



CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Finite-size supercell correction schemes for charged defect calculations

Hannu-Pekka Komsa, Tapio T. Rantala, and Alfredo Pasquarello

Phys. Rev. B **86**, 045112 — Published 13 July 2012

DOI: [10.1103/PhysRevB.86.045112](https://doi.org/10.1103/PhysRevB.86.045112)

# Investigation of finite-size supercell correction schemes for charged defect calculations

Hannu-Pekka Komsa<sup>1</sup>, Tapio T. Rantala<sup>2</sup>, and Alfredo Pasquarello<sup>1</sup>

<sup>1</sup>*Chaire de Simulation à l'Echelle Atomique (CSEA),*

*Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*

<sup>2</sup>*Department of Physics, Tampere University of Technology, FIN-33101 Tampere, Finland*

(Dated: June 19, 2012)

Various schemes for correcting the finite-size supercell errors in the case of charged defect calculations are analyzed and their performance for a series of defect systems is compared. We focus on the schemes proposed by Makov and Payne (MP), Freysoldt, Neugebauer, and Van de Walle (FNV), and Lany and Zunger (LZ). The role of the potential alignment is also assessed. We demonstrate a connection between the defect charge distribution and the potential alignment, which establishes a relation between the MP and FNV schemes. Calculations are performed using supercells of various sizes and the corrected formation energies are compared to the values obtained by extrapolation to infinitely large supercells. For defects with localized charge distributions, we generally find that the FNV scheme improves upon the LZ one, while the MP scheme tends to overcorrect except for point-charge-like defects. We also encountered a class of defects, for which all the correction schemes fail to produce results consistent with the extrapolated values. These are found to be caused by partial delocalization of the defect charge. We associate this effect to hybridization between the defect state and the band-edge states of the host. The occurrence of defect charge delocalization also reflects in the evolution of the defect Kohn-Sham levels with increasing supercell size. We discuss the physical relevance of the latter class of defects.

PACS numbers: 61.72.J-,71.15.-m,71.55.-i

## I. INTRODUCTION

The modeling of point defects has become an invaluable tool for understanding materials properties in many areas of applied physics and materials science.<sup>1</sup> Due to the typically low concentration of defects in crystalline materials, it is often of interest to consider the properties of a single defect in an otherwise pristine crystal. Periodic boundary conditions are ideal for the description of pristine crystals and are thus often employed in density-functional-theory (DFT) calculations. However, when defects are introduced in the computational cell, the cell size has to be sufficiently large to neglect the interactions between the periodic images of defects. In the case of charged defects, the strong and long-ranged Coulomb interaction between the localized charge distributions converges very slowly. Consequently, the supercell sizes required to yield converged energies become prohibitively large.

Several approaches have been proposed over the years to overcome this problem. The straightforward approach is to perform calculations for supercells of varying size and to extrapolate to the limit of infinitely large supercell.<sup>2-5</sup> When this is excessively demanding from the computational point of view, one generally resorts to using small supercells and to adopting correction schemes. Apart from the local moment countercharge method which accounts for the multipole moments of the defect within the self-consistent electronic-structure cycle,<sup>6</sup> the most common approaches consist in correcting *a posteriori* the calculated formation energies.<sup>5-11</sup> However, a critical examination of the performance and applicability of these schemes has hitherto been missing.

In this work, we critically examine the performance of

various commonly used *a posteriori* correction schemes when applied to the formation energies of a large set of defects within different host materials. We consider the schemes proposed by Makov and Payne,<sup>8</sup> Freysoldt, Neugebauer, and Van de Walle,<sup>10</sup> and Lany and Zunger.<sup>9</sup> We aim at showing the mutual relation between the various schemes and at defining the conditions for their application. In order to assess the quality of the correction schemes, we consider supercells of increasing sizes and extrapolate the formation energies to the limit of infinitely large supercell. Our study includes three host materials: diamond, GaAs, and MgO, representing materials with covalent bonding, small band-gap, and ionic bonding, respectively. Within these hosts, we examine vacancies, antisites, substitutional impurities, and interstitial impurities, thus covering a large variety of cases. We consider both unrelaxed and relaxed defect geometries. In addition, we examine the underlying reasons which lead to an unsatisfactory convergence behavior for some defect systems. In particular, we highlight the role of defect charge delocalization and illustrate its relation with the band-edge positions of the host material.

This paper is organized as follows. A thorough description of the schemes is given in Sec. II. We also provide an analysis which gives a comprehensive description of defect corrections, connecting the considered schemes to each other and to other related work in the literature. The results of the calculations and the comparison of the correction schemes are presented in Sec. III. In particular, we distinguish two classes of defects showing different charge-localization behavior. The conclusions are drawn in Sec. IV.

## II. DEFECT CALCULATIONS

### A. Formation energy

The defect calculation is traditionally approached through the concept of formation energy,<sup>11,12</sup> which in the case of neutral defects can be written as

$$E^f[X^0] = E_{\text{tot}}[X^0] - E_{\text{tot}}[\text{bulk}] - \sum_i n_i \mu_i \quad (1)$$

where  $E_{\text{tot}}[X^q]$  is the total energy of the supercell with the defect  $X$  in charge state  $q$  and  $E_{\text{tot}}[\text{bulk}]$  the total energy of the non-defective bulk supercell. For each species  $i$ , the chemical potential  $\mu_i$  of the  $n_i$  added atoms allows us to describe various growth conditions. When the defects are charge neutral, i.e., in absence of a charge monopole, the interactions between defects converge rapidly with supercell cell size, and fairly reliable estimates for the formation energies of isolated defects can be obtained from relatively small supercell calculations.

When the defects are charged, the situation is more complicated for several reasons. First, the periodically repeated defect system effectively includes a homogeneous neutralizing background charge (jellium) which ensures that the electrostatic energy per unit cell remains finite. Second, as the charge of the defect system no longer matches that of the neutral reference systems, the formation energy can properly be defined only through the introduction of a chemical potential for the electrons.<sup>11,12</sup> This brings an additional term to the formation energy:

$$E^f[X^q] = E_{\text{tot}}[X^q] + E_{\text{corr}}^q - E_{\text{tot}}[\text{bulk}] - \sum_i n_i \mu_i + q[\epsilon_F + \epsilon_v + \Delta v_{0/b}] \quad (2)$$

where  $\epsilon_v$  is the valence band maximum (VBM) as calculated in the non-defective bulk supercell and  $\epsilon_F$  is the Fermi energy (the electron reservoir energy), which is customarily given with respect to the VBM.  $\Delta v_{0/b}$  is a term used for aligning the electrostatic potentials of the bulk and the defective supercells: in practice, this term effectively places the VBM in the supercell with the neutral defect at  $\epsilon_v + \Delta v_{0/b}$ . We note that  $\Delta v_{0/b}$  appearing in Eq. (2) carries an opposite sign with respect to the conventional electrostatic potential,  $\Delta V_{0/b} = -\Delta v_{0/b}$ , which we preferentially adopt in the following and distinguish by an upper case letter. Third, the interaction between localized charges in a neutralizing background is long-ranged and converges slowly with respect to the supercell size  $L$ . The energy related to this interaction can be sizable even in fairly large supercells. Thus, we have explicitly added to Eq. (2) the term  $E_{\text{corr}}^q$  which accounts for the finite-size supercell correction.

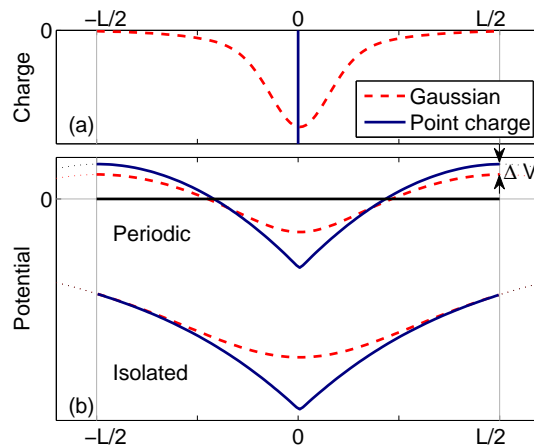


FIG. 1. (Color online) (a) Charge distributions for a point charge (solid line) and for a Gaussian charge (dashed line), and (b) the corresponding electrostatic potentials subject to open and periodic boundary conditions.

### B. Potential alignment

The alignment term  $\Delta V_{0/b}$  can be obtained through the comparison of the electrostatic potential in the bulk-like region far from the neutral defect and in the bulk calculation:  $\Delta V_{0/b} = V_{0|\text{far}} - V_b$ . Indeed, although the electronic structure far from the neutral defect would be very similar to that of the bulk, and consequently also the charge density, the average electrostatic potential in this region might differ by a constant. This is due to the convention of setting the average Hartree potential over the supercell to zero in DFT calculations and possibly to variations in the average value of the local pseudopotential which depends on the number and kind of atoms in the supercell. Hence, a localized change in the potential in the near-region needs to be compensated by a constant shift of the potential across the simulation cell

$$\frac{1}{\Omega} \int_{\text{near}} \delta V(\mathbf{r}) d\mathbf{r} = -\Delta V, \quad (3)$$

where  $\Omega$  is volume of the supercell and  $\delta V(\mathbf{r})$  is some potential change confined to the near-region. This relation is illustrated in Fig. 1 for a point charge and a Gaussian charge distribution. For a fixed near-region which encompasses the localized changes in the potential, the shift  $\Delta V$  then scales at large  $L$  like the inverse volume of the simulation cell, i.e.  $\Omega^{-1}$  or  $L^{-3}$ .

Traditionally,  $\Delta V_{0/b}$  is obtained from the comparison between the neutral and bulk potentials and then also used in the charged defect calculations. The potential in a charged defect calculation converges very slowly and can not be used directly to obtain the potential alignment. To illustrate this behavior, we plot in Fig. 2 the electrostatic potentials along the (001) direction for the carbon vacancy in diamond in the neutral, +2, and -4

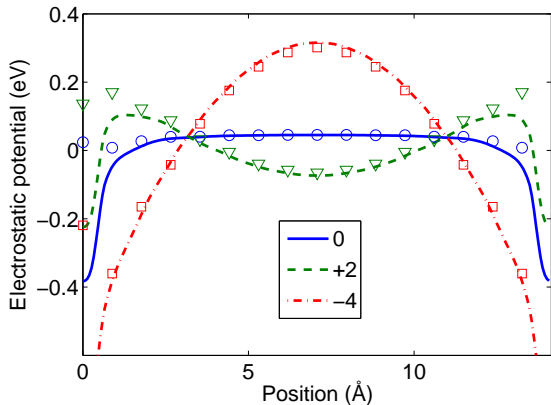


FIG. 2. (Color online) The electrostatic potential for the carbon vacancy in diamond for the neutral (solid, blue), +2 (dashed, green), and -4 (dash-dotted, red) charge states in the unrelaxed geometry. The electrostatic potential is represented through averages over transversal planes (lines) and over atomic core regions (symbols). The defect is located at the origin and its first periodic image is found at 14.1 Å.

charge states. For the neutral defect, the potential converges rapidly to a constant value, and  $\Delta V_{0/b}$  can generally be determined in an unambiguous way. At variance, in the charged defect calculations, the potential is far from having reached a converged value. In Fig. 2, two commonly used methods are used to illustrate the evolution of the potential, viz. through the average over transversal planes and through averages over the core regions around the nuclei. Far from the defect, these two ways of picturing the potential generally lead to very similar alignments. In the case of relaxed geometries, the core-average method might be preferable as the planar-average method would require an additional smoothing procedure. Alternatively,  $\Delta V_{0/b}$  could be obtained through a spatial partitioning scheme such as based on Voronoi cells or through the use of Wannier functions to estimate the potential averages, as discussed by Corsetti and Mostofi.<sup>13</sup>

