme, they do not converge at the same rate. In particular, it is shown that the use of a single sampling point at the origin of reciprocal space is especially inefficient. A **k**-point sampling scheme is proposed, which is computationally efficient and its efficacy relative to other common approaches is demonstrated. [S0163-1829(96)02923-2]

I. INTRODUCTION

Total-energy calculations of periodic systems are generally performed on a single unit cell, using periodic boundary conditions.[1] The eigenstates of such a system can be labeled by the reciprocal-lattice vectors, **k**, in the first Brillouin zone.[2] Since the system is infinite, the quantum numbers **k** are continuous. Calculation of the total energy requires a self-consistent calculation of the potential and the eigenvalues, which are performed at a finite number of points in the Brillouin zone. Procedures for generating optimal sets of **k** points have been proposed and discussed.[3–7] The dependence of the eigenvalues (as well as other quantities) on **k** reflects the chemical binding between atoms in the unit cell under consideration and their counterparts in neighboring cells.[8] In the absence of any interaction, the energy bands would be flat.

Aperiodic systems, such as defects and surfaces, are also often calculated in a periodic arrangement using a supercell. The supercell contains part of the system of interest, which is then periodically repeated in space. If the supercell is chosen to be large enough, the interactions with neighboring cells will be insignificant and the energy calculated will be that of the aperiodic system. The advantages of using periodic boundary conditions for aperiodic systems include easy representation of a condensed-matter environment and the use of numerical schemes developed for periodic structures, in particular, the plane-wave representation.

However, in practice, it is difficult to perform calculations in cells large enough for the total energy of the aperiodic system to be fully converged with respect to cell size. This is because the computational effort increases rapidly with cell size: e.g., for a defect in a solid the effort increases as V^3, where V is the volume of the supercell, and even in the best case, of an isolated atom or molecule, the effort increases as $V\ln V$.[1] If the calculation is performed using a supercell which is too small the calculated energy contains a spurious contribution from interactions between the neighboring supercells. Therefore, it is relevant to consider which calculational scheme leads to results for the total energy that converge fastest with respect to cell size. In a previous paper,[9] we considered this question for purely electrostatic interactions. We showed that for any system, with suitable choice of the energy functional, the calculated energy could be made to converge to its limiting result as $O(L^{-5})$, where L is the linear dimension of the supercell. However, it is possible that in the limit of large cells the interaction is not dominated by the remaining electrostatic interactions (quadrupole-quadrupole, dipole-octupole, etc.), but instead by chemical interactions. The existence of such interactions causes the band structure to be dispersive, with the energy varying through the Brillouin zone associated with the repeated supercell. Therefore, it is necessary to examine **k**-point sampling schemes for performing the integration over the Brillouin zone also for aperiodic systems.

Unlike periodic systems, a converged calculation of an aperiodic system is possible only in a very large supercell. In the limit of an infinitely large supercell, the bands are flat and all points in reciprocal space are equivalent, but for any finite cell, the band structure is dispersive. In determining the set of **k** points to be sampled in a calculation of an aperiodic system, an important consideration is minimizing the number of **k** points, because of the computational effort required for large supercells. Many aperiodic calculations have low or no symmetry and, therefore, the first Brillouin zone is irreducible and the sampling must extend over the entire zone. Therefore, it is preferable to sample at points of high symmetry in the Brillouin zone, in order to reduce the total number of points sampled. These considerations have often led to the choice of single-point sampling at the origin (Γ point), which has the additional advantage of real wave functions. This simple scheme will indeed converge to the exact result in the limit of a large enough supercell, but it is not clear that its rate of convergence will be especially fast. More extensive **k**-point sampling has also been reported for aperiodic calculations, but, to the best of our knowledge, no analysis of **k**-point sampling for aperiodic systems has been reported to date.

