Relation between image charge and potential alignment corrections for charged defects in periodic boundary conditions

T. R. Durrant, S. T. Murphy, M. B. Watkins, and A. L. Shluger

---

Articles you may be interested in

Perspective: Basic understanding of condensed phases of matter via packing models
The Journal of Chemical Physics **149**, 020901 (2018); 10.1063/1.5036657

Adaptive resolution molecular dynamics technique: Down to the essential
The Journal of Chemical Physics **149**, 024104 (2018); 10.1063/1.5031206

Linear response time-dependent density functional theory without unoccupied states: The Kohn-Sham-Sternheimer scheme revisited
The Journal of Chemical Physics **149**, 024105 (2018); 10.1063/1.5030652

Communication: Truncated non-bonded potentials can yield unphysical behavior in molecular dynamics simulations of interfaces
The Journal of Chemical Physics **147**, 121102 (2017); 10.1063/1.4997698

Communication: Coupled cluster and many-body perturbation theory for fractional charges and spins
The Journal of Chemical Physics **148**, 221103 (2018); 10.1063/1.5040164

Aqueous solvation from the water perspective
The Journal of Chemical Physics **148**, 234505 (2018); 10.1063/1.5034225

Relation between image charge and potential alignment corrections for charged defects in periodic boundary conditions

T. R. Durrant,[1] S. T. Murphy,[1,2] M. B. Watkins,[1,3] and A. L. Shluger[1]

[1]*Department of Physics and Astronomy and London Centre for Nanotechnology, University College London, Gower Street, London WC1E 6BT, United Kingdom*
[2]*Engineering Department, Lancaster University, Lancaster LA1 4YW, United Kingdom*
[3]*School of Mathematics and Physics, University of Lincoln, Lincoln LN6 7TS, United Kingdom*

Charged defects are often studied within the periodic density functional theory (DFT), but this introduces strong finite-size artifacts. In this work, we develop an electrostatic image interaction correction (IIC) method based on the direct solution of the Poisson equation for charge models constructed directly from DFT calculations. These IICs are found to be detail-insensitive, depending almost entirely on bulk dielectric properties. As these IICs are not able to fully explain the observed finite-size scaling, we explore potential alignment in detail and introduce a novel decomposition to separate out different contributions. We find that the two main sources of potential alignment are defect image interactions and changes in the number of atoms present in the supercell. This first effect is accurately predicted by the periodic part of our IIC. The second contribution is unrelated to the IIC and justifies the common observation that the magnitude of finite-size dependence can strongly vary between vacancy and interstitial defects. It can be approximately predicted using atomic radius, but is strongly sensitive to the pseudopotential employed. Combined, these developments provide a new justification for known finite-size scaling rules. Our results suggest that for cubic supercells, the Lany-Zunger IIC, combined with simplified potential alignment between neutral systems, can yield accurate corrections in spite of the simplicity of the approach. © 2018 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/1.5029818

I. INTRODUCTION

Accurate defect formation energies are vital to understanding the properties of materials and electronic devices, as they are required to calculate defect concentrations and charge transition levels.[1] Defect calculations are often carried out using periodic boundary conditions in classical simulations[2] or at different levels of many-electron theory, notably density functional theory (DFT). Periodic translations allow defects to interact with their own images in the neighboring cells of the simulation. This leads to strong electrostatic interactions between a defect and its images when the defect has net charge.[3] The Coulomb interaction is long ranged, and these image interactions may remain significant even when relatively large supercells are used. The use of periodic boundary conditions also creates an additional problem that the bulk electrostatic potential is not properly recovered far from the defect.[4]

In practice, it is not possible to use sufficiently large supercells in which these effects can be neglected, especially with the high computational costs associated with hybrid functionals, now in common usage.[5] This problem is typically solved by applying *post hoc* corrections to the final energy. In this way, highly developed standard DFT codes can still be employed at a reasonable computational cost.

Post hoc correction methods have been developing over the last 25 years, and there are now many approaches available and compared in the literature (see, e.g., Refs. 3 and 6–13 which is by no means a complete list). Detailed comparison of different methods is complicated by a realization that image interactions contribute to the need of potential alignment, meaning that the two problems must be considered together.[8–10] It seems likely that some of the suggested techniques are reliable, but many disagreements in the literature mean that the current position is rather unclear, especially for non-experts. This is perhaps one of the reasons why it has recently been suggested that formation energies can only be rigorously calculated for neutral defects.[14]

By contrast, the actual finite-size scaling of charged defect formation energies has been observed to exhibit some common features across many kinds of materials and defects. For cubic systems, for example, the main errors in the formation energy are observed to scale in terms of the length and volume of the supercell.[15] Furthermore, vacancy and interstitial defects in the same charge state often appear to demonstrate quite different finite-size scaling.[10,13] The origins of these observations are not yet fully understood. In this paper, we attempt to separate image interaction corrections and potential alignment corrections and to use this approach to provide a new rational for these observations.

First, we introduce a new image interaction correction (IIC), where we solve the Poisson equation directly for the additional charge present in the DFT supercell. This enables us to explore the degree of finite-size dependence that is

introduced by electrostatic interactions with images and the inclusion of a jellium background, without recourse to an approximate description of the image charge distribution. Such an approach predicts that these image interactions introduce a length dependent finite-size error, for the case of cubic supercells.

Next, we consider potential alignment and introduce a new decomposition of the total potential alignment into three parts. We find that the image correction we constructed exactly predicts one of these parts, leaving two other components that are unrelated to image interactions. These alignment components are introduced by the removal (or addition) of atoms from the cell, and ionic relaxation of the cell, respectively. We observe that the first component can be predicted by calculations on isolated atoms and that the potential alignment correction required for isolated atoms scales with the volume of the supercell. Additionally, this potential alignment correction has an opposite sign between vacancies and interstitials. It is observed that inclusion of the second component often introduces additional finite-size dependence.

We apply these methods to vacancy defects in three materials, diamond, MgO, and SnO_2, where our examples are chosen both to represent materials with different types of bonding and defect states that are quantitatively different. The relative magnitude of the IICs and potential alignment corrections is substantively different between these examples. However, the two corrections outlined above are shown to be sufficient, without any empirical parameters.

The rest of the paper is organized as follows. In Sec. II, we give a brief overview of finite-size dependence and some of the existing approaches to calculating image charge corrections. Then in Sec. III, we propose and justify a new image charge correction method. In Sec. IV, we introduce a new decomposition of the potential alignment. In Sec. V, we apply these developments to defect calculations in diamond, MgO, and SnO_2 and compare our results with those obtained using several common correction methods. The discussion and conclusions of our findings are provided in Sec. VI.