### C. Charge corrections in a supercell of finite size

In electrostatics correction schemes for charged defects, three systems are generally considered: (1) a reference system which corresponds to the pristine host, (2) a system similar to the reference, but with the addition of a periodic array of defect charges represented through some localized charge distribution  $\rho_c$  and of a neutralizing background jellium  $n$ , and (3) a system similar to the reference, but with the addition of a single isolated defect charge distribution  $\rho_c$ . The desired correction corresponds to the difference between the electrostatic energies of the isolated and periodic defect charge distribu-

tions:

$$E_{\text{corr}} = E_{\text{iso}} - E_{\text{per}} \quad (4)$$

where  $E_{\text{iso}}$  is the selfenergy of the isolated charge distribution  $\rho_c$ , and  $E_{\text{per}}$  the electrostatic energy of the system subject to periodic boundary conditions

$$E_{\text{per}} = \frac{1}{2} \int_{\Omega} V_{\text{per}}(\mathbf{r}) \rho_c(\mathbf{r}) d\mathbf{r}, \quad (5)$$

including both the selfenergy as well as the interaction with the periodic images and with the background charge. The integral in Eq. (5) is carried out over the simulation cell  $\Omega$ . The factor  $\frac{1}{2}$  is required to remove the double-counting. The potential  $V_{\text{per}}$  is obtained by solving the Poisson equation for the defect charge distribution *and* the background charge density  $n$  subject to periodic boundary conditions. In other words, under the assumption that the selfinteraction of the charge distributions in the isolated and in the periodic cases are the same and cancel, the opposite of  $E_{\text{corr}}$  corresponds to the energy resulting from the interaction of the defect charge with its periodic images and with the background charge. This is precisely what needs to be removed from the calculated formation energy in order to recover the formation energy of the isolated defect. The correction energy can also be written as<sup>14</sup>

$$E_{\text{corr}} = \frac{1}{2} \int_{\Omega} V_{\text{corr}}(\mathbf{r}) \rho_c(\mathbf{r}) d\mathbf{r}, \quad (6)$$

where  $V_{\text{corr}} = V_{\text{iso}} - V_{\text{per}}$  only includes the potential from image defects and from the background, but not from the defect itself.

When the localized charge distribution  $\rho_c$  is known, as this is the case for atoms and molecules in vacuum, the correction  $E_{\text{corr}}$  can be obtained in a straightforward fashion.<sup>7,8</sup> However, in the case of defects in solids, there are several complicating aspects. Most importantly, the Coulomb interaction between localized charges is screened by the host. The screening has to be accounted for in the evaluation of the energy correction. Depending on the level of sophistication, various models for the screening could be employed. The simplest solution consisting in scaling the Coulomb interaction by the macroscopic dielectric constant  $\epsilon$  of the bulk is often sufficient.

The screening response also significantly changes the total charge density of the system. We may conceptually divide the total charge difference into three contributions,  $\rho_c + \rho_{\text{scr}}^{\text{loc}} + \rho_{\text{scr}}^{\text{deloc}}$ , although such a partition might practically be difficult to realize. The localized defect charge distribution  $\rho_c$  is screened by a localized charge distribution  $\rho_{\text{scr}}^{\text{loc}}$ , which corresponds to a total screening charge of about  $-(1 - 1/\epsilon)q$  accumulating on length scales on which local fields are important. Furthermore, since the charge neutrality within the simulation cell is preserved, the localized screening charge has to come from the supercell. This gives rise to a completely delocalized, quite uniform charge distribution,<sup>9,15</sup> which is here described

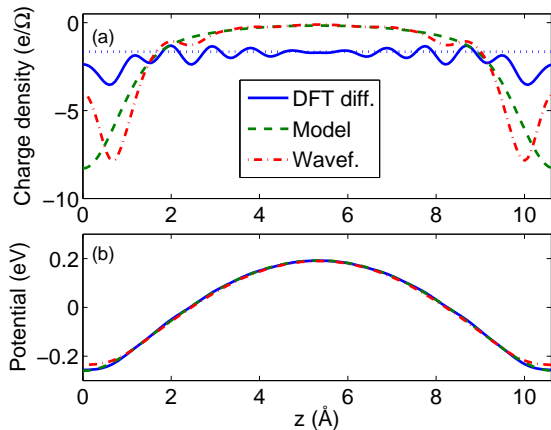


FIG. 3. (Color online) Comparison between (a) three charge distributions and (b) their relative electrostatic potentials for the  $V_C^{-2}$  defect in a 216-atom supercell: the charge difference between the charged and the neutral DFT calculations (solid, blue), a model charge density distribution (dashed, green), and the partial charge density associated to the defect wave function (dashed-dotted, red). The horizontal dotted line indicates the level of the delocalized screening charge density of  $(1 - 1/\epsilon)q/\Omega$ . Except for the result of the DFT calculation, the screening in the calculation of the potentials is accounted for by the consideration of the macroscopic dielectric constant  $\epsilon$ .

by  $\rho_{\text{scr}}^{\text{deloc}}$ . For an isolated defect, this charge is pushed to infinity (or to the surface in realistic situations) and can be ignored.

This detailed picture of the screening response is fully consistent with the expected long-range behavior of the potential, which scales as  $q/\epsilon r$ . This is illustrated in Fig. 3, where we display the charge difference between the charged and the neutral calculations, the partial charge density associated to the defect wave function  $\rho_d = |\psi_d|^2$ , and the charge density distribution of a Gaussian model charge. We also show in Fig. 3, the respective electrostatic potentials obtained by solving the Poisson equation. Note that the DFT potential is directly obtained from the charge density (i.e. with  $\epsilon = 1$ ) since the latter already accounts for all the screening effects, but in the other two cases the screening is included through the explicit use of the macroscopic dielectric constant  $\epsilon$ . The obtained potentials agree very well with each other despite the different appearance of the charge distributions.

A further difficulty encountered when treating charged defects in simulation cells of finite size results from the effect of the image potentials on the Kohn-Sham levels associated to the defect. Depending on the strength of interaction with the periodic images, these Kohn-Sham levels undergo significant shifts which may bring them to hybridize with the electronic bands of the host. This effect may lead to a partial delocalization of the defect charge throughout the simulation cell, which effectively alters the nature of the studied defect. The underestima-

tion of the band gap in conventional semilocal approximations to DFT further enhances this problem. We note that this effect is distinct with respect to the quantum-mechanical interaction between image defect wave functions, which generally scales exponentially with supercell size but which might affect the detailed convergence properties.<sup>5</sup> The screening, hybridization, and quantum-mechanical effects concurrently contribute to the nontrivial behavior of defect calculations in simulation cells of finite size.

### 1. Scheme of Makov and Payne

In the following, we introduce the correction schemes considered in this work and focus in particular on the estimation of  $E_{\text{corr}}$ . We first consider the case of a finite system in vacuum in which the extra charge can be well described by a point charge. In the ideal case of a periodically repeated point charge within a neutralizing jellium, the electrostatic energy due to periodic images is the well known Madelung energy

$$E_{\text{corr}} = \frac{q^2 \alpha}{2L} \quad (7)$$

where  $\alpha$  is the Madelung constant which depends on the Bravais lattice.<sup>7,16</sup> For example, for sc, bcc, and fcc lattices,  $\alpha$  assumes the value of 2.8373, 3.6392, and 4.5848, respectively.<sup>7</sup> The value of  $\alpha$  can be calculated for any Bravais lattice through the use of the Ewald method.<sup>14,17</sup>

Makov and Payne derived an extension of the Madelung-type lattice sum for a more extended but localized charge distribution.<sup>8</sup> Focusing on cubic simulation cells, they found that the leading correction terms for the energy read:<sup>8</sup>

$$E_{\text{corr}} = \frac{q^2 \alpha}{2L} - \frac{2\pi q Q}{3\Omega} \quad (8)$$

where

$$Q = \int r^2 \rho_c(\mathbf{r}) d\mathbf{r} \quad (9)$$

is the second radial moment of the localized charge distribution  $\rho_c$ . This form of electrostatics correction composed of two terms is usually referred to as Makov-Payne (MP) correction. The first term corresponds to the Madelung energy in Eq. (7). The second term scales like  $L^{-3}$  and results from the interaction of the localized charge distribution with the uniform compensating background.<sup>18</sup>

The extension of the correction to a defect charge distribution embedded in a host is then obtained by scaling the result in Eq. (10) by the macroscopic dielectric constant of the medium:<sup>8</sup>

$$E_{\text{corr}}^{\text{MP}} = E_{\text{MP1}} + E_{\text{MP2}} = \frac{q^2 \alpha}{2\epsilon L} - \frac{2\pi q Q}{3\epsilon \Omega}, \quad (10)$$

where we denoted the first and second term MP1 and MP2, respectively. However, we note that the MP2 term is rarely applied because of difficulties encountered in defining the localized charge density of the defect. When the charge distribution becomes very wide, the MP expansion in Eq. (10) is no longer expected to apply. In fact, in the limit of a completely delocalized defect charge, the interaction between image charges is completely cancelled because of the compensation due to the background charge.

## 2. Scheme of Freysoldt, Neugebauer, and Van de Walle

The approach taken by Freysoldt, Neugebauer, and Van de Walle<sup>10</sup> is more suitable for the case of charged defects within a dielectric medium. In this scheme, the DFT potential is explicitly used to obtain an improved model of the electrostatics. Following Ref. 15, the correction is expressed as

$$E_{\text{corr}}^{\text{FNV}} = E_{\text{lat}} - q\Delta V_{q/0}. \quad (11)$$

The lattice energy  $E_{\text{lat}}$  corresponds to the correction energy defined as in Eq. (4) for a model defect charge distribution. For instance, in case a point-charge model is adopted,  $E_{\text{lat}}$  would correspond to the MP1 term of the Makov-Payne expression. The potential alignment term  $q\Delta V_{q/0}$  is obtained by comparing the potential from the model charge to the potential difference (charged vs. neutral) in the DFT calculation:

$$\Delta V_{q/0} = (V_q^{\text{DFT}} - V_0^{\text{DFT}})|_{\text{far}} - V^{\text{model}}|_{\text{far}}. \quad (12)$$

Note that, in case one takes as reference the pristine host rather than the neutral defect calculation, the second term in Eq. (11) becomes  $V_{q/0} + V_{0/b} = V_{q/b}$ , and one achieves an equivalent description by which the potential alignment introduced in Eq. (2) is no longer necessary.<sup>10,15</sup>

Different charge models give different  $E_{\text{lat}}$ . Nevertheless, the FNV scheme yields corrected formation energies that are largely independent of the adopted charge model insofar the defect charge is well localized within the simulation cell. As is desirable, only the defect charge state  $q$  matters. This property stems from the fact that the interaction of the defect with its periodic images is dominated by the long-range Coulomb potential  $1/\epsilon r$ . This behavior is captured by any reasonable model charge distribution. As the local effects due to the defect are difficult to capture within a model description, differences between the DFT potential and the model potential occur but are short ranged. These differences are accounted for in the FNV scheme through the alignment-like energy term. The connection between the model defect charge distribution and the potential alignment is further discussed in Sec. IID.

The FNV scheme provides a rather general framework for calculating energy corrections for defects in supercells.

The scheme applies to supercells and defect charge distributions of any shape. Furthermore, the comparison between the model and the DFT potentials provides a mean to verify the consistency of the applied approach. In practical applications, it might be difficult to extract  $\Delta V_{q/0}$  from the comparison between the model and the DFT potentials, as the former generally shows a smooth behavior while the latter undergoes strong variations. The alignment of the model and the DFT potentials also becomes problematic when the defect charge distribution significantly delocalizes over the simulation cell.

