Physica B 407 (2012) 3063–3067

Contents lists available at SciVerse ScienceDirect

Physica B

journal homepage: www.elsevier.com/locate/physb

Comparison between various finite-size supercell correction schemes for charged defect calculations

Hannu-Pekka Komsa [a,b,*], Tapio Rantala [b], Alfredo Pasquarello [a]

[a] Chaire de Simulation à l'Echelle Atomique (CSEA), Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
[b] Department of Physics, Tampere University of Technology, P.O. Box 692, FI-33101 Tampere, Finland

ARTICLE INFO

Available online 22 August 2011

Keywords:
Defect
Finite-size supercell
DFT

ABSTRACT

We present a comparison of the most common finite-size supercell correction schemes for charged defects in density functional theory calculations. Considered schemes include those proposed by Makov and Payne (MP), Lany and Zunger (LZ), and Freysoldt, Neugebauer, and Van de Walle (FNV). The role of the potential alignment is also assessed. Supercells of various sizes are considered and the corrected formation energies are compared to the values obtained by extrapolation to large supercells. For defects with localized charge distributions, we generally find that the FNV scheme slightly improves upon the LZ one, while the MP scheme generally overcorrects except for point-charge-like defects. We also encountered more complex situations in which the extrapolated values do not coincide. Inspection of the defect electronic structure indicates that this occurs when the defect Kohn–Sham states are degenerate with band-edge states of the host.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Charged defect calculations subject to periodic boundary conditions are plagued by strong supercell-size-dependent errors caused by the interaction of the defect with its periodic images and the compensating background charge [1–7]. Calculations of isolated charged defects would require supercells that are prohibitively large to be treated within density-functional-theory (DFT) calculations. Several schemes have therefore been proposed to correct for these errors. The most common *a posteriori* schemes for correcting the errors in the total energy are those proposed by Makov and Payne (MP) [2], Lany and Zunger (LZ) [6], and Freysoldt, Neugebauer, and Van de Walle (FNV) [7]. However, a careful comparison of the performance and applicability of these schemes has hitherto been missing.

In this paper, we present a comparison of the correction schemes for a fairly extensive set of defect systems. In addition, we investigate the role of the potential alignment, which is integral to all of these schemes. Our study includes three host materials: GaAs, diamond, and MgO, representing materials with small band-gap, covalent bonding, and ionic bonding, respectively. As defects, we examine vacancies, antisites, substitutional impurities, and interstitial impurities, thus covering a large variety of cases. The defects are here considered in their unrelaxed geometries. The quality of the correction schemes are assessed by considering supercells of increasing sizes and by extrapolating the formation energies to the limit of infinitely large supercell.

2. Method

2.1. Formation energy

For the neutral defects, the interaction between the periodic replicas decreases fairly rapidly with increasing supercell size and reliable total energies can be calculated. The defect formation energy is defined as [4]

$$E^f[X^0] = E_{\mathrm{tot}}[X^0] - E_{\mathrm{tot}}[\mathrm{bulk}] - \sum_i n_i \eta_i, \tag{1}$$

where $E_{\mathrm{tot}}[X^0]$ is the total energy of the supercell with the neutral defect X and $E_{\mathrm{tot}}[\mathrm{bulk}]$ is the total energy of the bulk supercell without the defect. The chemical potentials η_i of n_i added atoms allow us to describe various growth conditions.

If the defects are charged, the situation gets more complicated for several reasons. First, a neutralizing background charge (jellium) has to be introduced in order to avoid the divergence of the electrostatic energy. Second, the interaction of the resulting localized charges in the neutralizing background is long-ranged and converges like $1/L$ with respect to the supercell side L. In charged defect calculations, an electron chemical potential μ corresponding to the electron reservoir energy needs to be defined to account for charge variations.

* Corresponding author. Tel.: +41 21 693 5108.
E-mail address: hannu.komsa@epfl.ch (H.-P. Komsa).