In this paper, we consider the question of obtaining a sampling scheme for which computed energy will converge rapidly to the limiting value for the infinitely large supercell. For simplicity we focus on neutral point defects, without dipole moments, in systems with gaps at the Fermi level (i.e., semiconductors or insulators), which are neutral and do not have net dipole moments. We consider the energy-band structure of these systems in the limit of large supercells, using the tight-binding approximation to describe the weak interactions between defects in different cells. From this picture, we show how to obtain simple **k**-point sampling schemes, which efficiently converge the calculated energy with respect to supercell size in the limit of very large cells. In particular, we confirm that the choice of the origin of reciprocal space as a single sampling point has a particularly slow convergence with respect to cell size. For cells which are not very large, we argue that better convergence is obtained by increasing the **k**-point sampling and propose a method for determining the required sampling. We illustrate our approach and its efficacy with several examples.

II. k-POINT SAMPLING

The total energy of the aperiodic supercell containing a defect can be considered as the sum of three contributions: (i) the energy of the underlying lattice, (ii) the interaction of the defect with the lattice, and (iii) the interaction between the defects in neighboring supercells. The latter is just the spurious energy which disappears in the limit of an infinitely large cell. Therefore, for any given **k**-point sampling, the total energy of an aperiodic system will *always* converge to its limiting value in a large enough cell. In contrast, the energy of the underlying lattice is well known to converge for any cell size as the number of **k** points sampled is increased, until a certain (system dependent) **k**-point density is reached. The convergence of the defect-lattice interaction requires both a large cell size, so that the isolated defect is adequately represented, and a sufficiently high **k**-point density, so that the underlying density is also adequately represented. Note that, in contrast to periodic systems, in an aperiodic system it is *not* possible to converge the energy only by increasing the density of **k**-point sampling, but it is also necessary to converge with respect to supercell size.

Consider first the regime where the defect-defect interactions dominate the convergence of the energy. This occurs if the cell size is large relative to the spatial extension of the defect state and if we assume, temporarily, that the lattice-defect interaction is either insignificant or is well converged for any **k**-point sampling scheme in these large cells. For localized defects, the defect interaction can be described in the tight-binding approximation.[8] Defect states, described by Wannier functions, will combine to form a band. The **k** dependence of the band will be of the form

$$\varepsilon(\mathbf{k})=\varepsilon+\int d^3r|\phi(\mathbf{r})|^2\Delta u(\mathbf{r}) + \sum_{\mathbf{R}\neq 0}\int_{\substack{\text{unit}\\ \text{cell}}} d^3r\,\phi(\mathbf{r})\Delta u(\mathbf{r})\phi(\mathbf{r}-\mathbf{R})\cos(\mathbf{k}\cdot\mathbf{R}), \tag{1}$$

where the sum is over the supercell lattice vectors **R**, ϕ are the Wannier functions, Δu is the interaction potential, and ε is a constant. The interaction potential Δu decreases to zero as the supercell size increases. Therefore, the second and third terms on the right-hand side of Eq. (1) vanish as the supercell size becomes infinite, which implies that ε is the energy of interaction of the defect with the lattice. The second term, which is a linear response term, is independent of **k**. This term contains the residual size dependence in the limit of the best **k**-point sampling scheme. The third term is a sum of interactions between overlapping defect states localized on neighboring cells.

To calculate the energy, it is necessary to calculate the integral,[10]

$$E=\frac{V_c}{8\pi^3}\int d^3k\ \varepsilon(\mathbf{k}), \tag{2}$$

which in the limit of finite sampling corresponds to the sum:

$$E=\frac{\sum_i w_i\varepsilon(\mathbf{k}_i)}{\sum_i w_i}, \tag{3}$$

where $\{\mathbf{k}_i\}$ are the sampling points and w_i are the weights. It is obvious that by choosing a set of **k** points, such that all the cosine factors in the sum of $\varepsilon(\mathbf{k})$ are zero, we obtain an optimal sampling scheme in the sense that sampling at additional **k** points will not improve the result. This optimal set of **k** points is infinite. However, the overlaps in the third term in (1) decrease rapidly with increasing distance, so for large cells the dominant terms are the nearest-neighbor terms. Thus, in the limit of large cells, it is necessary to consider only the finite set of **k** points which zero the cosine terms for the nearest-neighbor interactions in the third term in (1). If we use this set of **k** points, then the size dependence of the energy is concentrated in the second term in Eq. (1). Note that, if we sample only at the origin of reciprocal space, we obtain the result