II. CALCULATION OF DEFECT FORMATION ENERGIES

A. Cell size dependence

One standard definition[16] of the defect formation energy E_f is given as

$$E_f = E_{\text{defect}}^{\text{DFT}} - E_{\text{bulk}}^{\text{DFT}} - \sum_i n_i \mu_i$$
$$+ E_{\text{IIC}} + q(\mu_e + \Delta V_{\text{NAP}}), \quad (1)$$

where $E_{\text{defect}}^{\text{DFT}}$ and $E_{\text{bulk}}^{\text{DFT}}$ are the total energies of the supercell containing the defect and an equivalently shaped supercell of the perfect bulk crystal, respectively. Additionally, this raw energy difference is modified by the energy required to exchange n_i ions with some reservoir of chemical potential μ_i. The first three terms of Eq. (1) are often fully adequate for calculating the formation energy of neutral defects, although it neglects other sources of errors such as elastic relaxation or defect state hybridization. Additional terms which describe these effects have been suggested in Refs. 10 and 17.

When a defect has net charge q, additional terms appear in Eq. (1). First, the energy cost to exchange q electrons with an electron chemical potential μ_e must be applied. We follow the usual convention and use the valence band maximum (VBM) of the bulk crystal as our chemical potential such that $\mu_e = E_{\text{VBM}}$. Second, the spurious interaction energy E_{IIC} due to the extra interaction of the defect both with its own charged images and the neutralizing jellium background (neutralizing charge is required in all charged periodic calculations to avoid a divergence of the total energy) must be removed. Finally, the average electrostatic potential cannot be directly calculated using periodic DFT and is conventionally set to be zero.[18–20] This neglect of the average potential (NAP) introduces a numerical error into the formation energy of the charged defect. A potential alignment correction $q\Delta V_{\text{NAP}}$ is often applied to account for this problem.

The two terms E_{IIC} and ΔV_{NAP} are finite-size corrections, which tend to zero as the size of the DFT simulation cell is increased, and are the focus of this paper. Their performance can be assessed by exploring the convergence of E_f with increasing supercell size. Ideally, once corrected, there would no longer be a supercell size dependence in E_f. Hence, the simplest measure of the success of these corrections is the supercell size dependence of the corrected energies.

Additionally, detailed studies on the supercell size dependence of the formation energy of charged defects have been carried out. The study of Castleton *et al.*[15] has produced the following empirical scaling rule for cubic supercells:

$$E_f(L) = E_f(L \rightarrow \infty) + \frac{a_1}{L} + \frac{a_3}{L^3}, \quad (2)$$

which demonstrates an inverse size dependence on both the length of the considered supercell, L, and the volume of the supercell $\Omega = L^3$. It is particularly useful if the physical origins of terms a_1 and a_3 can be identified, and several plausible suggestions have been offered in the literature.[10,13] [We note the similarity between Eqs. (2) and (4), which was discovered earlier.[3] However, we will later argue that these two equations are distinct.]

The expression in Eq. (2) is often used to extrapolate charged formation energies to the dilute limit.[15] This is then used to test the performance of charge correction methods in several papers.[6,9–11] We do not report these extrapolations in this paper, as we believe that the uncertainties in these extrapolations are about the same as the differences between the correction methods considered. Extrapolating a three parameter model from four supercells provides one point that can be exuded from the fit, to check the consistency of the extrapolation, but we always observe a poor prediction of this point. As we will later justify (in Table I), this is because the L^{-3} scaling of the a_3 term is only approximately obeyed, causing the accuracy of this extrapolation to vary between the different defects considered.

B. Existing image charge corrections

As the use of periodic boundary conditions replaces an isolated charged defect with an array of periodic images,

calculations using the supercell methodology also include an energy error (if one is interested in the dilute limit) due to the interactions with images, E_{IIC}. In order to remove these unwanted interactions, they can be modeled separately using classical electrostatics and then be removed. This is shown pictorially in Fig. 1.

The simplest electrostatic model that can be applied to construct an IIC is approximating the defect as a point charge (PC). In Fig. 1, this represents the charges q as ideal points. These corrections were first applied to defect calculations using classical molecular dynamics by Leslie and Gillan.[2] In the context of DFT, this is the first-order Makov-Payne (MP) correction for cubic supercells,[3]

$$E_{IIC}^{MP} = \frac{q^2 \alpha_M}{2\varepsilon L}, \tag{3}$$

where ε is the dielectric constant of the material containing the defect, L is the separation between defects, and α_M is the Madelung constant, which depends only on the shape of the supercell containing the defect. This provides the exact interaction energy of an array of point charges immersed in neutralizing jellium. The appropriate dielectric constant is the static dielectric constant ε_0 when ionic relaxation of the defect is performed or the high frequency dielectric constant ε_∞ when the ionic structure is kept fixed. This simple method has been extended to general supercell shapes and to anisotropic dielectrics.[21] For cubic cells, this produces a L^{-1} dependent term that could explain the origin of the a_1 term in Eq. (2).

However, a point charge is a highly idealized model, which is known to overestimate the size of the unwanted image interactions. Charged defects are often not point-like, but instead have charge distributed over some defect dependent regions, such as a vacancy site, dangling bonds, or neighboring ions. Additionally, the finite size of the supercell may affect the degree of dielectric screening.

Makov and Payne[3] were motivated by these considerations to suggest a multipole expansion to generate higher-order terms that could also be calculated. Such an expansion yields

$$E_{IIC}^{MP2} = \frac{q^2 \alpha_M}{2\varepsilon L} - \frac{2\pi q Q}{3 L^3} + O(L^{-5}), \tag{4}$$
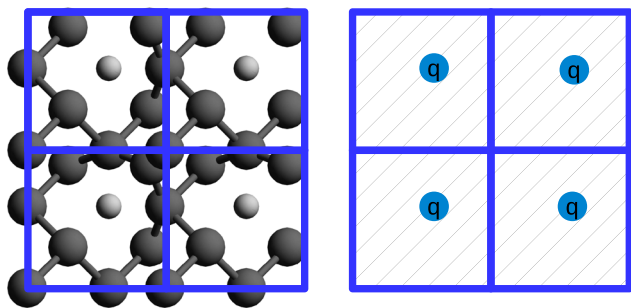

FIG. 1. A periodic DFT calculation containing a defect is shown on the left, which is used to construct the simplified classical model on the right. The model charge q interacts through a classical dielectric background, represented by the shaded background. The interaction of the classical model is removed from the final formation energy. Two possible models for q are shown in Fig. 2.

where Q is the quadrupole moment of the defect charge model. This expression contains only odd L terms because the even L terms are zero for charge distributions with cubic symmetry. Hence, Eq. (4) is correct up to the fifth-order for cubic supercells. However, the MP paper[3] does not provide a systematic way of calculating Q, so this quadrupole term is rarely calculated within the MP method.

The Lany-Zunger (LZ) correction[6,7] directly attacks this problem by making a practical approximation to Q. This method is derived by assuming that the defect draws bound screening charge, as demanded by classical electrostatics[52,53]. This bound charge is drawn uniformly from the supercell, leading to a uniform charging given by

$$q_{screen}^{LZ}(\mathbf{r}) \approx \frac{q}{\Omega}\left(1 - \frac{1}{\varepsilon}\right). \tag{5}$$

This charge present across the whole supercell is expected to be the dominant charge component far from the defect (note this is not the whole bound charge), is used to define the quadrupole moment Q, and leads to a reduction in the magnitude of the correction. In fact, this is shown to reduce the MP correction by about a third. Charge models constructed from DFT density differences are found to include this bound charge, as shown in Fig. 2(a).