0921-4526/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.physb.2011.08.028

The formation energy of the charged defect then becomes

$$E^f[X^q] = E_{\text{tot}}[X^q] + E^q_{\text{corr}} - E_{\text{tot}}[\text{bulk}] - \sum_i n_i \eta_i + q[\mu + E_v + \Delta V] \tag{2}$$

$$E^f[X^q] = E^f[X^0] + E^q_{\text{corr}} + q[\mu + E_v + \Delta V_{0/b}], \tag{3}$$

where μ is given with respect to the valence band maximum (VBM) E_v, obtained in a separate bulk calculation. ΔV is a term used to align the potentials between the supercells of the bulk and the neutral defect. This equation now also includes the finite-size supercell correction E^q_{corr}.

2.2. Potential alignment

The purpose of ΔV in Eq. (3) is to place the VBM energy in the defect supercell at $E_v + \Delta V$. The potential alignment between the supercells of the bulk and the neutral defect can easily be obtained through the comparison of the electrostatic potential far from the defect [4]. Although the electronic structure far from the defect is very similar to that of the bulk, the average electrostatic potential in this region is not necessarily the same in the two supercell calculations. This is due to the convention of setting the Hartree potential (or the total local potential) averaged over the supercell equal to zero in DFT calculations, consequently establishing a connection between the potentials near the defect and far from defect:

$$\int_{\text{near}} V(\mathbf{r})\, \mathrm{d}\mathbf{r} = -\int_{\text{far}} V(\mathbf{r})\, \mathrm{d}\mathbf{r}. \tag{4}$$

A localized change of the potential around the defect leads to a constant shift of the potential in the far region. The potential integral then becomes $\Omega_{\text{far}}\Delta V$, where Ω_{far} corresponds to the volume of the far region within the simulation cell. Hence, the potential alignment term scales like L^{-3} for large L.

2.3. Correction schemes

For large simulation cells, the correction of the total energy of a defect within a supercell can be expressed as

$$E_{\text{corr}} = E_{\text{iso}} - E_{\text{per}}, \tag{5}$$

where E_{iso} and E_{per} correspond to the electrostatic energy in the isolated case and in the periodic case, respectively. E_{iso} is the self-interaction of the isolated charge distribution. E_{per} is the energy of the system subject to periodic boundary conditions and includes both the self-interaction and the interaction with the periodic images and with the background charge. Contrary to atoms or molecules in vacuum, in the case of defects in solids, it is not trivial to distinguish the charge distribution of the defect from that of the host. In addition, the screening response needs to be accounted for. Not only is the potential screened by the host, but the defect charge is surrounded by a screening charge, which comes from the rest of the supercell [6]. Due to these complications, a model for the defect charge and for the dielectric response is needed in order to estimate E_{corr}.

The simplest model would be that of a point charge. In the ideal case of point charges in a neutralizing jellium, the electrostatic energy due to periodic images is the well-known Madelung energy [1]:

$$E_{\text{corr}} = \frac{q^2\alpha}{2\varepsilon L}, \tag{6}$$

where ε is the dielectric constant and α the Madelung constant. For a single point charge, α depends only on the Bravais lattice. For example, α is equal to 2.8373, 3.6392, and 4.5848, for sc, bcc, and fcc, respectively. The Madelung energy for any Bravais lattice can easily be calculated through the Ewald method.

Makov and Payne considered a localized charge distribution in vacuum [2]. Considering the dielectric constant of the solid, they then found the leading terms of the correction to be

$$E_{\text{corr}} = \frac{q^2\alpha}{2\varepsilon L} - \frac{2\pi q Q}{3\varepsilon\Omega}, \tag{7}$$

where

$$Q = \int \rho_d(\mathbf{r}) r^2\, \mathrm{d}\mathbf{r} \tag{8}$$

is the second radial moment of the defect charge distribution ρ_d [2]. In the following we will denote the first and second terms of this scheme as MP and MPQ.