$$\varepsilon(\mathbf{k})=\varepsilon+\int d^3r|\phi(\mathbf{r})|^2\Delta u(\mathbf{r}) + \sum_{\mathbf{R}\neq 0}\int_{\substack{\text{unit}\\ \text{cell}}} d^3r\,\phi(\mathbf{r})\Delta u(\mathbf{r})\phi(\mathbf{r}-\mathbf{R}), \tag{4}$$

which contains all the interactions with neighboring defects. Minimal sampling schemes which give zero-nearest-neighbor interactions for some common lattice types are given in Table I. Note that the optimal sampling schemes are determined by a combination of the lattice geometry and the symmetry of the lattice basis. Furthermore, particular bases can reduce the sampling density even further, e.g., the eight-

TABLE I. Minimal sets of **k** points necessary to zero-nearest-neighbor interactions for several types of lattices, units of $2\pi/a$ where a is the lattice constant. The points are determined in a Cartesian basis. Low-symmetry basis implies no symmetry assumed ($P1$ space group) and high-symmetry basis implies that the basis point group has all the symmetry of the lattice point group. Note that these points are not unique (Ref. 3).

Lattice type (real space)	**k** points for low-symmetry basis	**k** points for high-symmetry basis
Simple cubic	$(0, 0, 0)$,$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$
bcc	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$
fcc	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, $(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ $(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2})$, $(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$

atom cubic cell of the diamond structure has its nearest-neighbor interactions zeroed just by the $(0, 0, 0)$ point or the $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ point, instead of by both points together. This is because of a symmetry-induced ''accidental degeneracy'' in the band structure.

As an illustration of this procedure, consider an isolated atom in a cubic supercell. The atom is a ''defect'' in an empty lattice, the energy of the underlying lattice is always zero, and, therefore, the convergence is dominated by the defect-defect interactions. The nearest-neighbor interactions in Eq. (1) are zero if we sample at the points $(0, 0, 0)$ and $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ in reciprocal space. If we sample at four of the eight points $(\pm\frac{1}{4}, \pm\frac{1}{4}, \pm\frac{1}{4})$, then the contributions from the second and third-nearest neighbors are zero as well. In Fig. 1, we see the results of such calculations for a Mg atom. It is clear that zeroing the second-nearest-neighbor interactions does not cause the energy to converge any faster. For comparison, we also show results for the case of single-point sampling at the origin, where the convergence is much slower. This procedure can also be adopted in the case of a defect in a solid, with equal effectiveness as illustrated in Fig. 2. (In this example, we ignore the convergence of the lattice-defect interaction, because the sampling densities are in the vicinity of those required for convergence to the accuracy shown—ca. 54 **k** points in the primitive unit cell. For further details see below.)

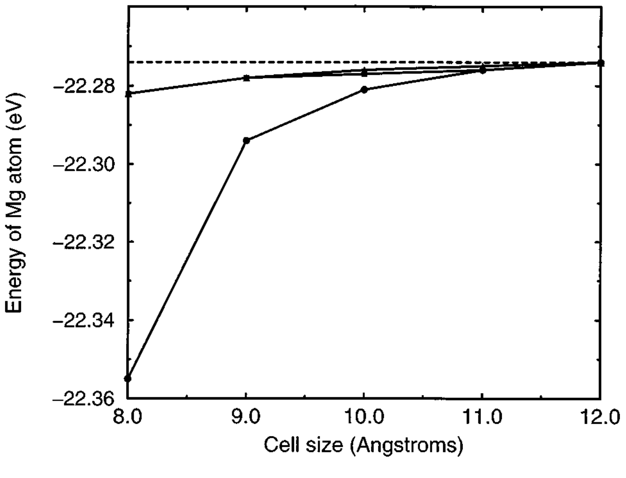

FIG. 1. Convergence of the total energy of Mg with respect to supercell size. For each system, three sampling schemes are shown: Γ point only (circles), Γ point and $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, point (squares), and four $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$-type points (triangles). The horizontal dashed line shows the fully converged result. Cutoff energy of 150 eV.