Generally, we may express this reduction in the IIC by defining an internal polarization ratio f. The MP correction is too large and needs to be reduced by a scaling factor. Then, the complete interaction correction is given as

$$E_{IIC} = f E_{IIC}^{MP}. \tag{6}$$

The LZ method provides a practical approximation to f in the form

$$f^{LZ} = 1 - c_{sh}\left(1 - \varepsilon^{-1}\right), \tag{7}$$

where the shape factor, c_{sh}, depends only on the shape of the defect containing the supercell. These factors have been calculated for common supercells.[7] As the LZ method is derived from Eq. (4), it is also partly underpinned by an assumption of a cubic supercell. For cubic cells with a high dielectric constant ($\varepsilon \to \infty$), this yields $f^{LZ} = 0.631$. By contrast, the original MP correction is recovered when $\varepsilon \to 1$. As the LZ method reduces to a scaling of the original MP correction, it also leads to a L^{-1} dependent correction that could produce the a_1 term in Eq. (2).

The interpretation of the LZ IIC is simple—the finite size of the supercell prevents ideal screening behavior being observed. Instead, the expected dielectric screening is enhanced by the electrostatic bound charge present in the supercell. This additional screening depends on the magnitude of the bound charge, which in turn depends on the dielectric constant. The LZ method requires the dielectric of the system to be constant and isotropic, presently limiting its application to bulk systems.

We note that the LZ method was also introduced with a potential alignment technique based on atomic site averages, performed over a large volume of the cell. Hence, when LZ results are presented in the literature, they may also include this alignment procedure. More complex methods for performing potential alignment fall beyond the scope of this
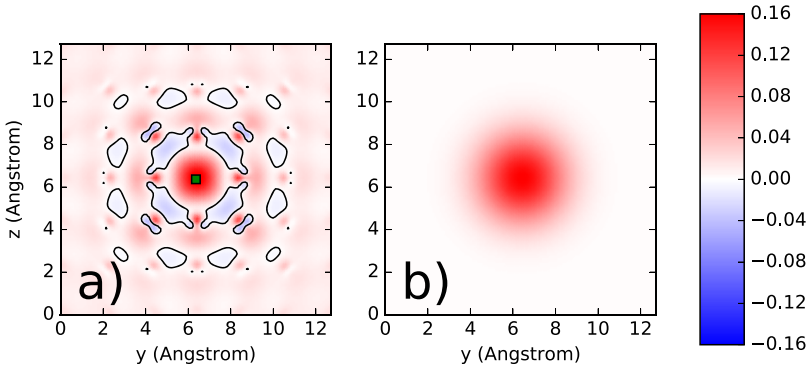
FIG. 2. Comparison of defect charge models for the V_O^{+2} defect in MgO. The total charge is plotted in y and z. (a) is a model constructed as a charge density difference between two DFT calculations, and (b) is a Gaussian model, as commonly used in the FNV method. The zero density isosurface is shown.

paper, and we limit ourselves to planar averages. Additionally, we will later argue that the LZ IIC does well describe the image electrostatics, which encourages simpler alignment procedures.

Another distinct conceptual approach is the defect correction method of Freysoldt, Neugebauer, and Van de Walle (FNV).[8] The FNV method is different to the other approaches as both the IIC and the potential alignment problem are tackled simultaneously. In this method, the charge induced by the defect is modeled as a Gaussian distribution like that shown in Fig. 2(b) (or other simple localized distribution), which normally make only a very minor change to the interaction strength E_{IIC}, and produces an energy almost identical to E_{IIC}^{MP}. Then, the electrostatic potentials of the model system and the DFT calculation are matched far from the defect, introducing extra polarization effects into the correction.

In this method, the correction is given as

$$E_{IIC/PA}^{FNV} = E_{IIC}^{FNV} - q\left(\Delta\tilde{V}_{q/0}\right), \tag{8}$$

where E_{IIC}^{FNV} is the interaction energy of the defect charge model used and $\Delta\tilde{V}_{q/0}$ is the shift in potential between the charged and neutral defect, beyond that introduced by the calculation of E_{IIC}^{FNV}. (Potential alignment is discussed in more detail in Sec. IV.) Equation (8) contains two terms, where E_{IIC}^{FNV} can be identified with the a_1 term and $-q\left(\Delta\tilde{V}_{q/0}\right)$ can be associated with the a_3 term in Eq. (2).

Out of the considered methods, the FNV approach offers the most flexibility because both the charge model used and the region over which potential alignment is calculated can be varied, effectively leading to a whole family of corrections. These options can allow some fine tuning of the magnitude of the correction to the specific defect being studied. More recent work on the method has looked at fitting these charge models to defect state orbitals directly.[22]

The FNV method has been extended to surfaces and interfaces.[23] However, it can be complicated in practice by the effects of atomic relaxation, as the movement of ions introduces long range changes in the potential not related to alignment effects.[9,11] The Kumagai-Oba (KO) method[9] has been suggested to circumvent these limitations.

One important question is whether we should expect IICs to have different magnitudes for positively and negatively charged defects. The corrections based on analytic theory (MP and LZ) are identical for positively and negatively charged defects. The FNV method is observed to break this

symmetry when it is applied, a flexibility introduced by the $\Delta\tilde{V}_{q/0}$ term.

III. PROPOSED IMAGE CHARGE CORRECTION

Above, we have seen electrostatic methods for improving the MP point charge description of defect-defect interactions. In this work, we calculate interaction corrections from DFT electronic densities directly, rather than approximating the defect charge distribution.

To do this, we follow very similar conceptual lines to the density countercharge (DCC) method of Dabo *et al.*,[24] where the charge due to the defect is modeled by an explicit charge distribution and a full electrostatic model system is solved (in the DCC method, corrections are then applied self-consistently within the DFT calculation, in contrast to our approach where corrections are applied after the calculation). In the DCC method, the spurious interaction with images is defined as

$$E_{IIC} = \frac{1}{2}\int q(\mathbf{r})v_{IIC}(\mathbf{r}), \tag{9}$$

where the interaction energy E_{IIC} is defined as being caused by the model charge $q(\mathbf{r})$ interacting electrostatically with its own images. This interaction potential v_{IIC} is broken down into two contributions, which can be computed more easily than evaluating v_{IIC} directly. These two components are

$$v_{IIC}(\mathbf{r}) = v_{iso}(\mathbf{r}) - v_{PBC}(\mathbf{r}). \tag{10}$$

These two terms are the potentials produced by the charge distribution $q(\mathbf{r})$, when it is interacting with itself in the absence of other charges, v_{iso}, and when it is surrounded by periodic images, v_{PBC}. We note that the periodic potential v_{PBC} also includes the contribution of a jellium background.