## 3. Scheme of Lany and Zunger

The last scheme that we consider was suggested by Lany and Zunger (LZ).<sup>9</sup> They proposed to calculate the second radial moment in the MP2 term using the charge difference obtained directly from the total charge densities between the charged and neutral DFT calculations, cf. the charge difference in Fig. 3(a). Since the charge difference beyond the immediate vicinity of the defect was found to be dominated by a delocalized screening charge of density  $n_s = (1 - 1/\epsilon)q/\Omega$ , the second radial moment could easily be calculated. For a tetragonal cell, one finds:

$$Q = \int n_s r^2 d\mathbf{r} = \frac{1}{12}(L_x^2 + L_y^2 + L_z^2) \left(1 - \frac{1}{\epsilon}\right) q \quad (13)$$

For a cubic cell this then gives an MP2 energy correction of

$$E_{\text{MP2}} = -\frac{\pi q^2}{6\epsilon L} \left(1 - \frac{1}{\epsilon}\right) = -E_{\text{MP1}} c_{\text{sh}} \left(1 - \frac{1}{\epsilon}\right) \quad (14)$$

where  $c_{\text{sh}}$  depends only on the shape of the supercell. For a cubic cell,  $\pi/3\alpha \approx 0.369$ . This correction can be combined with the MP1 term,<sup>19</sup>

$$E_{\text{corr}}^{\text{LZ}} = \left[1 - c_{\text{sh}} \left(1 - \frac{1}{\epsilon}\right)\right] \frac{q^2 \alpha}{2\epsilon L}, \quad (15)$$

and effectively results in the MP1 energy term scaled by a factor of about 0.65.<sup>9</sup> More accurate evaluations require the calculation of  $c_{\text{sh}}$  for the adopted cell shape.<sup>19</sup>

However, the use of the DFT charge density difference leads to some conceptual difficulties. First, the adopted MP formulation of the energy correction relies on an expansion of increasing powers of  $1/L$ , for which defect charge distribution is required to be localized. This clearly does not apply to the adopted charge distribution. As a consequence the term in the MP expansion scaling like  $L^{-3}$  is found to scale like  $L^{-1}$ . Also, it is not clear how higher order terms in the expansion should be treated.<sup>19</sup> Second, the original MP formulation states that the second radial moment is to be calculated for the ‘‘immersed’’ defect charge without any screening effects,<sup>8</sup> which is instead included in the LZ formulation. Third,

the MP1 term should be exact for a point charge, whereas LZ scheme still gives an additional MP2 term.

We remark that Lany and Zunger apply their energy correction in combination with a potential alignment scheme for charged defects. In the latter scheme, the alignment is achieved through the determination of the average of the electrostatic potential evaluated at all atomic sites but those in the immediate neighborhood of the defect.<sup>19</sup> This alignment scheme is similar to that based on the average of the Kohn-Sham potential adopted by Taylor and Bruneval, which ensures that the Coulomb potential associated to the defect charge averages to zero over the simulation cell.<sup>5</sup> In this respect, the exclusion region considered by Lany and Zunger appears to account for the modified chemistry but unduly discards the associated local contribution of the electrostatic potential to the average. For these reasons and for compatibility with the application of the MP and FNV schemes, we here prefer to apply the LZ scheme in conjunction with the alignment procedure described in Sec. II B.

#### D. Connection between MP and FNV schemes

In this section, we first demonstrate the equivalence of the MP and FNV schemes for the case of finite systems in vacuum. We then show that the FNV scheme is more amenable to be extended to defect systems since it does not require the unscreened defect charge distribution.

Following Taylor and Bruneval,<sup>5</sup> we add the MP energy correction corresponding to a finite system in vacuum, given in Eq. (8), to the DFT functional and derive the selfconsistent potential. Within the Kohn-Sham formalism, we obtain the effective potential as a functional derivative of the energy with respect to the electron density  $n_e(\mathbf{r})$ :

$$v_{\text{KS+corr}}(\mathbf{r}) = \frac{\delta(E[n] - T[n] + E_{\text{corr}}^{\text{MP}})}{\delta n_e(\mathbf{r})} \quad (16)$$

The first two terms yield the customary Kohn-Sham potential. Since the MP correction energy is a functional of the total charge density, we introduce the charge distribution

$$\rho(\mathbf{r}) = \sum_i Z_i \delta(\mathbf{r} - \mathbf{R}_i) - n_e(\mathbf{r}), \quad (17)$$

where  $\mathbf{R}_i$  and  $Z_i$  are the position and the charge of the  $i$ th nucleus, respectively. Application of the chain rule gives the potential  $v_{\text{corr}}^{\text{MP}}$  acting on the electronic wave functions and the corresponding electrostatic potential  $V_{\text{corr}}^{\text{MP}}$ :

$$v_{\text{corr}}^{\text{MP}} = -\frac{\delta E_{\text{corr}}^{\text{MP}}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} = -V_{\text{corr}}^{\text{MP}}. \quad (18)$$

Denoting  $q = \int \rho(\mathbf{r}) d\mathbf{r}$  and  $Q = \int \rho(\mathbf{r}) r^2 d\mathbf{r}$ , we obtain

$$V_{\text{corr}}^{\text{MP}}(\mathbf{r}) = \frac{q\alpha}{L} - \frac{2\pi}{3\Omega} \left[ \frac{\delta q[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} Q + q \frac{\delta Q[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \right] \quad (19)$$

$$= \frac{q\alpha}{L} - \frac{2\pi Q}{3\Omega} - \frac{2\pi q}{3\Omega} r^2. \quad (20)$$

This expression coincides with the correction potential derived from a different starting point by Dabo *et al.*,<sup>14</sup> who gave the leading orders for a set of point charges.<sup>20</sup> We note that the first and third terms correspond to the correction potential for an ideal point charge  $q$ ,<sup>14</sup> the second term being an additional correction  $\Delta V$  accounting for the finite extent of the defect charge distribution:

$$\Delta V = -\frac{2\pi Q}{3\Omega}. \quad (21)$$

We now address the energy correction associated to the same charge distribution from the perspective of the FNV scheme. We first adopt a model charge distribution consisting of a point charge. This implies that  $E_{\text{lat}} = q^2\alpha/2L$ . As far as the alignment term is concerned, we need to evaluate the difference between the potential due to the point charge  $q$  and that due to the actual defect charge distribution at a large distance from the defect. For both systems, these potentials can be expressed through the correction potential given in Eq. (20). Since  $V_{\text{corr}} = V_{\text{iso}} - V_{\text{per}}$ , the shift in the periodic potential is just the opposite of the shift in the correction potential, leading to  $\Delta V_{q/0} = -\Delta V$ . Thus, the second term in the FNV energy correction,  $q\Delta V_{q/0}$ , precisely corresponds to the second term in MP energy correction given in Eq. (8). This demonstrates in a general way the correspondence between the FNV and MP schemes, which holds for any model charge distribution adopted in the FNV scheme. Application of the FNV and MP schemes to finite charge distributions thus leads to equivalent corrections.

In order to show more directly that the second term in MP energy correction solely results from an alignment effect, we obtain the result in Eq. (21) following an alternative path. According to Eq. (3) in Sec. II B, the potential shift can be obtained from the average of the potential difference between a model charge and a point charge:

$$\Delta V = \frac{1}{\Omega} \int d\mathbf{r} \left[ V(\mathbf{r}) - \frac{q}{r} \right] \quad (22)$$

$$= \frac{1}{\Omega} \int d\mathbf{r} \int d\mathbf{r}' \left[ \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} - \frac{\rho(\mathbf{r}')}{r} \right]. \quad (23)$$

The Coulomb kernel may be expanded in spherical harmonics:<sup>21</sup>

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{m=\ell} \frac{1}{2\ell+1} \frac{r_{<}}{r_{>}^{\ell+1}} Y_{\ell m}^*(\theta', \phi') Y_{\ell m}(\theta, \phi), \quad (24)$$

where  $r_>$  and  $r_<$  are the higher and lower value among  $r$  and  $r'$ , respectively. Because of the integral on  $d\mathbf{r}$  appearing in Eq. (22), only the  $\ell=0$  term survives, and yields without loss of generality:

$$\Delta V = \frac{1}{\Omega} \int d\mathbf{r} \int d\mathbf{r}' \left[ \frac{\rho(\mathbf{r}')}{r_>} - \frac{\rho(\mathbf{r}')}{r} \right]. \quad (25)$$

For  $r_> = r$ , the integrand vanishes and we thus need to consider in the following only the region where  $r_> = r' > r$ . We define  $\tilde{\rho}$  as the spherically averaged charge density,

$$\tilde{\rho}(r) = \frac{1}{4\pi} \int \rho(\mathbf{r}) \sin\theta d\theta d\phi, \quad (26)$$

and obtain

$$\Delta V = \frac{(4\pi)^2}{\Omega} \int_0^\infty dr r^2 \int_r^\infty dr' r' \tilde{\rho}(r') \quad (27)$$

$$- \frac{(4\pi)^2}{\Omega} \int_0^\infty dr r \int_r^\infty dr' r'^2 \tilde{\rho}(r'). \quad (28)$$

Straightforward integration by parts then yields

$$\Delta V = \frac{4\pi}{\Omega} \left( \frac{1}{3} - \frac{1}{2} \right) \int_0^\infty dr 4\pi r^4 \tilde{\rho} = -\frac{2\pi}{3\Omega} Q, \quad (29)$$

in agreement with Eq. (21). In the Appendix, this derivation is explicitly illustrated for commonly used model charge distributions.

It is of interest to extend the present correction schemes to the case of charged defects embedded in host materials. Makov and Payne generalized their scheme to this situation by considering the dielectric constant of the host.<sup>8</sup> However, the calculation of the second radial moment appearing in their energy correction requires the unscreened defect charge distribution which is not directly available in DFT calculations. At variance, the FNV scheme only uses a simple alignment procedure which does not require the unscreened defect charge distribution,<sup>10</sup> yet achieves equivalent results to the MP correction scheme, as demonstrated above.

Recently, the validity of applying a potential alignment in the case of charged defects has been questioned due to possible double counting effects.<sup>5</sup> To address this issue, we follow the derivation of Dabo *et al.*<sup>22</sup> We evaluate the energy correction through the expression in Eq. (6) which corresponds to the interaction of the defect charge distribution  $\rho_c$  with the correction potential  $V_{\text{corr}}$  generated by  $\rho_c$  itself. The double counting is explicitly avoided through the inclusion of the factor  $\frac{1}{2}$ . We express the defect charge distribution as  $\rho_c = \rho_0 + \Delta\rho$ , where  $\rho_0$  corresponds to a point charge and  $\Delta\rho$  is a neutral charge distribution. The correction potential can correspondingly be expressed as  $V_0 + \Delta V$ , where  $V_0$  is the potential resulting from the image point charges and the neutralizing background, and  $\Delta V$  corresponds to the constant potential shift in Eq. (21) accounting for the difference between  $\rho_c$  and the point-charge density  $\rho_0$ . The energy

correction reads:

$$E_{\text{corr}} = \frac{1}{2} \int (\rho_0 + \Delta\rho)(V_0 + \Delta V) d\mathbf{r} \quad (30)$$

$$= \frac{1}{2} \int \rho_0(V_0 + \Delta V) d\mathbf{r} + \int \Delta\rho(V_0 + \Delta V) d\mathbf{r} \quad (31)$$

$$= \frac{1}{2} \int \rho_0 V_0 d\mathbf{r} + \frac{1}{2} q \Delta V + \frac{1}{2} \int \Delta\rho V_0 d\mathbf{r}, \quad (32)$$

where we used that  $\Delta V$  is a constant shift,  $q$  the integral of  $\rho_0$ , and  $\Delta\rho$  a neutral charge distribution. The first term corresponds to the Madelung correction, the second term to a potential alignment, and the third term to the interaction of the defect charge with the parabolic potential appearing in Eq. (20). Hence,

$$E_{\text{corr}} = \frac{\alpha q^2}{2L} + \frac{1}{2} q \Delta V - \frac{\pi q}{3\Omega} Q \quad (33)$$

$$= \frac{\alpha q^2}{2L} + q \Delta V \quad (34)$$

where we could combine the second and third terms through the relation in Eq. (21). Thus, despite the absence of the  $\frac{1}{2}$  factor in the potential alignment term, this expression does not suffer from any double counting effects.