The MP scheme is not concerned with difficulties associated to the screening. Indeed, the MP scheme, at least when only the first term is used, has been found to overcorrect [5]. In response, Lany and Zunger (LZ) proposed to use the charge difference between the charged and neutral defect to evaluate the second radial moment for the MPQ term [6]. In addition, the alignment between the potentials of the bulk and the neutral defect is applied. The charge difference away from the defect primarily consists of a delocalized screening charge of density $n_s = (1-\frac{1}{\varepsilon})q/\Omega$ [6,8], which dominates the second radial moment and gives (for a cubic cell)

$$Q = \int n_s r^2\, \mathrm{d}\mathbf{r} = -\frac{q}{4}L^2\left(1-\frac{1}{\varepsilon}\right). \tag{9}$$

This then gives an energy contribution of

$$\Delta E = -\frac{\pi q^2}{6\varepsilon L}\left(1-\frac{1}{\varepsilon}\right). \tag{10}$$

According to this viewpoint, both MP and MPQ terms have the same L^{-1} scaling and can be combined to a single correction term. For a large ε, this corresponds to the single MP term scaled by a factor of 0.631 [6,8].

The long-range behavior of the interaction potential between image charges scales like $1/r$, leading to an interaction energy that should be fairly independent of the charge distribution as long it is sufficiently localized. However, as discussed above, the changes in the potential near the defect lead to a potential shift far from the defect, which affects the energy E_{per}. A scheme for obtaining the energy and the alignment in a consistent manner was proposed by FNV in Refs. [7,9]. The correction is given as

$$E_{\text{corr}} = E_{\text{lat}} + q\Delta V_{q/0}. \tag{11}$$

The lattice-energy E_{lat} corresponds to the energy correction [cf. Eq. (5)] associated to a model charge distribution for the defect. For a point-charge model, E_{lat} reduces to the MP term. The potential alignment $\Delta V_{q/0}$ in Eq. (11) is obtained by comparing the potential from the model charge far from the defect to the potential difference between the charged and the neutral defect calculations (neutral-defect reference) [7,9]. Alternatively, one could determine the potential alignment $\Delta V_{q/b}$ by comparing the potential from the model charge to the potential difference between the charged defect and the bulk calculations (bulk reference) [7,9]. In this case, there is no need to consider the potential alignment in Eq. (3) explicitly, since it is now included in $E_{\text{corr}} = E_{\text{lat}} + q\Delta V_{q/b}$, where $V_{q/b} = V_{q/0} + V_{0/b}$. We note that, unlike the other schemes, the possibility of choosing the bulk reference in the FNV scheme allows one to apply the potential alignment without considering the neutral defect. When the defect charge distribution and the dielectric screening of the host are well described, $\Delta V_{q/0} \approx 0$ and the FNV scheme reduces to $E_{\text{corr}} = E_{\text{lat}}$.

2.4. Computational details

In all calculations, we use the plane-wave DFT code VASP within the formulation based on partial augmented waves

[10,11]. The exchange correlation energy is described within the local density approximation. We perform calculations in three host materials: zinc-blende GaAs, diamond, and cubic MgO. These represent a small band-gap semiconductor, a large band-gap covalent system, and a system with ionic bonding. In the GaAs calculations, we use a Ga pseudopotential with 3*d* electrons included in the core and an energy cutoff of 400 eV. In the diamond and the MgO calculations, the energy cutoff is 500 eV. We use 64-atom supercells with $4\times4\times4$ **k**-points, 216-atom supercells with $2\times2\times2$ **k**-points, and 512-atom supercells with $2\times2\times2$ **k**-points, without including the Γ-point in any of the **k**-point meshes. Such meshes generally improve the convergence of the total energies. We use equilibrium lattice constants and dielectric constants consistently calculated within our electronic-structure scheme. Our calculations give lattice constants of 3.536, 5.627, and 4.155 Å and high-frequency dielectric constants of 5.76, 13.7, and 3.10, for diamond, GaAs, and MgO, respectively.

For all FNV corrections, we use a model charge consisting of a Gaussian with a width of 2 bohr. We generally use the bulk reference and estimate only the combined alignment term $V_{q/b}$. For the structurally relaxed defects, the DFT potential difference is smoothened by a suitable Gaussian convolution.

In all figures, the formation energies are plotted as a function of inverse supercell side L^{-1}. In this representation, ideally corrected values should exhibit horizontal flat lines. In addition, we perform a regression of the calculated formation energies to the form $aL^{-1}+bL^{-3}+c$ [12,5], which gives an extrapolated value c in the limit of infinitely large supercells, $L^{-1}\to0$. To the extent that the potential alignment term scales like L^{-3}, the MP, the LZ, and the uncorrected results extrapolate by construction to the same value. When the FNV method also extrapolates to the same value, we assume that the extrapolated values correspond to the limit of isolated defect.