Previously, the DCC method has been developed for molecules in vacuum, where the dielectric $\varepsilon = 1$ through all space. In order to produce a *post hoc* correction for bulk materials, we need to include the effect of the dielectric of the host material. The appropriate version of the Poisson equation to use is

$$\nabla[\varepsilon(\mathbf{r})\nabla v(\mathbf{r})] = q(\mathbf{r}), \tag{11}$$

where the model charge distribution q is embedded in a host crystal with dielectric tensor ε, producing the potential v.

In the original DCC paper, the whole charge distribution is used in the corrective electrostatic calculation. But to apply similar ideas to a periodic material, we must distinguish

between electronic charge which is present in the bulk system (which in turn contributes to the dielectric properties) from that induced by the defect itself. In order to do this, we need to define the model charge distribution $q(\mathbf{r})$. When ions are not allowed to relax, this can be simply defined as the electronic density difference between the neutral defect and the charged defect in state q. This yields a charge distribution with total magnitude q,

$$q(\mathbf{r}) = n_{\mathrm{q=q}}^{\mathrm{defect}}(\mathbf{r}) - n_{\mathrm{q=0}}^{\mathrm{defect}}(\mathbf{r}), \qquad (12)$$

where $n_{\mathrm{q=q}}^{\mathrm{defect}}$ is the electronic density of the supercell containing the charged defect and $n_{\mathrm{q=0}}^{\mathrm{defect}}$ is the electronic density when the defect is in a charge neutral state. If relaxation around the defect is not allowed, the atomic structure is identical in both calculations.

When the defect geometry is allowed to relax, there is more flexibility in the construction of the defect charge model, q. In this paper, we fully relax the atomic geometry of the charged defect in the calculation of $n_{q=q}^{\mathrm{defect}}$. Then, we use this same ionic geometry in the calculation of $n_{q=0}^{\mathrm{defect}}$. This matches the ionic structure that is present in the image cells. Effectively, we replace charged defects in the image cells with neutral defects of the same ionic structure. If the geometry is also relaxed in the calculation of $n_{q=0}^{\mathrm{defect}}$, large fluctuations are introduced into q, reflecting the movement of ions.

It must be stressed that these charge models do not reflect just the electrons added (or removed) from the system, but also contain information about the polarization of the supercell as a whole.

This charge model includes a jellium background only implicitly and hence is of magnitude q. When the interaction potential v_{IIC} is calculated, a jellium background is effectively introduced in exactly the same way as for DFT calculations. This is because we follow the same convention as the DFT calculation that generated the charge model[19,20] and set the average value of the periodic part of the interaction correction model v_{PBC} to zero, in the same way as the average electrostatic potential V within the DFT calculation is set to zero.

Once $\varepsilon(\mathbf{r})$ and $q(\mathbf{r})$ are defined, the interaction energy can be calculated through Eq. (11). In this paper, we use the open source libraries DL_MG[25,26] and PSPFFT[27] to solve these electrostatic problems directly, without the introduction of further approximations. DL_MG is an iterative multigird solver.[28] PSPFFT is an FFT based solver for isolated boundary conditions.

IV. POTENTIAL ALIGNMENT

In Secs. II B and III, we have discussed interaction corrections E_{IIC}. However, it is readily observed that the use of simple interaction corrections (such as MP) is not always successful by itself. Significant improvements in cell size convergence can be seen if a potential alignment correction, $q\Delta V_{\mathrm{NAP}}$, is also applied.[4,11] How much of this is due to failures in interaction corrections E_{IIC} is not clear. This is a question of how much of V_{NAP} is introduced by the same image interactions treated in E_{IIC}. It has been suggested that no potential

alignment should be required as long as E_{IIC} can be determined with sufficient accuracy.[9]

The need for potential alignment is introduced into periodic DFT calculations by NAP, the convention of setting the average electrostatic potential to zero,[18–20] $\langle V \rangle = 0$. NAP can introduce non-physical shifts into the electrostatic potential, which effectively shift the position of the bulk VBM [μ_e in Eq. (1)] and hence lead to errors in the formation energy of charge defects. The average potential of the bulk supercell and the supercell containing the charged defect should be different, so NAP results in the electrostatic potential of the bulk crystal not being recovered far from the defect. An example of this behavior is shown in Fig. 3.

Generally speaking, when using PBC DFT, total energies of neutral systems can be directly compared between different DFT calculations, but the eigenvalues can always be shifted by an unknown amount, with this unknown amount being introduced by NAP. When the system is charged, both the eigenvalues and the total energy can be affected by NAP.

If the shift in the electrostatic potentials introduced by NAP can be identified, its non-physical contribution to the formation energy can be removed. In order to determine the shift, a potential difference between a supercell containing the charged defect and a supercell of the same size containing the bulk crystal can be calculated

$$\Delta V_{q/b}(\mathbf{r}) = V_q(\mathbf{r}) - V_b(\mathbf{r}), \qquad (13)$$

where the electrostatic potential of the supercell containing the charged defect, V_q, is compared against the potential of a bulk supercell of the same shape, V_b. As the average of both the electrostatic potentials V_q and V_b is zero, the average of $\Delta V_{q/b}$ over the entire supercell will also be zero. To determine a scalar potential shift, $\Delta V_{q/b}$, between the two configurations, it must be averaged over some region of the supercell that is chosen to reproduce bulk properties,[22]
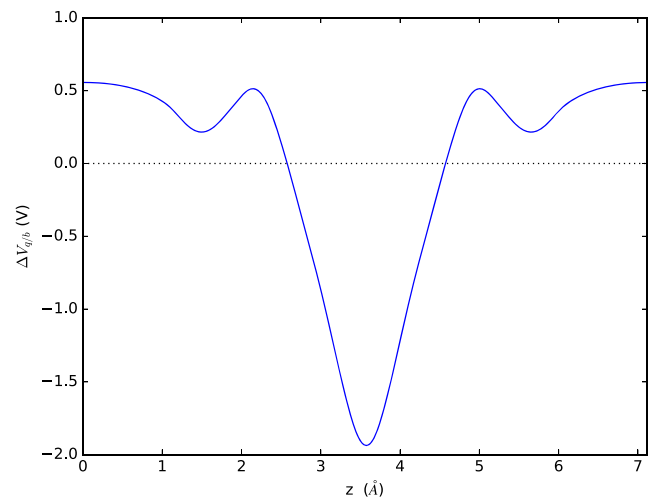

FIG. 3. The difference between the xy-averaged electrostatic potential of a V_C^{-2} defect in a $2 \times 2 \times 2$ diamond supercell (with full ionic relaxation) and an equivalent bulk cell of the same size. The defect is placed at the center of the cubic supercell. A potential difference of zero (denoted by the dotted line) is not recovered far from the defect. For this example, $\Delta V_{q/b} = 0.56$ V. The bulk potential is most strongly modified near to the defect at the center of the supercell.

$$\Delta V_{q/b} = \Delta V_{q/b}(\mathbf{r})\big|_{\text{far}}. \tag{14}$$

If the "far region" where $\Delta V_{q/b}$ is averaged over does reproduce bulk properties in the defective cell, then the difference between the potentials must be the unphysical potential shift introduced into the calculation by NAP. Hence,

$$\Delta V_{\text{NAP}} \approx \Delta V_{q/b}. \tag{15}$$

Equation (14) inevitably introduces a lot of choices in the way the far region that is averaged over is defined, potentially leading to different values of the scalar alignment shift $\Delta V_{q/b}$. In this paper, we limit ourselves to one of the simplest averaging techniques, calculating the xy planar average of $\Delta V_{q/b}(\mathbf{r})$ and then taking the z-value furthest from the defect. In this case, the far region is the xy plane furthest from the defect. Note that Eq. (15) is approximate, as the scalar potential shift introduced by NAP will only be correctly identified if the "far" region where the average is performed over does in fact reproduce bulk properties.