### III. RESULTS

In order to benchmark and compare these schemes, we perform calculations with varying supercell sizes for a fairly large set of defects in three different host materials. In particular, we consider zinc-blende GaAs, diamond, and cubic MgO. These materials correspond to a small band-gap semiconductor, a large band-gap covalent system, and a system with ionic bonding. Within these hosts, a range of different types of defects are considered. In diamond, we consider the carbon vacancy  $V_C$ , the hydrogen interstitial  $H_I$ , and the nitrogen substitutional defect  $N_C$ . In GaAs, we consider the gallium vacancy  $V_{\text{Ga}}$  and the arsenic antisite  $As_{\text{Ga}}$ . In MgO, we consider the oxygen vacancy  $V_O$  and the substitutional nitrogen at the oxygen site  $N_O$ . Some of these defects are considered in various charge states.

#### A. Computational details

The calculations are carried out within a plane-wave density-functional-theory scheme. Here, we use the local density approximation (LDA) with the projector augmented wave formalism, as implemented in the VASP package.<sup>23–25</sup> In the GaAs calculations, we adopt a PAW setup for gallium in which the  $3d$  electrons in the core are frozen and use an energy cutoff of 400 eV. In the diamond and MgO calculations, the energy cutoff is set at 500 eV. In Table I, we give bulk parameters of the



TABLE I. Calculated bulk parameters for diamond, GaAs, and MgO: the lattice constant  $a_{\text{lat}}$  (Å), the direct band gap  $E_g$  at the  $\Gamma$  point (eV), and the high-frequency ( $\epsilon_\infty$ ) and static ( $\epsilon_0$ ) dielectric constants. Experimental data are from Refs. 28–33.

	diamond	GaAs	MgO
$a_{\text{lat}}$	3.536	5.627	4.155
Expt.	3.567	5.648	4.207
$E_g$	5.62	0.50	5.06
Expt.	7.3	1.5	7.7
$\epsilon_\infty/\epsilon_0$	5.76/5.76	13.7/15.7	3.10/7.25
Expt.	5.7/5.7	11.1/13.1	3.0/9.6

host materials: the lattice constant, the direct band gap at the  $\Gamma$  point, and the high-frequency (ion-clamped) and the static (relaxed-ion) dielectric constants, together with the corresponding experimental values. We calculate these parameters using primitive unit cells. The dielectric constants are calculated via the density functional perturbation theory.<sup>26,27</sup>

For the defect calculations, three cubic supercells of increasing size are considered. The 64-atom supercells are sampled with a  $4 \times 4 \times 4$   $\mathbf{k}$ -point mesh, the 216-atom supercells with a  $2 \times 2 \times 2$  mesh, and the 512-atom supercells with a  $2 \times 2 \times 2$  mesh. The  $\mathbf{k}$ -point meshes are offset from the  $\Gamma$  point. The occupation of the defect bands has to be consistent among the different supercells. This is achieved by fully occupying the  $N$  lowest eigenstates at each  $\mathbf{k}$ -point, as opposed to occupying the levels up to Fermi level across the Brillouin zone. This also improves the total energy convergence with respect to the density of the  $\mathbf{k}$ -point mesh. In the present calculations, we ensure  $\mathbf{k}$ -point sampling convergence of the formation energies, as the importance of this issue has been highlighted several times in the study of defect corrections.<sup>2,34,35</sup>

The formation energies are plotted with respect  $L^{-1}$ , where  $L$  is the side of the supercell. In all figures, lines of the form  $aL^{-1} + bL^{-3} + c$  are fitted to the calculated values to obtain extrapolated values corresponding to the limit  $L \rightarrow \infty$ .<sup>2,3</sup> We note that the MP and LZ correction schemes only have  $L^{-1}$  and  $L^{-3}$  terms. This implies that the MP-corrected and the LZ-corrected formation energies extrapolate by construction to the same values as the uncorrected formation energies, provided the potential alignment is carried out in the same way in the two schemes. When formation energies with and without potential alignment are compared, their difference is expected to scale predominantly like  $L^{-3}$  leading to practically identical extrapolated values. Residual differences could be observed because of uncertainties associated to the determination of the potential shift. In view of the discussion in Sec. IID, the FNV scheme should in principle also extrapolate to the same values apart from small deviations resulting from the potential alignment. Thus, the adopted form of the fitting polynomial should be appropriate and all correction schemes are expected to give close if not identical extrapolated values.

To obtain reliable extrapolated values for the formation energies, it is important to minimize all possible computational errors since they affect the quality of the fit. In addition, there are other effects than pure electrostatics that affect the scaling. For instance, the interaction between periodic images could include quantum-mechanical effects due to the overlap between their defect wave functions, but are expected to decay exponentially with supercell size.<sup>5,36</sup> Elastic effects could also be significant in certain cases, but are of minor concern in this work as we mostly examine unrelaxed defects. Therefore, it is unlikely that the various correction schemes could produce exactly the same corrected formation energy, independent of the supercell size. Furthermore, as these effects could have a scaling behavior differing from  $L^{-1}$  or  $L^{-3}$ , they might affect the extrapolated result.

## B. Localized defect charge distributions

In the comparison of correction schemes, we give the formation energies related to the following cases: (1) the uncorrected result, (2) uncorrected, but with potential alignment (PA)  $V_{0/b}$ , (3) Makov-Payne (MP1) with potential alignment (in selected cases also without it), (4) LZ, and (5) FNV. We start by considering all these correction schemes in their simplest forms. This is the preferred method of application for the schemes, but also makes the comparison unambiguous. In particular, since the Makov-Payne method is sensitive to the width of the charge, we here only consider the point-charge correction MP1. For FNV, we use a model charge distribution consisting of a Gaussian with a fixed width of  $\beta = 2$  bohr [cf. second term in Eq. (38) of the Appendix], but the correction for a localized defect is expected to be insensitive to the adopted model charge.

### 1. Point charge defects

We start with defects with the simplest type of charge distribution, namely the point charge. We consider two different defects in diamond: the +1 charge state of interstitial hydrogen and the +1 charge state of substitutional nitrogen. In both cases, the defect corresponds to adding a positive unitary point charge to the bulk.

The formation energies are shown in Fig. 4. The uncorrected and the potential-alignment corrected curves extrapolate to the same value in the limit of infinitely large supercell. The potential alignment is found to have only  $L^{-3}$  dependency as expected (cf. Sec. IIB). Similarly, the formation energies corrected through the FNV scheme extrapolate to the same limit, thus indicating that the adopted form for the fitting polynomial is appropriate and that the extrapolated value can reliably be taken as a benchmark.

All the correction schemes clearly improve upon the uncorrected results. Based on the discussion in Sec. II,

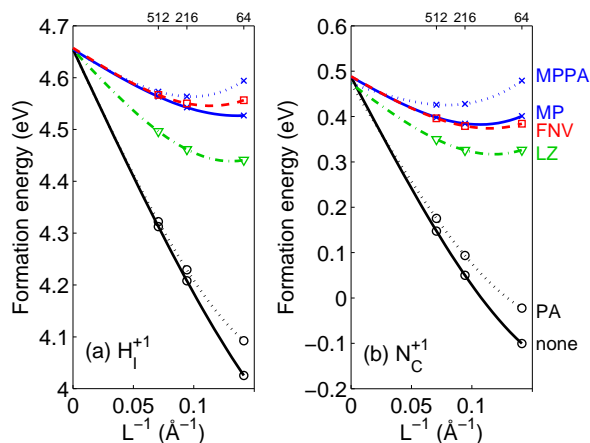


FIG. 4. (Color online) Formation energies of the unrelaxed  $H_I^{+1}$  and  $N_C^{+1}$  defects in diamond vs inverse supercell size: uncorrected (circles, solid, black), potential alignment (circles, dotted, black), Makov-Payne without PA (crosses, solid, blue), MP with PA (crosses, dotted, blue), FNV (squares, dashed, red), and LZ (triangles, dash-dotted, green).

the Madelung-type correction is expected to work well for point-like charges, and this is indeed found to hold for both defects. The slight undercorrection observed for the MP scheme at finite  $L$  might either result from effects neglected in the correction scheme or from the limitation of the extrapolation form, and thus gives a measure of the accuracy of our benchmarking. The FNV scheme is found to yield very similar results. Furthermore, the model potential agrees well with its counterpart derived from the DFT calculations (not shown). The LZ scheme produces always smaller corrections than the MP one, and thus systematically underestimates the correction for point charges.

While MP and FNV energies are close, they do not coincide because of the potential alignment term  $\Delta V_{q/0}$ . The MP scheme should correctly account for a point-charge distribution, but microscopic screening effects due to local fields could produce changes in the local potential around the defect, thereby affecting the alignment term in the FNV scheme.

### 2. Defects with extended but localized charge distribution

Next, we consider more extended defect charge distributions, which nevertheless are still well localized within the supercell. In Fig. 5, we show the formation energies vs. inverse supercell size for  $V_C^{-2}$  in diamond and  $N_O^{-1}$  in MgO.

Again, the curves extrapolate to the same values and a clear improvement is observed upon the uncorrected results. In these cases, the application of the MP scheme results in a small overcorrection, which has often been encountered in the literature.<sup>3,10,37</sup> The situation could

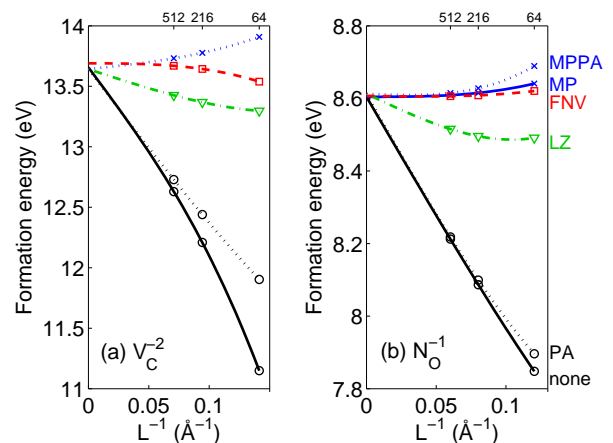


FIG. 5. (Color online) Formation energies of the unrelaxed  $V_C^{-2}$  defect in diamond and of the unrelaxed  $N_O^{-1}$  defect in MgO vs. inverse supercell size. Same notation as in Fig. 4.

be improved by taking into account the second MP term with a proper charge distribution. However, this approach can practically not be realized because of difficulties in determining the unscreened defect charge. The FNV scheme overcomes this problem by introducing a proper alignment term. This results in markedly improved corrections. The LZ scheme also performs reasonably well in these cases giving corrections which underestimate the extrapolated values only slightly.

### 3. Relaxation of defect geometry

All the defects considered so far have been treated without allowing for structural relaxation. In this section, we address relaxation effects. In calculations of unrelaxed defects, it is appropriate to use the high-frequency (ion-clamped) dielectric constant for describing the screening. When the defect geometry is allowed to relax, the relevant dielectric constant is the low-frequency (relaxed-ion) one. Among the materials considered in this work, diamond has the same high and low frequency dielectric constants, while GaAs shows only slightly differing dielectric constants. At variance, the difference between high and low frequency dielectric constants is large for the highly ionic compound MgO. Thus, we focus in the following on MgO for which the choice of the right dielectric constant is more critical.