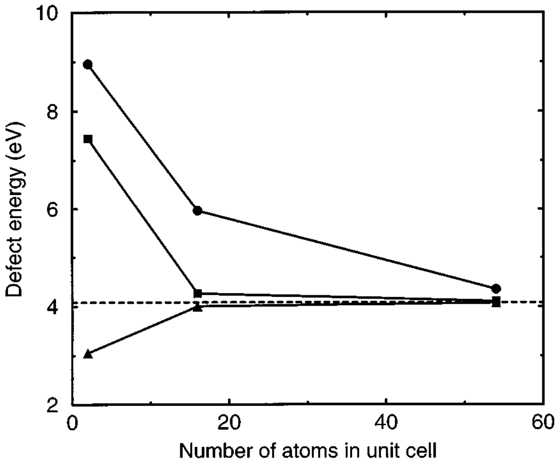

FIG. 2. Convergence of the defect energy with supercell size for three different **k**-point sampling methods. The defect is substitution of an I atom for a Cl atom in a NaCl lattice with CsCl structure. Cutoff energy 250 eV and lattice parameter of 3.05 Å.

Now, we return to consider the convergence of the lattice-defect interaction. For a given cell size, this interaction will converge as the **k**-point sampling is increased [e.g., see Fig. 3(a)]. In principle, one can use a fixed **k**-point sampling scheme (e.g., the nearest-neighbor scheme discussed above) and by increasing the supercell size, cause the energy to converge. An alternative is to increase the **k**-point sampling until the lattice energy converges. The latter is much more efficient computationally if one knows in advance how many **k** points are necessary.

A practical solution to this problem is to first converge the calculation at a fixed cell size with respect to the **k**-point sampling density. The **k** points should be chosen to systematically account for interactions with more distant neighbors. Once the relevant **k**-point density has been determined, one converges the calculation with respect to supercell size at the same, fixed sampling density, i.e., by decreasing the number of **k** points as the cell size is increased. If the cell size is increased to the limit where only a very small number of **k** points is required, then the considerations outlined at the beginning of this section should be employed.

We illustrate this approach by calculating the energy of the ideal vacancy defect in Si. This system has been studied extensively[11–14] and is thought to require very large supercells for convergence, of the order of 200 atoms.[13] The significant dispersion of the defect band is thought to be the dominant factor in limiting the rate of convergence.[11,12] In Fig. 3(a), we present the results for the defect energy in a 16-atom fcc cell, as a function of the **k**-point sampling. One can see that the energy converges as the sampling density is increased. From these results, we deduce that sampling, with a density equivalent to 864 **k** points in the first Brillouin zone of the primitive unit cell, is sufficient to converge the defect energy to an accuracy of better than 0.1 eV. In Fig. 3(b), we show the convergence of the defect energy with respect to cell size at this constant density. We find that using the