Possibly, some of the total potential alignment shifts present in Fig. 3 are caused by the same electrostatic interactions with image defects that were discussed in Secs. II B and III. Hence, we should be careful that the potential changes included in the calculation of E_{IIC} are not double counted in the potential alignment correction $q\Delta V_{\text{NAP}}$. Avoiding this double counting is the main motivation of the FNV method.[8] We will demonstrate that the IIC we introduced above can be used to predict part of the total potential alignment and that significant contributions to the potential alignment remain that are caused by effects entirely separate to IICs.

A. Decomposition of potential alignment

In the standard approach, ΔV_{NAP} is calculated by comparing the electrostatic potentials of a supercell of the bulk material and a supercell containing the charged defect, via Eqs. (13)–(15). However, when the problem is approached in this way, it remains unclear how NAP is introducing a potential shift. In order to further explore the causes of potential alignment, we split the process of forming a charged defect into steps and compare the potentials of four DFT calculations. Such a procedure is illustrated in Fig. 4. As the start and end DFT calculations are identical to the standard approach, splitting up the alignment in this way produces a total potential alignment $\Delta V(\mathbf{r})$ which is identical. This leads to the following expression for the total alignment:

$$\Delta V_{q/b}(\mathbf{r}) = \Delta V_{\text{AB}}(\mathbf{r}) + \Delta V_{\text{BC}}(\mathbf{r}) + \Delta V_{\text{CD}}(\mathbf{r}), \tag{16}$$

where the total change in potential of the bulk and defective systems $\Delta V_{q/b}$ is composed of the three alignment terms between the four configurations introduced via Fig. 4. Evaluating the potential in this way does require two additional single-point DFT calculations that would not otherwise be carried out (configurations B and C), whereas traditional approaches use only configurations that have already been calculated to evaluate the formation energy. We note that the charge model we defined in Eq. (12) is the charge density difference between configurations C and D. We show an example of this decomposition of the total potential difference in Fig. 5, for a V_C^{-2} defect in diamond.
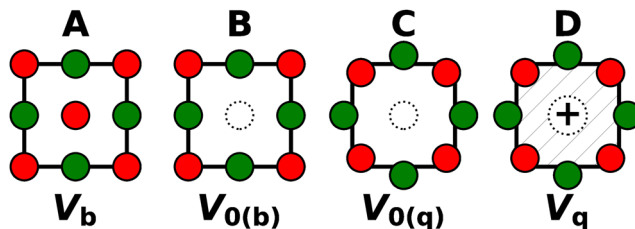

FIG. 4. Schematic process of forming a charged defect in a periodic DFT calculation. Each lettered square denotes a DFT single-point calculation. In A, the host material is simulated without a defect. In B, atoms are removed to form a vacancy (or added to form an interstitial), but the ionic structure is not yet relaxed. In C, the final ionic structure of the relaxed charged defect is adopted, but the charge is kept neutral. Finally, in D, electrons are exchanged to reach the required charge state. The ionic geometry in C is maintained, which is the relaxed geometry of this system. Configurations A and D are real states of the system, but B and C are artificial configurations used to break the defect formation process into steps. Only the final configuration D has a formal net charge compensated by a jellium background. The notation given below the figure is chosen to be similar to the conventional notation, but we prefer the alphabetical labels in this paper, to emphasize that we are not only using the conventional reference configurations.

B. The ΔV_{AB} component

Interestingly, in our example in Fig. 5, we observe that a significant potential shift is introduced far from the defect by the removal of a carbon atom from the supercell alone, producing the ΔV_{AB} component. This particular alignment component is shown in Fig. 6, for the diamond supercell.

The ΔV_{AB} component is extremely local to the defect site, so bulk properties are quickly recovered far from the defect. This is shown by how quickly ΔV_{AB} takes a constant value far from the defect. This ensures that the average in Eq. (14) can be carried out easily and without ambiguity—any sensible averaging technique will return the same value. As a result, for this component of potential alignment, we expect the approximate Eq. (15) to be very accurate.

ΔV_{AB} can be readily approximated by simply placing an atom of the type being exchanged in a periodic box of the same size, as justified in Fig. 7. This approximation is accurate to $\sim$30% for the examples we considered, as shown in Table I. Each atom present in the periodic DFT supercell makes a contribution to the average potential (which is conventionally neglected and set to zero), and hence changing the number of atoms in the supercell modifies the potential shift introduced by NAP.

The electrostatic potentials depend on the exact distribution of electric charge within the atom. As a consequence, any component of a DFT calculation that modifies the electron density can also modify the potential shift associated with each atom. The atomic species is clearly the most important, but the chemical environment in which the atom is present in also makes an important contribution to the distribution of its electric charge.

We will further develop these ideas in Sec. IV F, in order to predict how ΔV_{AB} scales with the size of the supercell.

C. The ΔV_{BC} component

The ΔV_{BC} component is the most challenging to interpret. This potential change is introduced when the final ionic

TABLE I. Comparison between alignment constants calculated between configurations A and B, with the potential alignment introduced by an isolated atom of the type being removed. The percentage errors reveal that there is a strong relationship between the alignment constants, but that this is a significant variance. The V_{atom} terms scale ideally with inverse volume. The ΔV_{AB} potential alignment terms retain an approximate inverse volume scaling, but the introduced noise prevents highly accurate extrapolation via Eq. (2).

System	Supercell	ΔV_{AB} (V)	V_{atom} (V)	Error (%)
Diamond (C vacancy)	$1 \times 1 \times 1$	−1.87	−2.35	−25.4
	$2 \times 2 \times 2$	−0.32	−0.30	8.4
	$3 \times 3 \times 3$	−0.10	−0.09	15.7
	$4 \times 4 \times 4$	−0.05	−0.04	18.4
MgO (O vacancy)	$1 \times 1 \times 1$	−0.98	−1.24	−26.7
	$2 \times 2 \times 2$	−0.12	−0.15	−32.7
	$3 \times 3 \times 3$	−0.03	0.05	−31.5
	$4 \times 4 \times 4$	−0.02	−0.02	−20.5
	$5 \times 5 \times 5$	−0.01	−0.01	−15.2
SnO$_2$ (Sn vacancy)	$1 \times 1 \times 1$	−3.73	−3.44	7.9
	$2 \times 2 \times 2$	−0.66	−0.59	10.0
	$3 \times 3 \times 3$	−0.23	−0.16	28.6
	$4 \times 4 \times 4$	−0.10	−0.07	32.7

configuration of the charged defect is introduced. Hence, this component is zero by definition when the formation energy of an unrelaxed defect is considered. An example of all three components along with the total change in potential is shown in Fig. 5. For the example of diamond, this last alignment component makes the smallest contribution to the total potential shift far from the defect. In this example, the relaxation of ions makes almost no change to the electrostatic potential far from the defect. In this case, it appears that NAP does not introduce a meaningful potential shift between configurations B and C.