In order to demonstrate the effects of these differences, we show in Fig. 6 the formation energies of the +1 charge state of the unrelaxed and relaxed O vacancy in MgO. The generic features of the correction schemes are similar for the unrelaxed and relaxed defect. However, the use of the incorrect dielectric constant in the MP1 correction scheme would lead to a dramatic worsening of the corrected result as illustrated in Fig. 6. When the low-frequency dielectric constant is used for the unrelaxed

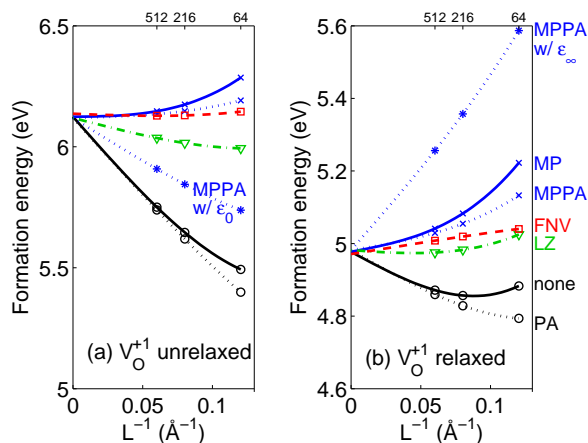


FIG. 6. (Color online) Formation energies vs. inverse supercell size for the (a) unrelaxed and (b) relaxed  $V_{\text{O}}^{+1}$  defect in MgO. Same notation as in Fig. 4. In (a), the MP1 correction is also applied with the inappropriate dielectric constant  $\epsilon_0$  (stars, dotted, blue). Similarly, the MP1 correction with  $\epsilon_\infty$  is shown in (b) (stars, dotted, blue).

defect, we observe a significant underestimation of the corrected result [Fig. 6(a)]. At variance, the use of  $\epsilon_\infty$  for the relaxed defect leads to a large overestimation [Fig. 6(b)].

Note, that since the atomic positions are different in the relaxed and bulk calculations, the resulting DFT potential difference can vary rapidly. This is especially true for an ionic host such as MgO. Consequently, determining the potential alignment in the FNV scheme involves some difficulty. To overcome this problem, we convolute the DFT potential with a suitable Gaussian function.

For reference, we list in Table II calculated formation energies for a set of relaxed defects. The formation energies of neutral defects correspond to the largest supercell considered, while the formation energies of charged defects correspond to extrapolated values of uncorrected results without potential alignment. In the latter case, the electron chemical potential is set at the valence band edge. The chemical potentials adopted here are as follows:  $\mu_{\text{C}}$  from bulk diamond,  $\mu_{\text{N}}$  from the  $\text{N}_2$  molecule,  $\mu_{\text{O}}$  from the  $\text{O}_2$  molecule,  $\mu_{\text{Ga}}$  from bulk Ga, and  $\mu_{\text{As}}$  from  $\mu_{\text{GaAs}} - \mu_{\text{Ga}}$ .

#### 4. Quantification of errors

In order to quantify the performance of the correction schemes, we define three measures of the error:

$$\delta E_1^{\text{err}} = E^f[64\text{-atom sc}] - E^f[\text{extrap.}], \quad (35a)$$

$$\delta E_2^{\text{err}} = E^f[512\text{-atom sc}] - E^f[\text{extrap.}], \quad (35b)$$

$$\delta E_3^{\text{err}} = E^f[216\text{-atom sc}] - E^f[512\text{-atom sc}], \quad (35c)$$

TABLE II. Calculated formation energies for a set of relaxed defects. For charged defects, the electron chemical potential is set at the valence band edge.

defect	$E^f$	defect	$E^f$
$\text{N}_{\text{C}}^0$	3.98	$\text{V}_{\text{O}}^0$	7.59
$\text{N}_{\text{C}}^{+1}$	0.44	$\text{V}_{\text{O}}^{+1}$	4.98
$\text{V}_{\text{C}}^{+2}$	5.92	$\text{V}_{\text{O}}^{+2}$	3.24
$\text{V}_{\text{C}}^{+1}$	6.01	$\text{V}_{\text{Ga}}^{-3}$	3.30
$\text{V}_{\text{C}}^0$	7.12	$\text{As}_{\text{Ga}}^0$	1.43
$\text{V}_{\text{C}}^{-2}$	12.68	$\text{As}_{\text{Ga}}^{+2}$	1.04
$\text{V}_{\text{C}}^{-4}$	22.30		

where the used formation energies correspond to the various supercell sizes used and to the extrapolated limit through a self-explanatory notation. In particular, we consider not only differences between corrected and extrapolated formation energies, but also differences between corrected formation energies for two different supercell sizes. These error estimators are evaluated for a set of unrelaxed and relaxed defects in Tables III and IV, respectively. We only consider defects in which the defect charge is found to be well localized within the supercell. This condition generally ensures that all correction schemes give close extrapolated values.

For each estimator, we calculate mean absolute errors (MAEs) as given in the tables. For the considered set of unrelaxed defects, the FNV scheme always provides the smallest MAE, with an estimated accuracy of 0.09, 0.04, and 0.02 eV for  $\delta E_1^{\text{err}}$ ,  $\delta E_2^{\text{err}}$ , and  $\delta E_3^{\text{err}}$ , respectively. However, in the particular case of point-charge defects like  $\text{H}_{\text{I}}^{+1}$  and  $\text{N}_{\text{C}}^{+1}$ , the MP scheme performs even better than FNV. Also, for some specific defects, the LZ scheme offers a competitive description compared to the FNV one. For the relaxed defects, the FNV scheme again generally gives the smallest MAEs. The LZ scheme appears competitive when considering the  $\delta E_3^{\text{err}}$  estimator.

### C. Delocalized defect charge distributions

#### 1. Failures of correction schemes: Role of delocalized charge

In our study, we encounter several cases in which the correction schemes apparently fail. This occurs when the various correction schemes no longer give close extrapolated values or when the corrected formation energies do not fall close to their extrapolated values.

One of such cases corresponds to the As antisite in GaAs. The formation energies of the +2 charge state are given in Fig. 7(a) for the various correction schemes. The uncorrected results, with or without the potential alignment  $\Delta V_{0/b}$ , produce slightly different extrapolated values. Since the potential alignment  $\Delta V_{0/b}$  is obtained as a difference between the neutral defect and the bulk potentials, the problem should be ascribed to the neutral charge state of the defect. Inspection of the Kohn-Sham

TABLE III. Error estimators for unrelaxed defects. The smallest errors are indicated in bold.

defect	$E^f[64\text{-atom sc}] - E^f[\text{extrap.}]$			$E^f[512\text{-atom sc}] - E^f[\text{extrap.}]$			$E^f[216\text{-atom sc}] - E^f[512\text{-atom sc}]$		
	MP	LZ	FNV	MP	LZ	FNV	MP	LZ	FNV
$\text{H}_\text{I}^{+1}$	<b>-0.06</b>	-0.21	-0.10	<b>-0.08</b>	-0.16	-0.09	<b>-0.01</b>	-0.03	-0.02
$\text{N}_\text{C}^{+1}$	<b>-0.01</b>	-0.16	-0.10	<b>-0.06</b>	-0.14	-0.09	<b>0.00</b>	-0.02	-0.02
$\text{V}_\text{C}^{-2}$	0.25	-0.36	<b>-0.12</b>	0.08	-0.23	<b>0.01</b>	0.05	-0.06	<b>-0.03</b>
$\text{V}_\text{C}^{-3}$	0.37	<b>-0.04</b>	0.20	0.15	-0.05	<b>0.04</b>	0.06	<b>-0.01</b>	0.04
$\text{V}_\text{O}^{+1}$	0.07	-0.13	<b>0.02</b>	0.01	-0.09	<b>0.00</b>	0.01	-0.02	<b>0.00</b>
$\text{N}_\text{O}^{-1}$	0.08	-0.11	<b>0.02</b>	0.01	-0.09	<b>0.00</b>	0.01	-0.02	<b>0.00</b>
MAE	0.14	0.17	<b>0.09</b>	0.07	0.13	<b>0.04</b>	<b>0.02</b>	0.03	<b>0.02</b>

TABLE IV. Error estimators for relaxed defects. The smallest errors are indicated in bold.

defect	$E^f[64\text{-atom sc}] - E^f[\text{extrap.}]$			$E^f[512\text{-atom sc}] - E^f[\text{extrap.}]$			$E^f[216\text{-atom sc}] - E^f[512\text{-atom sc}]$		
	MP	LZ	FNV	MP	LZ	FNV	MP	LZ	FNV
$\text{N}_\text{C}^{+1}$	<b>-0.04</b>	-0.20	-0.17	<b>-0.11</b>	-0.19	-0.14	<b>0.00</b>	-0.02	-0.02
$\text{V}_\text{C}^{-2}$	0.52	-0.09	<b>-0.03</b>	0.12	-0.18	<b>0.04</b>	0.09	<b>-0.01</b>	0.02
$\text{V}_\text{O}^{+1}$	0.16	<b>0.05</b>	0.06	0.05	<b>-0.00</b>	0.03	0.03	<b>0.01</b>	<b>0.01</b>
$\text{As}_\text{Ga}^{+2}$	<b>-0.02</b>	-0.18	-0.04	<b>-0.05</b>	-0.13	-0.06	<b>-0.00</b>	-0.03	-0.01
MAE	0.18	0.13	<b>0.07</b>	0.09	0.13	<b>0.07</b>	0.03	<b>0.02</b>	<b>0.02</b>

spectrum of the neutral defect indicates that the occupied defect band is resonant with the conduction band leading to hybridization between defect states and band-edge states of the host. This effect depends on the defect band width which varies with supercell size. The problem also manifests in the potential difference between the neutral and bulk calculation (not shown), which does not converge rapidly to a constant value when moving away from the defect. As shown in Fig. 8(a), this results from a spurious charge delocalization over the full simulation cell, which slightly alters the bulk charge density at large distance from the defect. We note that when the geometry of the defect is relaxed or when the band gap is opened, e.g. through the use of hybrid functionals, the Kohn-Sham levels move out of resonance and the problem disappears.<sup>38</sup> Also, since the problem is specific to the neutral defect, it can completely be circumvented in the calculation of charged defect levels by using the FNV scheme and taking the bulk as reference for the potential alignment  $\Delta V_{q/b}$ .

Another even more problematic case is the  $\text{V}_\text{C}^{+2}$  defect in diamond, which has often been discussed in the literature.<sup>10,37</sup> The formation energies are shown in Fig. 7(b). All the correction schemes are found to result in significant overcorrections. Furthermore, the FNV and the MP schemes do not extrapolate to the same limit. Among the considered correction schemes, the LZ one appears to give the best corrections. The situation is similar for the  $-4$  charge state of the same defect [Fig. 7(c)].

In the case of charged defects in which the correction schemes apply successfully, the delocalized charge density across the supercell corresponds to the  $(1 - 1/\epsilon)q/\Omega$  level (cf. Sec. II C). For reference, we illustrate this behavior in Figs. 3 and 8(b) for two of such well-behaved defects,

the  $\text{V}_\text{C}^{-2}$  and  $\text{N}_\text{C}^{+1}$  defects in diamond, respectively. We verified that this property holds for all well-behaving defects in this work. For the problematic  $\text{V}_\text{C}^{+2}$  and  $\text{V}_\text{C}^{-4}$  defects, the charge density difference at a large distance from the defect markedly differs from the expected value for the delocalized screening charge. Denoting this level with  $(1 - \eta/\epsilon)q/\Omega$ , we infer a fairly significant scaling factor  $\eta$  of about 0.8 for both  $\text{V}_\text{C}^{+2}$  and  $\text{V}_\text{C}^{-4}$ . The degree of charge delocalization thus appears to be a robust indicator for identifying defects for which the energy correction schemes are less effective.

The origin of this anomalous behavior leading to  $\eta \neq 1$  could in principle either result from the limitation of classical electrostatics in describing the screening around the defect or from defect charge leakage which would effectively reduce the charge localized at the defect site. However, the first explanation contrasts with the good correspondence between the classical prediction and the DFT charge density level observed for well-behaved defects. Furthermore, it is expected that classical screening behavior is recovered within at most a few bond lengths from the defect.<sup>39</sup> We thus focus in the following section on the extent of the defect charge delocalization.