3. Results

We first consider defects which essentially behave like point charges with respect to the bulk reference: the interstitial hydrogen defect and the substitutional nitrogen defect in their respective $+1$ charge state in diamond. Their formation energies are shown in Fig. 1. All the curves extrapolate to approximately the same value, which indicates that the adopted supercells are sufficiently large and that the extrapolated value can be taken as a benchmark. In this case, the correction schemes that work best are the MP (with or without potential alignment) and the FNV ones. For these defects, the MP correction is indeed expected to perform well. The FNV scheme which assumes a Gaussian charge rather than a point charge is able to correct itself through the inclusion of the $q\Delta V_{q/0}$ term. Despite the treatment of the screening charge, LZ does not perform as well. However, all correction schemes improve upon the uncorrected values either with or without potential alignment.

We now consider defect charge distributions which cannot be described as point charges. We additionally impose that the defect Kohn–Sham orbitals are well localized and that their levels always fall within the band gap. This kind of defect is here illustrated through two examples, the -2 charge state of the vacancy in diamond and the $+1$ charge state of the O vacancy in MgO. Their formation energies vs. L^{-1} are given in Fig. 2. For both defects, all the corrections schemes give close extrapolated values and all improve upon the uncorrected results. At finite L, the MP scheme generally overcorrects, as noted previously [5]. The FNV scheme gives the best corrections, but the improvement over the MP and LZ schemes is only slight.

For certain defects, the correction schemes fail in achieving good estimates for the formation energies at finite L. As a first example, we focus on the $+2$ charge state of the vacancy in diamond, for which the formation energies are given in Fig. 3(a). The extrapolated values obtained with the various correction

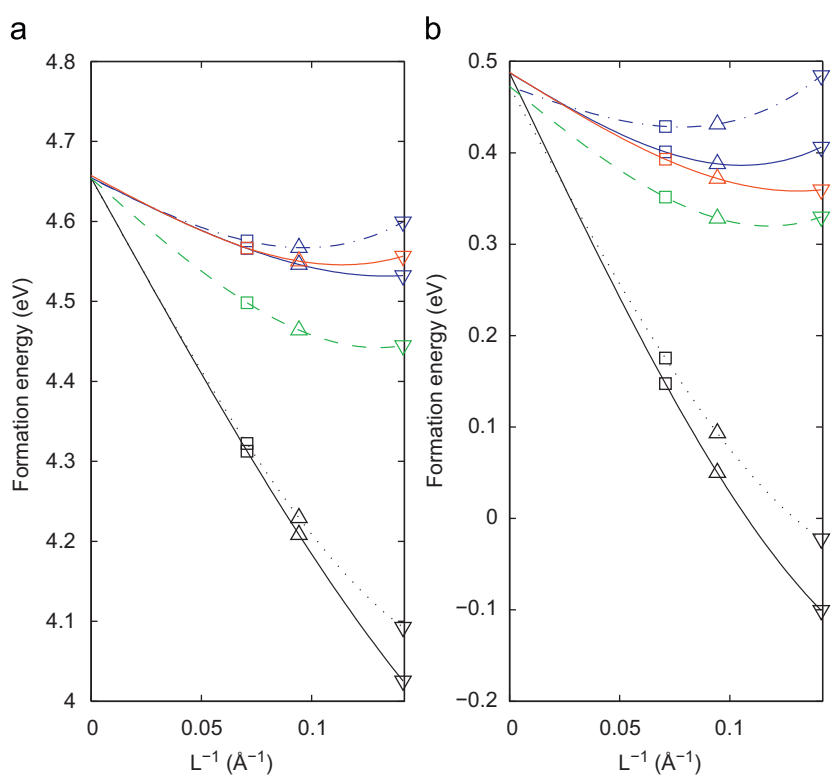

Fig. 1. Formation energies of the unrelaxed (a) H interstitial and (b) N substitutional defects in the $+1$ charge states in diamond: uncorrected (black, solid), potential alignment (black, dotted), MP (blue, solid), MP with potential alignment (blue, dash-dotted), FNV with bulk reference (red, solid), LZ (green, dashed). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