In some other materials, such as MgO, $\Delta V_{BC}(\mathbf{r})$ is large, far from the defect. However, it never tends to a constant value, making it difficult to select a far region for Eq. (14). In turn, this will damage the approximation made in Eq. (15).
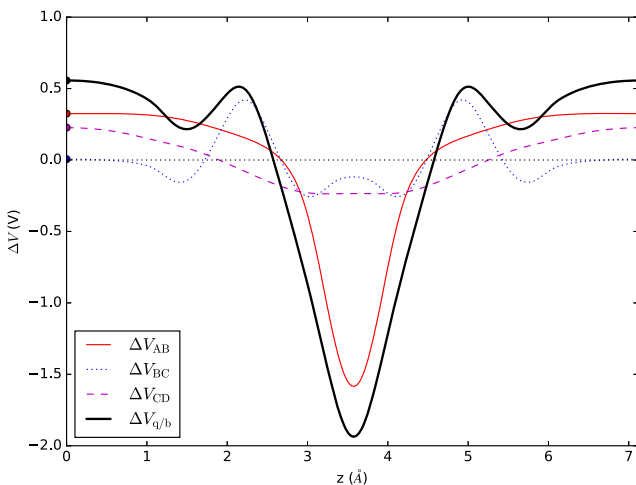
FIG. 5. Comparison of all the potential delta components of a V_C^{-2} defect in a $2 \times 2 \times 2$ diamond supercell. The ΔV_{AB} component is explored in Fig. 6, and the ΔV_{CD} component is explored more detail in Fig. 8. For this example, $\Delta V_{AB} = 0.32$ V, $\Delta V_{BC} = 0.01$ V, and $\Delta V_{CD} = 0.23$ V.
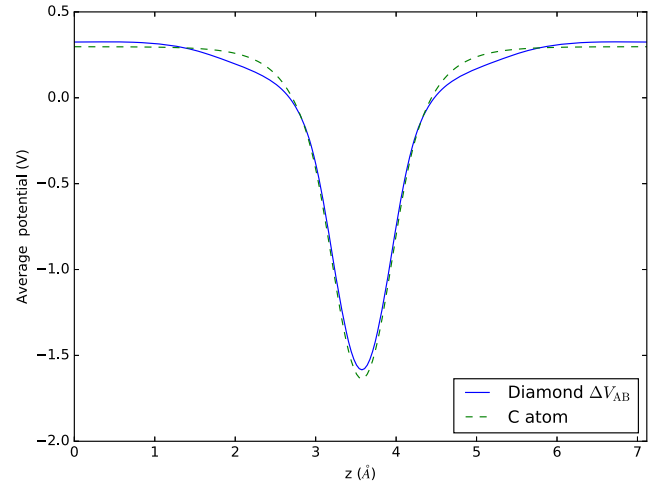
FIG. 6. Comparison between the ΔV_{AB} alignment components of a $2 \times 2 \times 2$ supercell of diamond when a carbon atom is removed with the potential of a carbon atom placed in an empty periodic cell of the same size. In this example, $\Delta V_{AB}(z=0) = 0.325$ V and $V_{C\,atom}(z=0) = 0.297$ V.

When the contribution of ΔV_{BC} to ΔV_{NAP} is entirely neglected, our corrections appear to perform as reliably for both examples where ΔV_{BC} is large and small. This suggests that this component does not in fact introduce a significant NAP shift. We can justify this observation using the following argument—as the number of electrons and ions in the cell is kept fixed, as well as their total density, there is not a large change in the neglected average potential. We note that the ΔV_{BC} is large in materials where the original formulation of the FNV method is believed to perform poorly and is the alignment component most strongly altered by the use of the KO far averaging method.[9] It would be interesting to explore in detail how the KO sampling method affects the different alignment components we define.

D. The ΔV_{CD} component

Finally, we consider the potential difference ΔV_{CD} between the charged defect and a neutral reference structure (as defined in Fig. 4). We observe the following relationship between our corrective potential, as introduced in Eq. (10), and the ΔV_{CD} potential alignment component:

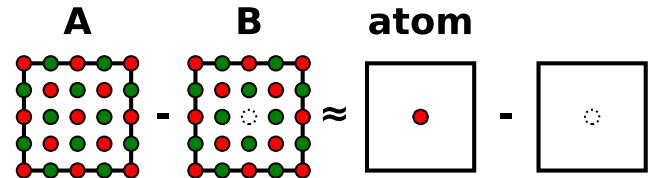$$\Delta V_{CD}(\mathbf{r}) \approx 2\varepsilon v_{PBC}(\mathbf{r}). \qquad (17)$$

FIG. 7. Schematic diagram of the approximation of the ΔV_{AB} potential alignment component, as in Fig. 6. Our results show that this is approximated by a calculation on an isolated atom. This is justified as shown in the diagram because the crystal appears to play only a secondary role. It changes the electrostatic distribution of charge associated with the atom, but this is a modification rather than a total change in character. The case of an isolated atom is simple enough to attack analytically and leads to the result that we expect, the potential shift introduced to be volume dependent, as in Eq. (23).

The electrostatic potential difference present in the DFT calculation is experienced by all the electrons in the calculation and therefore is not screened, whereas the interaction model v_{PBC} is experienced only by the additional defect charge and hence is screened by the dielectric constant ε (which is ε_0 when the ionic structure is relaxed and ε_∞ when it is held fixed). Although this factor of ε is justified by this electrostatic argument, it is not included in the derivation of the FNV method.[8] We note that the factor of two in this expression is due to double counting of the electron Hartree potential in CASTEP and may not be required in other codes depending on the electrostatic convention used (this factor is not required in CP2K). Finally, only the periodic part of the interaction model, v_{PBC}, contributes. The periodic DFT calculation does not contain the isolated component of the image interaction, v_{iso}—as a result, ΔV_{CD} does not contain enough information to infer E_{IIC}.

An example of this behavior is shown in Fig. 8. Across the defects we considered in this paper in isotropic dielectrics, we find Eq. (17) is satisfied to within 1% (for a material with an anisotropic dielectric, ε should be replaced with a tensor, but we did not explore how this could be done in this paper). This correspondence demonstrates that the potential alignment component ΔV_{CD} is already fully incorporated in our interaction correction and need not be considered separately. Note that this is the only part of potential alignment that is influenced by the jellium background present in charged calculations.