## 2. Improving the model charge in the FNV scheme

Within the FNV scheme, the model parameters can be set to yield a model charge that extends beyond the supercell size. It was suggested in Ref. 10 that adding an exponential tail to the model charge could yield improved corrected energies in the case of the  $\text{V}_\text{C}^{+2}$  defect. The tail was obtained by fitting to the defect wave function. When the defect charge delocalizes, several different model charges may yield the same amount of charge at

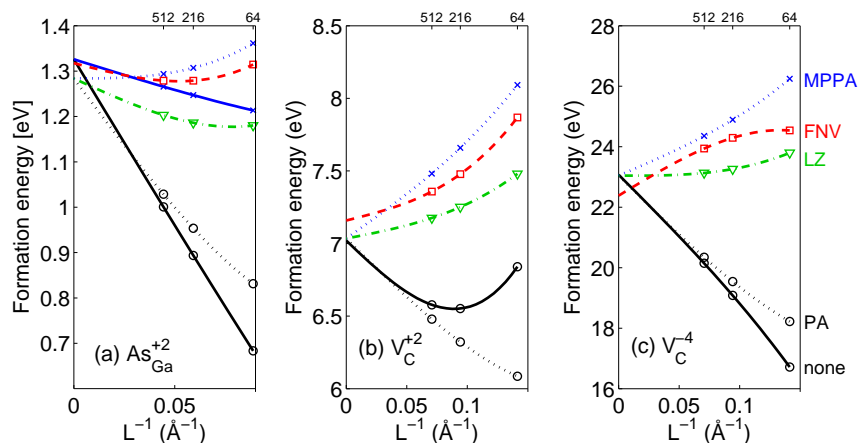


FIG. 7. (Color online) Formation energies vs. inverse supercell size for (a) the  $\text{As}_{\text{Ga}}^{+2}$  defect in GaAs, and (b) the  $\text{V}_{\text{C}}^{+2}$  and (c)  $\text{V}_{\text{C}}^{-4}$  defects in diamond. The considered defects are unrelaxed. Same notation as in Fig. 4.

large distance from the defect and consequently provide an equally good description of the long-range potentials (disregarding constant shifts). However, the FNV energy corrections for such delocalized defect charge distributions are no longer independent of the adopted model. It is therefore of interest to understand how the model charge relates to the energy corrections and to the behavior of the potential at large distances from the defect.

In this section, we consider various model charge distributions. All considered distributions can be expressed through a linear combination of a Gaussian and exponential functional as given in Eq. (38) of the Appendix.<sup>10</sup> The first model charge that we consider consists of a Gaussian charge density [ $x = 0$  in Eq. (38)] with a fixed width ( $\beta = 2$  bohr), as used hitherto in this work. The second model charge consists of a linear combination of a Gaussian and an exponential function with parameters fitted to the defect wave function, as proposed in Ref. 15. In the third model, the charge is partitioned between a regular Gaussian ( $\beta = 2$  bohr) and a completely delocalized charge density:

$$\rho_c(r) = q(1-x)N_\beta e^{-r^2/\beta^2} + qx/\Omega \quad (36)$$

which can be obtained from Eq. (38) by setting the exponential decay length  $\gamma$  to infinity. In this model charge, only a fraction  $\eta = 1 - x$  of the defect charge is localized at the defect site. In the periodically repeated system, the energy and the potential associated to this third model charge are just the same as for a Gaussian with a charge scaled by  $1 - x$ , since the homogeneously distributed charge only has a  $\mathbf{G} = 0$  Fourier component, which is set to zero by convention. Similarly, in the isolated system, only the Gaussian charge gives an energy contribution. The energy correction thus corresponds to that of a Gaussian charge distribution with an integrated charge  $\eta q$ .

In Figs. 9 and 10, we show corrected formation en-

ergies obtained using the various model charges in the FNV scheme for the  $\text{V}_{\text{C}}^{+2}$  and  $\text{V}_{\text{C}}^{-4}$  defects in diamond. The figures also contain a comparison between the corresponding model potentials and the DFT difference potential. In the third model charge introduced above, we used  $\eta = 0.8$  for both charge states as found from the study of the delocalized charge densities in Fig. 8. This value of  $\eta$  appears to have the same value for all supercells studied, albeit its accurate determination is difficult in the case of the 64-atom supercell.

In the FNV scheme, a good model charge is expected to yield accurate formation energies at finite supercell size, but also to ensure a good correspondence between the model potential and the DFT difference potential. As seen in Figs. 9 and 10, the Gaussian model charge leads to sizable errors in the corrected formation energies. Furthermore, this kind of model charge fails in reproducing the DFT difference potential. To improve upon the Gaussian charge model, it is necessary to delocalize more defect charge in the region far from defect. Fitting a model charge to the defect wave function significantly improves the model potential as shown in Figs. 9(c) and 10(c), but the formation energy is barely affected. Finally, when a fraction of the defect charge is taken to be completely delocalized, the agreement between the model potential and the DFT difference potential further improves [cf. Figs. 9(d) and 10(d)] and the formation energies converge rapidly to their extrapolated value. However, the physical motivation of using such delocalized model charge is questionable especially in the limit of infinitely large supercells. This is discussed in more detail below in Sec. III C 4. Nevertheless, for delocalized defect charges, these examples illustrate that it is difficult to infer the quality of the energy correction from the comparison between the DFT difference potential and the model potential when a calculation at only one supercell size is available.

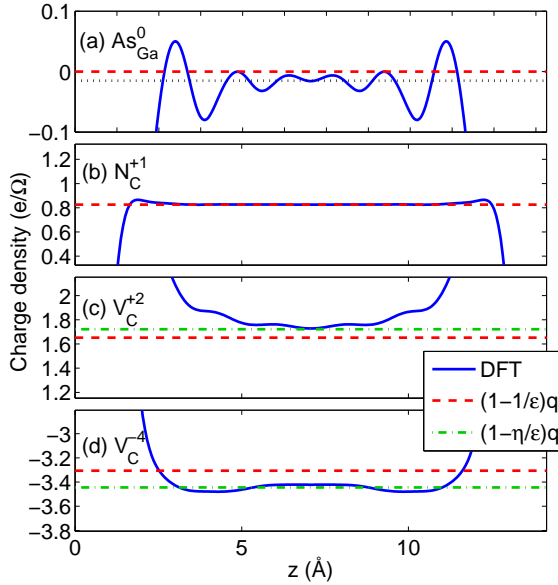


FIG. 8. (Color online) Planar-averaged charge density difference across the supercell for (a) the  $\text{As}_{\text{Ga}}^0$  defect in GaAs, and (b) the  $\text{N}_{\text{C}}^{+1}$ , (c)  $\text{V}_{\text{C}}^{+2}$ , and (d)  $\text{V}_{\text{C}}^{-4}$  defects in diamond. In (a), the difference is taken between the charge densities of the neutral defect and of the bulk; in (b–d), between those of the charged and neutral defect. The level corresponding to the homogeneously distributed screening charge  $(1 - 1/\epsilon)q/\Omega$  is indicated (dashed, red). The dash-dotted (green) line is obtained through a fit and corresponds to  $(1 - \eta/\epsilon)q/\Omega$  with  $\eta = 0.8$  for both  $\text{V}_{\text{C}}^{+2}$  and  $\text{V}_{\text{C}}^{-4}$ . The charge densities are broadened through a Gaussian function with a width of 1.5 bohr.

### 3. Hybridization between defect and band-edge states

The results in the previous section indicate that some degree of defect charge delocalization may be present in the calculations. Such delocalization could originate from the hybridization of the supposedly localized defect states with delocalized band-edge states. If the delocalization of the model charge is indeed caused by hybridization with the band edges, one would expect this effect to diminish upon the opening of the band gap of the host material.<sup>40,41</sup>

To examine this supposition, we perform hybrid-functional calculations for the vacancy in diamond. This kind of calculations are known to yield larger band gaps than the local density approximation. For consistency, we perform a comparison between the semilocal PBE functional<sup>42</sup> and the hybrid PBE0 functional<sup>43</sup>, which only differ by the way the exchange energy is described. These calculations are performed with plane-wave basis sets and normconserving pseudopotentials, as implemented in the QUANTUM-ESPRESSO package.<sup>44</sup> We used an energy cutoff of 50 Ry. In the case of hybrid functionals, the evaluation of the nonlocal Fock exchange

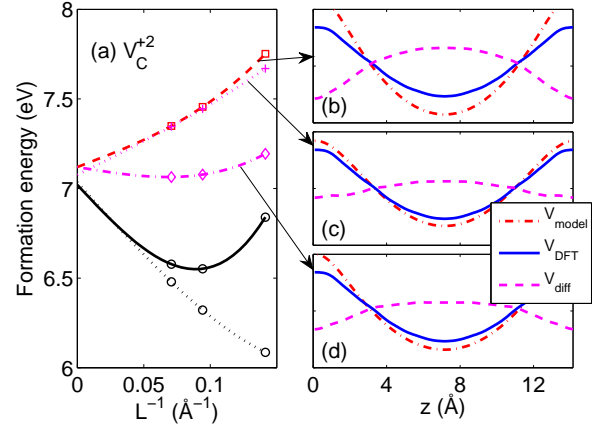


FIG. 9. (Color online) (a) Formation energies vs inverse supercell size for the  $\text{V}_{\text{C}}^{+2}$  defect in diamond as obtained using different model charges in the FNV scheme, and (b–d) corresponding potentials from the 216-atom supercell calculations. We considered a Gaussian model charge (squares, dashed, red) with its corresponding DFT difference potential ( $V_{\text{DFT}}$ ) and model ( $V_{\text{model}}$ ) potential in (b), a model charge fitted to the defect wave function (plus symbols, dotted, magenta) with potentials in (c), and a scaled Gaussian (diamonds, dash-dotted, magenta) with potentials in (d). The difference potentials  $V_{\text{diff}} = V_{\text{DFT}} - V_{\text{model}}$  are also given in (b–d). In (a), the uncorrected formation energies, without (solid, black) and with (dotted, black) potential alignment, are also given.

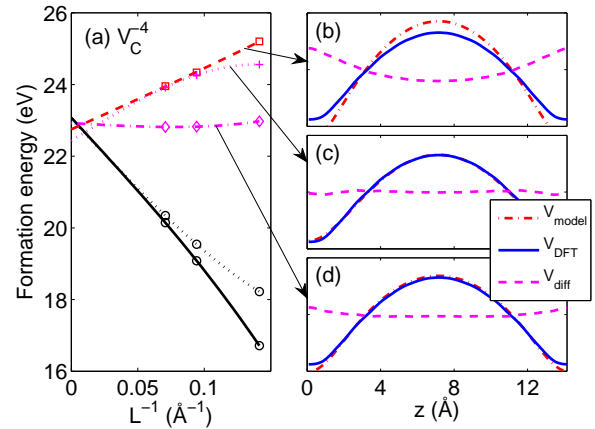


FIG. 10. (Color online) Same as in Fig. 9 but for the  $\text{V}_{\text{C}}^{-4}$  defect.

operator is carried out with a Brillouin-zone sampling in which the number of mesh points in each direction is halved compared to the employed  $\mathbf{k}$ -point sampling. We use the experimental lattice constant of 3.57 Å. We calculate the dielectric constants of diamond in both the PBE and PBE0 through the application of a finite electric field as defined within the modern theory of polarization,<sup>45</sup>



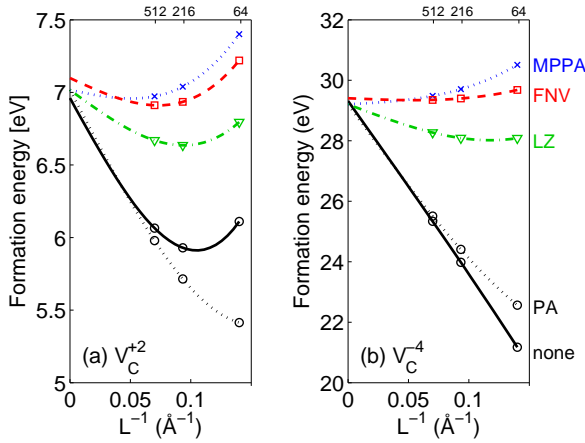


FIG. 11. (Color online) Formation energies vs. inverse supercell size for the  $V_C^{+2}$  and  $V_C^{-4}$  defects in diamond, as obtained with the PBE0 functional. Same notation as in Fig. 4.

finding 5.77 and 5.51, respectively. The direct band gaps at the  $\Gamma$  point are found to be 5.60 eV and 7.71 eV for PBE and PBE0, respectively. Thus, the PBE band gap is very close to the LDA band gap (5.62 eV) and noticeably underestimates the corresponding experimental band gap of 7.3 eV, while the PBE0 considerably opens the band gap bringing it in much closer agreement with experiment.