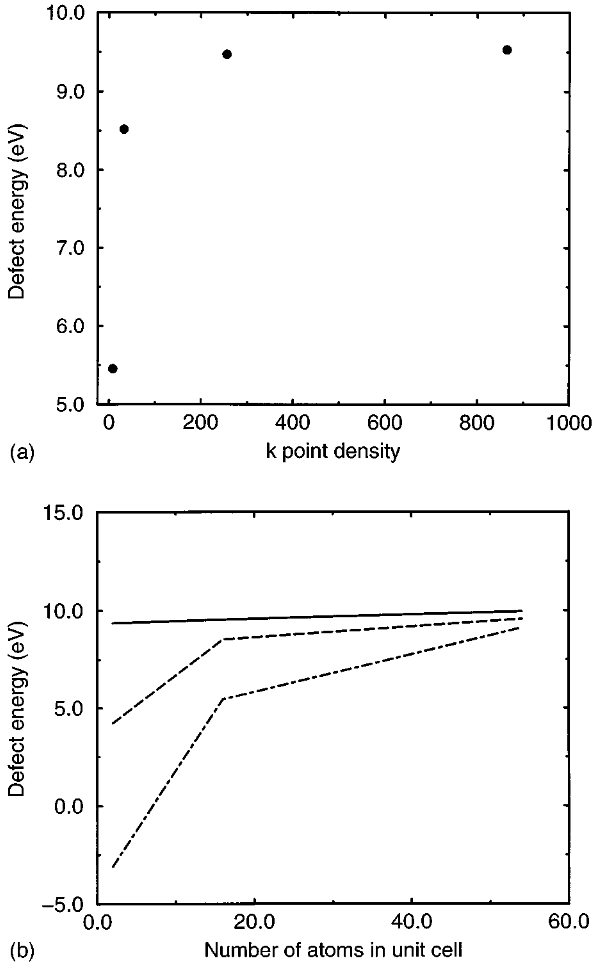

FIG. 3. (a) Convergence of the Si vacancy defect energy (defined as the energy to remove a Si atom to infinity) in a 16-atom fcc cell with respect to the density of **k**-point sampling. Several types of meshes were used following Table III of Ref. 7. Cutoff energy 136 eV and lattice parameter 5.43 Å. (b) Convergence of the vacancy defect, with respect to supercell size at different **k**-point densities. Solid line, 864 **k** points. Dashed line, nearest-neighbor interactions only, according to Table I. Dash-dot line, Γ-point sampling only.

present scheme, the defect energy converges to within ca. 0.5 eV at a cell size of 54 atoms. For comparison, we also show the much slower convergence obtained if only Γ-point sampling is used, or even if nearest-neighbor interactions are accounted for.

III. DISCUSSION

We have considered the question of efficient **k**-point sampling for aperiodic systems calculated in supercells with periodic boundary conditions. We have described a method for determining the **k** points to be sampled, so that overall computational efficiency is achieved, by keeping the density of **k** points fixed at a density, which converges the lattice-defect interaction. As the supercell size increases, the *number* of **k** points decreases. In the limit of large supercells where only very few **k** points are required, we have shown that choosing them to remove the dispersion of the band can considerably accelerate the convergence with respect to cell size. This method combines the two different requirements of such calculations: convergence with respect to supercell size and with respect to **k**-point sampling. Application of our approach was demonstrated to significantly improve convergence with respect to supercell size. The present analysis has concentrated on point defects which are aperiodic in all three space directions, but it can easily be generalized to other cases, e.g., surfaces which are aperiodic in one dimension only.

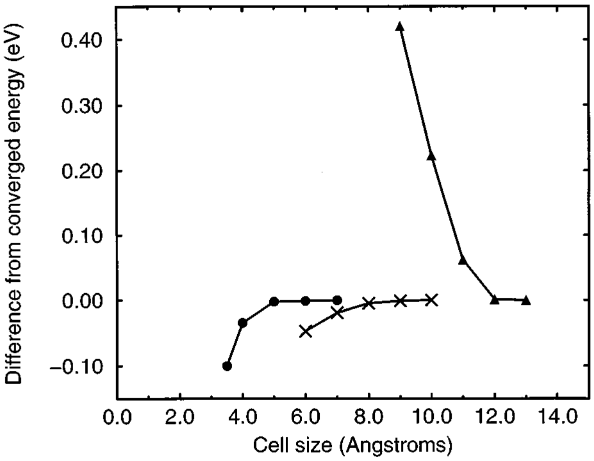

FIG. 4. Relative convergence of the total energy of three different atoms: O (filled triangles), Mg (crosses), and Na (filled circles). The supercell size at which convergence is achieved is correlated inversely with the ionization potential.

The present scheme accounts for the particular nature of the underlying lattice in an exact manner. Alternative schemes have been proposed, in which the energy is not converged with respect to **k**-point sampling at every supercell size, but this is partly corrected for by subtracting an ‘‘equivalently’’ unconverged lattice energy.[15] While such schemes may correct for the lattice energy, they do not account for the error introduced into the defect-lattice term through inaccurate description of the lattice density. This error can clearly be seen in Fig. 3(b).