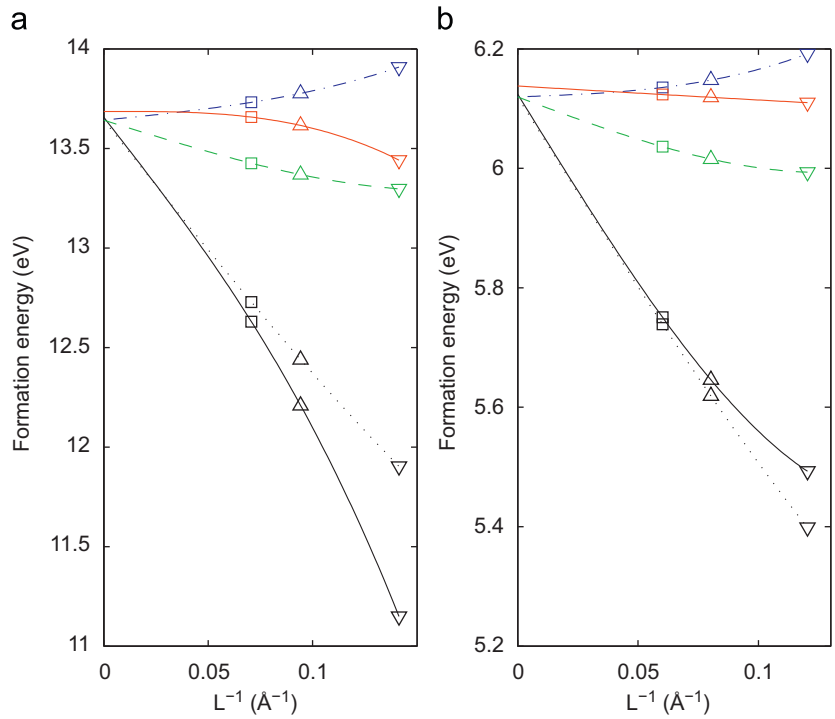

Fig. 2. Formation energies of (a) the −2 charge state of the unrelaxed vacancy in diamond and (b) the +1 charge state of the O vacancy in MgO. Same notation as in Fig. 1.

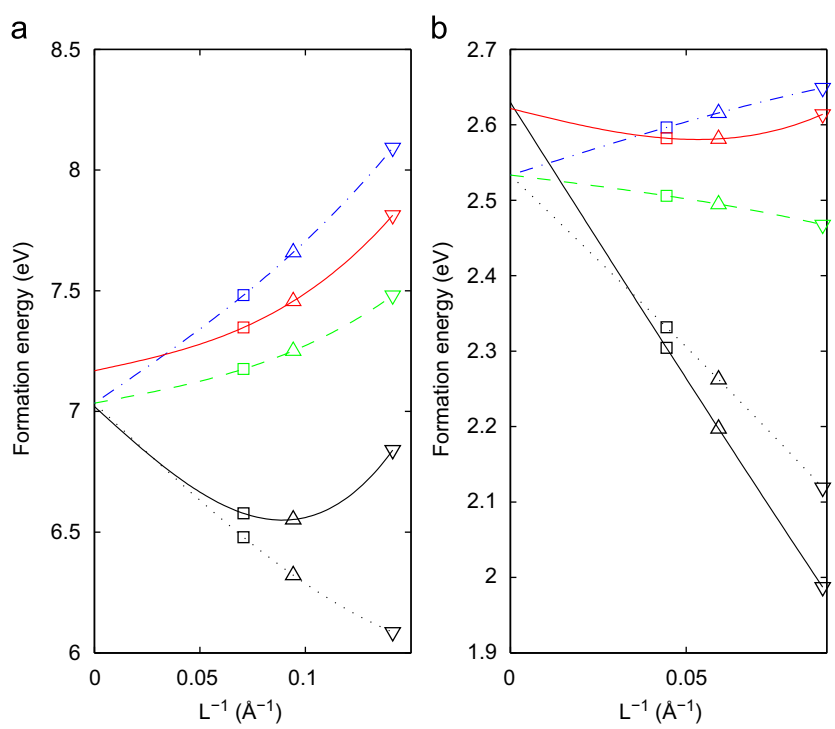

Fig. 3. Formation energies of (a) the +2 charge state of the unrelaxed vacancy in diamond and of (b) the +2 charge state of the unrelaxed As antisite in GaAs. Same notation as in Fig. 1.