E. Discussion of all components

We show an example of the finite-size scaling resulting when the various components of potential alignment are used to perform alignment corrections, in Fig. 9. We switch to the example of an oxygen vacancy in MgO, as ΔV_{BC} is significant for this system. By contrast, the ΔV_{AB} component is relatively small in this case. These results demonstrate that alignment with ΔV_{AB} makes an improvement to the size scaling, but that the ΔV_{BC} and ΔV_{CD} components appear to do more harm than
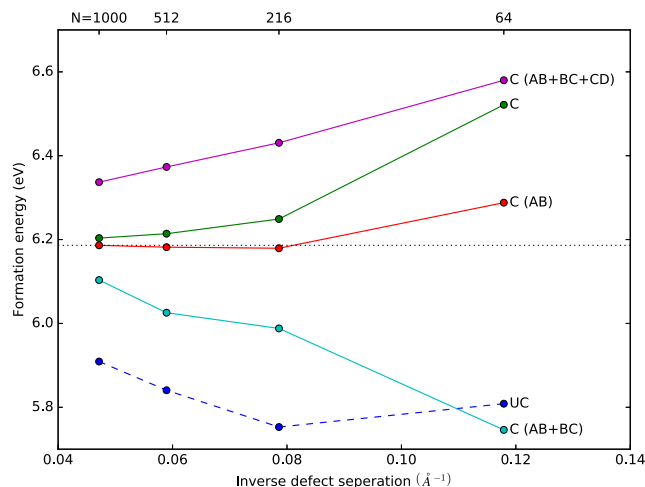
FIG. 9. Finite-size convergence of the formation energy of the V_O^{+2} defect in MgO, depending on the potential alignment utilized. Results are shown without any correction (UC) and with the Poisson IIC proposed above (C). Additionally, results also including a potential alignment correction are shown. Using only the ΔV_{AB} component yields the best finite-size scaling. Including the ΔV_{BC} and ΔV_{CD} components appears to make the scaling worse. A dotted line is shown for the result of the Poisson correction and ΔV_{AB} potential alignment for the largest supercell, to demonstrate how little size dependence remains in these results.

good. As we have argued, inclusion of ΔV_{CD} is expected to be harmful as this component is already fully included in the IIC. The negative contribution of the ΔV_{BC} contribution requires some further discussion.

The accurate determination of ΔV_{BC} is challenging for some materials, where we cannot evaluate Eq. (14) reliably. This equation requires the selection of a region of the supercell that reproduces bulk properties, but if relaxation of the defect ionic structure leads to long-range effects, we cannot identify such a region. For the example of diamond, we recover a bulk-like region (i.e., ΔV_{BC} tends to a constant far from the defect) and find that $\Delta V_{BC} \sim 0$. In the case of MgO, we cannot identify such a region, and the planar average we apply is certainly not sophisticated enough to accurately determine ΔV_{BC}. However, the results in Fig. 9 very strongly suggest that $\Delta V_{BC} \sim 0$. If this is the case, it is best not to include the ΔV_{BC} component in the potential alignment correction.

Taken together, all these results suggest a tempting premise. The ΔV_{AB} component is conveniently both the most significant source of the potential alignment error (once an image correction has been consistently applied) and the most well-behaved component. Almost any far averaging procedure used in Eq. (14) will unambiguously determine this component because the modification to the electrostatic potential is local to the defect site and does not lead to long-range changes in the material.

F. Finite-size scaling of the ΔV_{AB} component

Returning to the ΔV_{AB} component, simple model calculations can demonstrate that the exchange of atoms in periodic DFT leads to a potential shift. Consider an isolated atom in a large cubic box. If the atom is neutral, we expect that the same energy will be calculated under isolated and periodic boundary
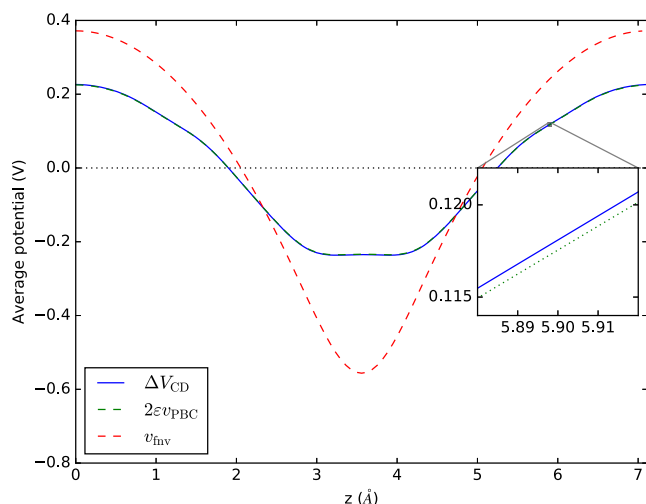

FIG. 8. Comparison of the ΔV_{CD} alignment component of a V_C^{-2} defect in a $2 \times 2 \times 2$ supercell, with the correction potentials constructed. The inset shows a $\times 50$ magnification.
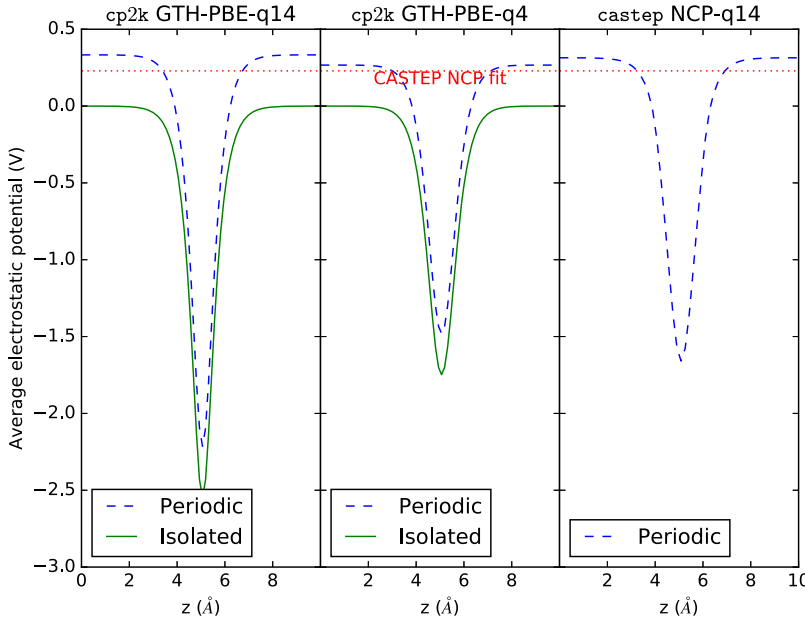
FIG. 10. Electrostatic potential of an Sn atom at the center of a 10 Å cubic box, using different pseudopotentials and DFT codes. The potential under isolated boundary conditions is shown in green solid line (only available for cp2k). The blue dashed line shows the potential when periodic boundary conditions are applied. The red dotted line shows the potential alignment predicted from the atomic radius, by the fit in Fig. 11. For each individual pseudopotential, the electrostatic potentials are identical in shape for both boundary conditions, but in the periodic system, the potentials are shifted upwards by ΔV_{atom}. The convention of setting $\langle V \rangle = 0$ introduces this shift ΔV_{atom} far from the atom, as the isolated atom has a non-zero average potential. Equivalently, calculations on atoms using isolated boundary conditions are always correctly referenced to the vacuum level, but this level is shifted within the PBC calculations. From left to right, V_{atom} = 0.33, 0.27, and 0.31 V. The castep on-the-fly pseudopotential definition for NCP-q14 is 3|2.4|12|14|16|50N:51N:42N(qc = 8).