Our analysis confirms that the common choice of single-point sampling at the origin is highly inefficient,[16] even for systems of low symmetry. In Table I, we have presented better alternatives for some common lattices. One interesting possibility is sampling at the L point for fcc lattices.[17] From Table I, one can see that, for highly symmetric lattice bases, the L point corrects for nearest-neighbor interactions, and, therefore, in the limit of sufficiently large supercells it is an appropriate choice for fcc lattices.

Our previous work[9] on charged and dipolar systems together with the present work shows that total-energy calculations converged to the so-called ‘‘chemical accuracy’’ (less than kT) can be achieved using moderately sized supercells. The size of the supercell at which convergence may be achieved is determined by the spatial extension of the localized defect state. In Fig. 4, we illustrate this for the case of neutral atoms, for which the density decays asymptotically as $\exp[-2(2I)^{1/2}r]/r$, where I is the ionization potential. The total energy of an O atom is seen to converge faster with respect to supercell size than that of a Mg atom, and even faster relative to the Na atom. This ordering reflects the ordering of the ionization potentials. In the case of defects in solids, the spatial extension of the defect state is determined by its location relative to the nearest conduction band.

Finally, we note that, for an isolated point defect, the most efficient calculational scheme is to create a bcc lattice of the defects. In the cases where this is possible, the defects can be maintained at a given separation in approximately 0.8 the volume of the unit cell needed in a simple cubic lattice of defects. For defects in solids, this implies approximately half the computational effort compared to a simple cubic arrangement. Furthermore, only one **k** point is necessary to zero the nearest-neighbor interactions in a bcc lattice, as opposed to two points in a simple cubic lattice (see Table I), which reduces the relative effort by another factor of 2, bringing the total reduction to a factor of 4.

ACKNOWLEDGMENTS

G.M. acknowledges the support of the MINERVA foundation. R.S. acknowledges EPSRC support.

---

[1] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).

[2] V. Heine, *Group Theory in Quantum Mechanics* (Pergamon, Oxford, 1977).

[3] A. Baldereschi, Phys. Rev. B **7**, 5212 (1973).

[4] D.J. Chadi and M.L. Cohen, Phys. Rev. B **8**, 5747 (1973).

[5] H.J. Monkhorst and J.D. Pack, Phys. Rev. B **13**, 5188 (1976).

[6] S. Froyen, Phys. Rev. B **39**, 3168 (1989).

[7] J. Moreno and J.M. Soler, Phys. Rev. B **45**, 13 891 (1992).

[8] N.W. Ashcroft and N.D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1981).

[9] G. Makov and M.C. Payne, Phys. Rev. B **51**, 4014 (1995).

[10] I. Robertson and M.C. Payne, J. Phys. Condens. Matter **2**, 9837 (1990).

[11] O. Sugano and A. Oshiyama, Phys. Rev. Lett. **68**, 1858 (1992), and references therein.

[12] G.A. Baraff and M. Schlüter, Phys. Rev. Lett. **41**, 892 (1978).

[13] W.Y. Ching and M.Z. Huang, Solid State Commun. **57**, 305 (1986).

[14] S.G. Louie, M. Schlüter, J.R. Chelikowsky, and M.L. Cohen, Phys. Rev. B **13**, 1654 (1976).

[15] M. Needels, Phys. Rev. Lett. **71**, 3612 (1993).

[16] D.A. Drabold, P.A. Fedders, O.F. Sankey, and J.D. Dow, Phys. Rev. B **42**, 5135 (1990); D.A. Drabold, P.A. Fedders, A.E. Carlsson, O.F. Sankey, and J.D. Dow, *ibid.* **42**, 5345 (1990).

[17] G. Rajagopal, R.J. Needs, S. Kenny, W.M.C. Foulkes, and A. James, Phys. Rev. Lett. **73**, 1959 (1994).