schemes fall within a range of several tenths of eV. This range is slightly wider than found in the previous cases, hinting to an underlying complication. The difficulties likely arise from the fact that the Kohn–Sham states move closer to and eventually become resonant with the band states of the host as the supercell size increases, as confirmed by inspection. For this defect state, the LZ scheme appears to give the smallest deviations from its respective extrapolated value. However, the results achieved within the FNV scheme could to some extent be improved through the consideration of a more extended model charge distribution for the defect, as illustrated in Refs. [7,9].

As a second example with underlying complications, we discuss the +2 charge state of the As antisite in GaAs [Fig. 3(b)]. Also in this case, the extrapolated values span an extended range.

However, when focusing in particular on the uncorrected formation energies, we notice that the sole application of the potential alignment leads to a different extrapolated value. Since the potential alignment only involves the neutral defect state, this effect can readily be correlated to Kohn–Sham states of the neutral defect being resonant with the band states of the host. Given the absence of a clear extrapolated value, it is difficult to assess the accuracy of the various schemes in this case. However, we note that the FNV results are not subject to the present complications involving the neutral defect state, since we adopted the bulk reference.

4. Conclusions

We compared the performance of various finite-size correction schemes for a set of charged defects in various hosts. We distinguished different kind of defects. The first type of defect closely corresponds to a point charge. In this case, the MP and FNV schemes are found to give slightly better results than the LZ scheme. The second type of defect has a more extended charge distribution, but its electronic structure is characterized by well localized Kohn–Sham defect states with associated energy levels falling in the band gap. For such defects, the MP scheme is found to overcorrect, while the FNV and LZ schemes give corrections of comparable accuracy. We also identified more complex situations in which the Kohn–Sham defect states are resonant with the band states of the host. In these situations, it is more difficult to achieve reliable extrapolated values and the accuracy of the various schemes cannot be assessed in a definite way. However, we remark that the FNV scheme offers extra flexibility in coping with such complications through the choice of model charge distribution of the defect and through the choice of reference for the alignment.

Acknowledgments

Financial support is acknowledged from the Swiss National Science Foundation (Grants nos. 200020-119733/1, 200020-134600, and 206021-128743). We used computational resources of CSC (Finland) and CSEA-EPFL.

References

[1] M. Leslie, M.J. Gillan, J. Phys. C: Solid State Phys. 18 (1985) 973.
[2] G. Makov, M.C. Payne, Phys. Rev. B 51 (1995) 4014.
[3] P.A. Schultz, Phys. Rev. Lett. 584 (2000) 1942.
[4] C.G. Van de Walle, J. Neugebauer, J. Appl. Phys. 95 (2004) 3851.
[5] C.W.M. Castleton, A. Höglund, S. Mirbt, Phys. Rev. B 73 (2006) 035215.
[6] S. Lany, A. Zunger, Phys. Rev. B 78 (2008) 235104.
[7] C. Freysoldt, J. Neugebauer, C.G. Van de Walle, Phys. Rev. Lett. 102 (2009) 016402.
[8] S. Lany, A. Zunger, Model. Simul. Mater. Sci. Eng. 17 (2009) 084002.
[9] C. Freysoldt, J. Neugebauer, C.G. Van de Walle, Phys. Status Solidi B 248 (2011) 1067.
[10] G. Kresse, J. Furthmüller, Comput. Mater. Sci. 6 (1996) 15.
[11] G. Kresse, J. Furthmüller, Phys. Rev. B 54 (1996) 11169.
[12] C.W.M. Castleton, S. Mirbt, Phys. Rev. B 70 (2004) 195202.