For the vacancy in diamond, the present PBE results are nearly identical to the LDA results described above, both in terms of formation energies and of the Kohn-Sham spectrum (not shown, cf. Figs. 5 and 7). The formation energies for  $V_C^{+2}$  obtained with the hybrid functional are given in Fig. 11(a). Compared to the PBE, the corrected formation energies are generally in much better agreement with the extrapolated values. The better agreement results from the fact that the separations between the uncorrected formation energies and their extrapolated value are larger than in the PBE. Similar considerations apply to the  $V_C^{-4}$  defect in Fig. 11(b).

The improved performance of the correction schemes in the PBE0 directly relates with the disappearance of the delocalized charge as shown in Fig. 12, where the charge densities in the PBE and PBE0 are compared. In the PBE0 the charge density level far from the defect closely corresponds to the  $(1 - 1/\epsilon)q/\Omega$  level, whereas this is not the case in the PBE. As illustrated in the figure, the difference between the dielectric constants in the PBE and PBE0 is insufficient to account for this behavior. The delocalized defect charge in excess with respect to the delocalized screening charge mainly results from the contribution of the defect wave function, as can be inferred from Fig. 12(c) for the  $V_C^{-4}$  defect.

One could argue that the higher localization observed in the PBE0 directly results from the reduced self-interaction in hybrid functional schemes. Indeed, hybrid functionals are known to cause Jahn-Teller distortions

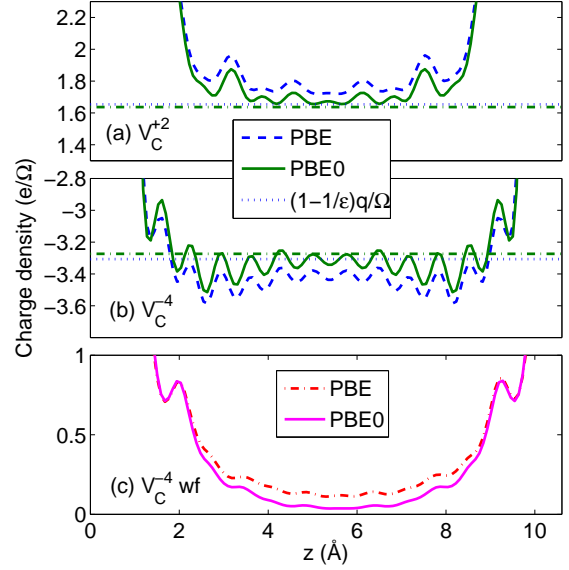


FIG. 12. (Color online) Planar-averaged charge density difference across the supercell for the (a)  $V_C^{+2}$  and (b)  $V_C^{-4}$  defects in diamond, as obtained with the PBE (dashed, blue) and PBE0 (solid, green) functionals. The difference is taken between the charge densities of the charged and neutral defects. The levels corresponding to the homogeneously distributed screening charge density  $(1 - 1/\epsilon)q/\Omega$  are indicated and differ in the PBE (dotted, blue) and PBE0 (dash-dotted, green) because of the different dielectric constants. (c) Planar-averaged charge distribution of the defect wave function of the  $V_C^{-4}$  in diamond, as obtained with the PBE (dashed-dotted, red) and PBE0 (solid, magenta) functionals.

in partially filled degenerate states, leading to enhanced localization with respect to semilocal functionals.<sup>46</sup> However, the present observations appear to have a different origin. In fact, in the  $V_C^{+2}$  defect, the degenerate manifold of the defect state is unoccupied and would thus give rise to an analogous defect state in PBE and PBE0, even if structural relaxation were allowed.<sup>38,47</sup>

The hybrid-functional results support the notion that the defect charge delocalization in the (semi)local schemes originates from hybridization between defect and band-edge states. It is interesting to investigate how the presence of such effects manifests in the Kohn-Sham energy spectrum. Kohn-Sham energy levels calculated in the PBE and PBE0 for four different charge states are shown in Fig. 13. In the case of PBE, the levels of the problematic  $V_C^{+2}$  and  $V_C^{-4}$  are closer to the band edges than those of the well-behaved  $V_C^0$  and  $V_C^{-2}$ . However, the former levels still fall well within the band gap for most supercells considered, despite the rapid evolution of the  $V_C^{-4}$  level towards the conduction band edge for increasing supercell size. At first sight, it is therefore not obvious that charge delocalization would occur for these charge states. Indeed, no apparent difference appears when comparing with the energy levels calculated

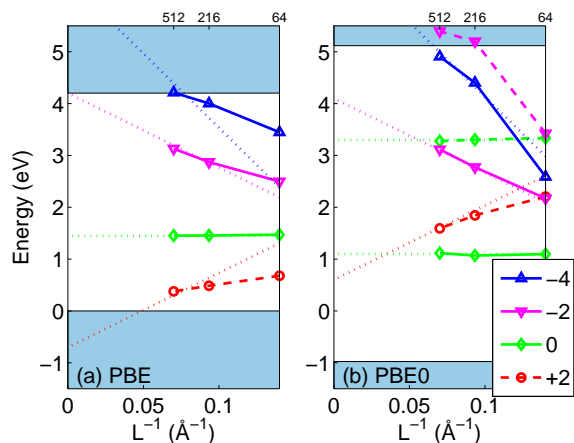


FIG. 13. (Color online) Kohn-Sham energy levels vs. inverse supercell size for various charge states of the vacancy defect in diamond, as obtained in (a) the PBE and (b) the PBE0. For each charge state, we show the occupied (solid) and unoccupied (dashed) energy levels in the band gap. The indicated level corresponds to the average over the defect band. In PBE, the unoccupied states of the neutral and  $-2$  charge states lie very close and are not distinguished. The dotted lines extrapolate the energy levels according to the scaling of the correction potential in the MP scheme.

in the PBE0, except for a separation between occupied and unoccupied states in  $V_C^0$  and  $V_C^{-2}$ .

However, the difference is revealed when the dependence of the Kohn-Sham energy levels with inverse supercell size is compared to that expected from the evolution of the first-order term of the correction potential given in Eq. (20):<sup>5,14,48</sup>

$$\epsilon_{\text{KS}}(L) - \epsilon_{\text{KS}}(L = \infty) \approx V_{\text{corr}}^{\text{MP}} = \frac{q\alpha}{\epsilon L}. \quad (37)$$

In Fig. 13, we compare the evolution of the Kohn-Sham levels with  $L^{-1}$  to the scaling of  $V_{\text{corr}}^{\text{MP}}$ . In the PBE, the scaling of the defect levels of the problematic  $V_C^{+2}$  and  $V_C^{-4}$  differs markedly from that of  $V_{\text{corr}}^{\text{MP}}$ . At variance, the  $V_C^{-2}$  and  $V_C^0$  levels behave as expected. It is thus the scaling behavior of the Kohn-Sham energy levels that can be seen as an indicator of charge delocalization rather than the position of their energy levels within the band gap. This correlation is supported by the hybrid functional results, where the calculated and the expected scalings are much closer, in accordance with the good performance of the energy correction schemes (cf. Fig. 11) and the absence of defect charge delocalization (cf. Fig. 12).

#### 4. Is delocalized defect charge physically meaningful?

We have shown that defect charge delocalization occurs in cases where the energy correction schemes do not perform optimally. We argued that this effect results

from defect Kohn-Sham levels lying in the proximity of the band edges and undergoing hybridization. The legitimate question arises whether the properties of such defect states are physically meaningful or whether we are confronted with unphysical features associated to the adopted electronic-structure scheme.

Let's consider first a defect showing charge delocalization within a semilocal DFT scheme. Such a situation might result either from a localized atomic state or from a shallow effective-mass-like state. In the former case, the delocalization is a consequence of the hybridization between the localized orbitals and the band-edge states. This effect then disappears when the band gap is opened as the separation between the defect level and the band-edge states is increased.<sup>40,49</sup> This is precisely what we observed in Sec. III C 3 when using hybrid functionals giving an improved description of the band gap. In this case, the semilocal DFT scheme is not expected to give a reliable description of the defect state. At variance, in the case of a shallow defect, the defect level would follow the band edge to which it is associated. The charge delocalization then results from the overlap of decaying defect wave functions from different periodic images and leads to quantum-mechanical interaction effects. In such cases, DFT supercells are generally too small to observe localization. It deserves further study to understand whether physical information about the isolated defect can be extracted in such circumstances. In either case, when charge delocalization is observed at the semilocal DFT level, a modeling approach based on simple electrostatics appears deficient.

However, one could imagine that the adopted electronic-structure scheme gives a faithful description of the experimental band gap, or, more generally, one might like to achieve a description of the defect state that is consistent with the adopted theoretical scheme. How should we handle a defect state showing charge delocalization at a finite value of the supercell size? We assume that a viable charge state of the defect carries integer charge in the limit  $L \rightarrow \infty$ . However, it is difficult to judge from the calculations at finite supercell size whether the delocalization would persist or disappear for  $L \rightarrow \infty$ . In the case of shallow defect, the adopted modeling scheme would not be appropriate as discussed above. In the alternative case in which the excess charge delocalization would persist in the limit  $L \rightarrow \infty$ , the defect state would not carry an integer charge and thus not be retained as a valid one.

We could also encounter charge defect states which do not show any sign of excess charge delocalization at finite  $L$ , but which could change their nature as the supercell is increased. To identify such cases, it is useful to extrapolate the calculated Kohn-Sham energy levels to the limit of infinite  $L$  through Eq. (37). In case its extrapolated energy level crosses a band edge, the defect state is expected to give rise to excess charge delocalization. For instance, the Kohn-Sham energy level of the  $V_C^{-4}$  defect in Fig. 13(b) is expected to hybridize with



the conduction band-edge states at supercell sizes larger than those considered. This situation would likely give rise to an invalid charge state in the limit of infinite  $L$ . In the unlikely case that the defect would turn into a shallow one, we would conclude that our modelling scheme is inappropriate. While the detection of excess delocalized defect charge is a robust indicator for a problematic defect, the extrapolation of the Kohn-Sham levels appears to be more effective for identifying valid charge states beyond the limit of those investigated with a finite simulation cell. We note that because of the dependence of  $q$  in Eq. (37), high-charge states are more likely to be discarded as valid defect states in the limit of infinite  $L$ .

In view of this discussion, it appears that, when the performance of energy correction schemes is not optimal, the concerned defect state either corresponds to an invalid one or to one that cannot properly be treated within the adopted DFT scheme. In such case, it therefore does not seem appropriate to spend efforts in modeling the extended defect charge through sophisticated charge model distributions, as would be possible in the FNV scheme. This consideration relieves the problem that the energy corrections strongly depend on the adopted model charge for this kind of defect states. At variance, when the defect state shows well-behaved charge localization, the energy correction in the FNV scheme is the same, irrespective of the adopted model charge.

#### IV. CONCLUSIONS

We first reviewed several of the most commonly used schemes for correcting finite-size supercell errors in charged defect calculations. In particular, we focused on the schemes proposed by (i) Makov and Payne,<sup>8</sup> (ii) Freysoldt, Neugebauer, and Van de Walle,<sup>10,15</sup> and (iii) Lany and Zunger,<sup>9</sup> and discussed the role of potential alignment in neutral and charged defect calculations. In our analysis, we elaborated a connection between the defect charge distribution and the potential alignment, which establishes a relation between the Makov-Payne and the Freysoldt-Neugebauer-VandeWalle schemes. Our comprehensive description also allowed us to make contact in a consistent manner with contributions due to Dabo *et al.*<sup>14</sup> and Taylor and Bruneval.<sup>5</sup>

In the second part of our investigation, we compared the performance of these correction schemes for a large variety of defects. We identified a class of defects for which the available correction schemes give good results. These correspond to defect states in which the charge density is well localized within the supercell. The Makov-Payne scheme works well for point-like defects, but results in overcorrections when the defect charge density is more extended. The Lany-Zunger scheme tends to undercorrect formation energies of defects with very localized charge distributions, but improves for defects with more extended charge distributions. The Freysoldt-Neugebauer-VandeWalle scheme was found to give good

formation energies for any defect with a localized charge distribution, without requiring particular attention in the choice of the model charge. Overall, based on our error estimators, this scheme appears to offer the best performance among those considered.

We also encountered a class of defects in which the available schemes failed to produce satisfactory results. These defects show charge delocalization in excess with respect to the level expected from the delocalized screening charge. The origin of this effect was assigned to the hybridization between the defect state and the band-edge states of the host. As far as the Kohn-Sham levels are concerned, they were mostly found to locate within the band gap. However, when excess delocalization occurs, their evolution with supercell size becomes anomalous. The evolution of the Kohn-Sham energy levels can also be used as an indicator for identifying defect states which are expected to show charge delocalization at larger supercell sizes than those actually considered in the calculations.

In conclusion, we found that the Freysoldt-Neugebauer-VandeWalle scheme is robust and meaningful for all defects showing well-behaved charge localization. At variance, the performance of this scheme for defects with excess charge delocalization is more ambiguous. Indeed, in the latter case, it is legitimate to question to what extent the study of defects with excess charge delocalization is physically meaningful. We concluded that these cases either refer to defect charge states which are not physically relevant or to situations for which the adopted DFT modeling scheme is inappropriate. In either case, such defect states do not seem to warrant further investigation within the adopted modeling scheme.

#### ACKNOWLEDGMENTS

Financial support is acknowledged from the Swiss National Science Foundation (Grants Nos. 200020-119733/1, 200020-134600, and 206021-128743) and Tekniikan edistämmissäätiö and Emil Aaltosen säätiö (H.-P.K.). We used computational resources of CSC (Finland) and CSEA-EPFL.

#### APPENDIX: POTENTIAL ALIGNMENT AND SECOND RADIAL MOMENT OF A MODEL CHARGE

In Sec. IID we demonstrated the connection between the potential alignment and the charge distribution for any charge distribution. Nevertheless, it is instructive to explicitly illustrate this equivalence for the commonly used model charge distribution consisting of a linear combination of a Gaussian and an exponential

function:<sup>10</sup>

$$\rho_c(r) = qxN_\gamma e^{-r/\gamma} + q(1-x)N_\beta e^{-r^2/\beta^2}, \quad (38)$$

where the normalization constants are  $N_\beta = (\pi\beta^2)^{-3/2}$  and  $N_\gamma = (8\pi\gamma^3)^{-1}$ , and the distribution approaches a  $\delta$ -function, as  $\beta \rightarrow 0$  and  $\gamma \rightarrow 0$ . The potential associated to the isolated  $\rho_c$  distribution is

$$V(r) = \frac{qx}{r} \left[ 1 - \frac{1}{2} e^{-r/\gamma} \left( \frac{r}{\gamma} + 2 \right) \right] + q(1-x) \frac{\text{erf}(r/\beta)}{r} \quad (39)$$

We evaluate the constant potential shift  $\Delta V_{\text{near}}$  with respect to the point-charge potential by averaging the difference potential  $V(r) - 1/r$  over the supercell assuming that the charge density is well localized therein:

$$\Delta V_{\text{near}} = x \frac{8\pi\gamma^2 q}{\Omega} + (1-x) \frac{\pi\beta^2 q}{\Omega}. \quad (40)$$

The shift in the simulation cell at a large distance from the defect is then given by  $\Delta V_{q/0} = -\Delta V_{\text{near}}$  according to Eq. (3). This shift can equivalently be obtained through the  $\mathbf{G} = 0$  Fourier component of the model potential.<sup>15</sup>

On the other hand, the change in the MP energy when going from a point-charge to the considered model charge reduces to the MP2 term. The second radial moment of the model charge  $\rho_c$  yields:

$$Q = xq(12\gamma^2) + (1-x)q \left( \frac{3}{2}\beta^2 \right) \quad (41)$$

Plugging this into the MP2 term of Eq. (8) then yields the associated energy correction:

$$E_{\text{corr}} = -\frac{2\pi q Q}{3\Omega} = -x \frac{8\pi\gamma^2 q^2}{\Omega} - (1-x) \frac{\pi\beta^2 q^2}{\Omega}, \quad (42)$$

which is exactly equivalent to  $q\Delta V_{q/0}$ . This also verifies that the potential shift obtained directly from Eq. (40) is equal to that appearing in the MP correction potential of Eq. (21).

- 
- <sup>1</sup> *Advanced Calculations for Defects in Materials: Electronic Structure Methods*, edited by A. Alkauskas, P. Deák, J. Neugebauer, A. Pasquarello, and C. G. Van de Walle (Wiley VCH, 2011).
- <sup>2</sup> C. W. M. Castleton and S. Mirbt, *Phys. Rev. B* **70**, 195202 (2004).
- <sup>3</sup> C. W. M. Castleton, A. Höglund, and S. Mirbt, *Phys. Rev. B* **73**, 035215 (2006).
- <sup>4</sup> N. D. M. Hine, K. Frensch, W. M. C. Foulkes, and M. W. Finnis, *Phys. Rev. B* **79**, 024112 (2009).
- <sup>5</sup> S. E. Taylor and F. Bruneval, *Phys. Rev. B* **84**, 075155 (2011).
- <sup>6</sup> P. A. Schultz, *Phys. Rev. Lett.* **84**, 1942 (2000).
- <sup>7</sup> M. Leslie and M. J. Gillan, *J. Phys. C: Solid State Phys.* **18**, 973 (1985).
- <sup>8</sup> G. Makov and M. C. Payne, *Phys. Rev. B* **51**, 4014 (1995).
- <sup>9</sup> S. Lany and A. Zunger, *Phys. Rev. B* **78**, 235104 (2008).
- <sup>10</sup> C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, *Phys. Rev. Lett.* **102**, 016402 (2009).
- <sup>11</sup> C. G. Van de Walle and J. Neugebauer, *J. Appl. Phys.* **95**, 3851 (2004).
- <sup>12</sup> S. B. Zhang and J. E. Northrup, *Phys. Rev. Lett.* **67**, 2339 (1991).
- <sup>13</sup> F. Corsetti and A. A. Mostofi, *Phys. Rev. B* **84**, 035209 (2011).
- <sup>14</sup> I. Dabo, B. Kozinsky, N. E. Singh-Miller, and N. Marzari, *Phys. Rev. B* **77**, 115139 (2008).
- <sup>15</sup> C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, *Phys. Status Solidi B* **248**, 1067 (2011).
- <sup>16</sup> E. Madelung, *Phys. Zs.* **19**, 524 (1918).
- <sup>17</sup> P. P. Ewald, *Annalen der Physik* **369**, 253 (1921).
- <sup>18</sup> Note that the term depending on  $Q$  appeared erroneously with the opposite sign in Ref. 8.
- <sup>19</sup> S. Lany and A. Zunger, *Modelling Simul. Mater. Sci. Eng.* **17**, 084002 (2009).
- <sup>20</sup> The dipole term in Eq. (13) of Ref. 14 vanishes for the origin used in the definition of  $Q$  in the MP scheme (Ref. 8).
- <sup>21</sup> J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, 1998) ISBN 047130932X.
- <sup>22</sup> I. Dabo, B. Kozinsky, N. E. Singh-Miller, and N. Marzari, *Phys. Rev. B* **84**, 159910(E) (2011).
- <sup>23</sup> G. Kresse and J. Hafner, *Phys. Rev. B* **47**, 558 (1993).
- <sup>24</sup> G. Kresse and J. Furthmüller, *Comput. Mat. Sci.* **6**, 15 (1996).
- <sup>25</sup> G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
- <sup>26</sup> S. Baroni and R. Resta, *Phys. Rev. B* **33**, 7017 (1986).
- <sup>27</sup> M. Gajdos, K. Hummer, G. Kresse, J. Furthmüller, and F. Bechstedt, *Phys. Rev. B* **73**, 045112 (2006).
- <sup>28</sup> *Numerical data and Functional Relationships in Science and Technology, Landolt-Börnstein*, edited by K.-H. Hellwege, O. Madelung, M. Schulz, and H. Weiss (Springer-Verlag, 1982).
- <sup>29</sup> V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, *Phys. Rev. B* **69**, 075102 (2004).
- <sup>30</sup> S. Adachi, *Optical Constants of Crystalline and Amorphous Semiconductors: Numerical Data and Graphical Information* (Kluwer Academic Publishers, 1999) ISBN 9780792385677.
- <sup>31</sup> J. Fontanella, R. L. Johnston, J. H. Colwell, and C. Andeen, *Appl. Opt.* **16**, 2949 (1977).
- <sup>32</sup> A. Mooradian and A. L. McWorter, *Phys. Rev. Lett.* **19**, 849 (1967).
- <sup>33</sup> J. R. Jasperse, A. Kahan, J. Plendl, and J. N. Mitra, *Phys. Rev.* **146**, 526 (1966).
- <sup>34</sup> M. J. Puska, S. Pöykkö, M. Pesola, and R. M. Nieminen, *Phys. Rev. B* **58**, 1318 (1998).
- <sup>35</sup> J. Shim, E.-K. Lee, Y. J. Lee, and R. M. Nieminen, *Phys. Rev. B* **71**, 245204 (2005).
- <sup>36</sup> W. Chen, C. Tegenkamp, H. Pfnür, and T. Bredow, *Phys. Rev. B* **82**, 104106 (2010).
- <sup>37</sup> J. Shim, E.-K. Lee, Y. J. Lee, and R. M. Nieminen, *Phys. Rev. B* **71**, 035206 (2005).
- <sup>38</sup> H.-P. Komsa and A. Pasquarello, *Phys. Rev. B* **84**, 075207 (2011).
- <sup>39</sup> F. Giustino and A. Pasquarello, *Phys. Rev. B* **71**, 144104 (2005).
- <sup>40</sup> A. Alkauskas, P. Broqvist, and A. Pasquarello, *Phys. Status Solidi B* **248**, 775 (2011).
- <sup>41</sup> A. Alkauskas and A. Pasquarello, *Phys. Rev. B* **84**, 125206 (2011).
- <sup>42</sup> J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- <sup>43</sup> J. P. Perdew, M. Ernzerhof, and K. Burke, *J. Chem. Phys.* **105**, 9982 (1996).
- <sup>44</sup> P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougousis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Scaluzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- <sup>45</sup> P. Umari and A. Pasquarello, *Phys. Rev. Lett.* **89**, 157602 (2002).
- <sup>46</sup> G. Pacchioni, F. Frigoli, D. Ricci, and J. A. Weil, *Phys. Rev. B* **63**, 054102 (2000).
- <sup>47</sup> A. Alkauskas and A. Pasquarello, *Physica B* **401**, 670 (2007).
- <sup>48</sup> S. Lany and A. Zunger, *Phys. Rev. B* **81**, 113201 (2010).
- <sup>49</sup> A. Alkauskas, P. Broqvist, and A. Pasquarello, *Phys. Rev. Lett.* **101**, 046405 (2008).