conditions—as well separated neutral atoms do not interact. In fact, this is the behavior we observe. Using cp2k,[29–32] the same DFT calculation can be performed with both boundary conditions.[33,34] For a calculation on an Sn atom in a 10 Å cubic box, the total energy is only modified by 0.5 meV. However, as shown in the first panel of Fig. 10, the electrostatic potential contains a significant shift of 0.33 V. This shift is observed to be sensitive to the details of the DFT calculation, depending on the atom type, pseudopotential, and xc functional. This sensitivity to the pseudopotential used is also observed in ΔV_{AB}, as shown in Table II for the example of diamond. This shift is introduced by NAP.

This same argument that the change in boundary conditions between the two calculations introduces an alignment shift can also be made mathematically. First, consider the constraints that the boundary conditions place on the electrostatic potential of a neutral atom,

$$\left\langle V_{atom}^{PBC}(\mathbf{r}) \right\rangle = 0, \qquad (18)$$

$$V_{atom}^{iso}(\mathbf{r} \to \infty) = 0 \implies \left\langle V_{atom}^{iso}(\mathbf{r}) \right\rangle = C, \qquad (19)$$

where $\left\langle V_{atom}^{PBC}(\mathbf{r}) \right\rangle$ is the average electrostatic potential of the isolated atom under periodic boundary conditions. For the same system under isolated boundary conditions, $V_{atom}^{iso}(\mathbf{r})$, we do not require an average potential of zero and instead have the average value C.

TABLE II. Dependence of the ΔV_{AB} neutral alignment component on PP, for the example of a $2 \times 2 \times 2$ diamond supercell containing a carbon vacancy.

PP definition	ΔV_{AB} (V)
c-opt[a]	0.249
USP[a]	0.324
NCP[b]	0.362
QC5[c]	0.327

[a]This PP is distributed with castep.
[b]castep PP definition: 1|1.2|17|20|23|20N:21L(qc=8).
[c]castep PP definition: 2|1.4|8|9|10|20:21(qc=5).

However, in both cases, we are describing the same physical situation and should still calculate the same energy. This requires that the only difference between the two potentials is a uniform shift,

$$V_{atom}^{PBC}(\mathbf{r}) + \left\langle V_{atom}^{iso} \right\rangle = V_{atom}^{iso}(\mathbf{r}). \qquad (20)$$

A uniform shift in the electrostatic potential of a charge neutral system does not change its energy. Following this reasoning, we would like to understand how this potential shift introduced by PBC scales with the size of our simulation supercell. In the case of isolated boundary conditions, the vacuum level remains well defined, so this shift must be a non-physical change in the PBC potential. We can calculate this shift as

$$\left\langle V_{atom}^{iso}(\mathbf{r}) \right\rangle = \frac{1}{\Omega} \int_{\Omega} V_{atom}^{iso}(\mathbf{r}) d\mathbf{r}, \qquad (21)$$

where the average is performed by taking an integral over the simulation cell of volume Ω. Note, however, that we can infer a lot about $V_{atom}^{iso}(\mathbf{r})$ using our knowledge of isolated atoms—it must be some kind of local well that binds electrons. As a local potential, it is only non-zero near the atom being considered. As long as our isolated box is large enough for the potential to properly decay to zero at the edges, then the actual size of the box does not affect the shape of this potential at all. This is equivalent to stating that

$$\int_{\Omega} V_{atom}^{iso}(\mathbf{r}) d\mathbf{r} = \alpha_{atom}, \qquad (22)$$

where the integral given in Eq. (21) evaluates to a constant α_{atom}, which is a property of the atom independent of the size of the simulation cell. Different choices of atoms, pseudopotentials and xc functionals will all have their own consistent value of α_{atom} across any large simulation supercell considered. By combining Eqs. (20)–(22), we discover the following size dependence for the non-physical potential shift:

$$\Delta V_{atom} = \left\langle V_{atom}^{iso}(\mathbf{r}) \right\rangle = \frac{\alpha_{atom}}{\Omega}. \qquad (23)$$
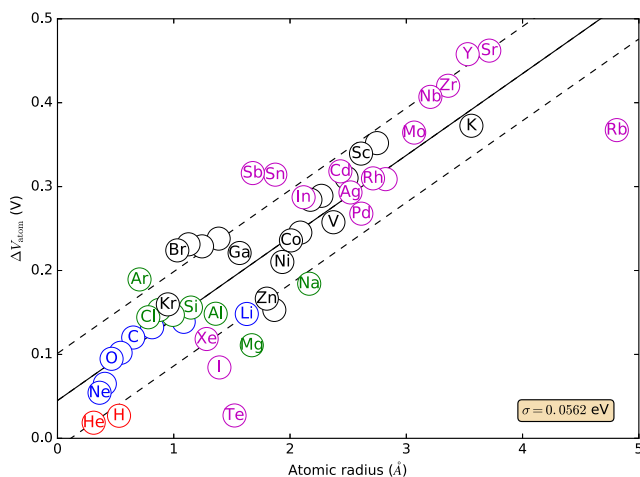
FIG. 11. Comparison between the potential alignment shifts introduced by an isolated atom in a 10 Å cubic box with the atomic radius of that species. Elements of the same row share the same color. We find that the potential alignment constants are dependent on the pseudopotential and XC functional used, but that all choices produced similar linear fits to that shown by the solid black line ($m = 0.097$ V/Å, $c = 0.045$ V). One standard deviation is shown by the dashed black line. Atomic radius is a characteristic of the element rather than the particular DFT approximations used.

Hence, this potential shift scales as a constant over the volume Ω. When we recall Eq. (1), this leads to an alignment correction of $q\Delta V$ and in fact would then explain the a_3 term in Eq. (2). Clearly this example is simpler than the actual situation of removing an atom from a lattice and leads to an ideal volume-dependent contribution to the total energy. When the defect is present within a material, this volume scaling is only approximately maintained.

As this potential shift is associated with isolated atoms, it is worth exploring if any atomic properties can be used to predict it. We find that atomic radius is the best predictor, as illustrated in Fig. 11. We present other descriptors we considered in Sec. S2 of the supplementary material. (We use atomic radii calculated by Ghosh and Biswas,[35] using Slater orbitals.) This suggests that this kind of potential alignment is likely to be more significant in calculations involving heavier elements. An important consequence of this argument is that we expect this component of potential alignment to swap signs when considering interstitials rather than vacancies. This behavior can be observed in the results of Shim *et al.*[13] and Taylor and Bruneval,[10] although these authors ascribe this behavior to different mechanisms (change in screening behavior and defect state hybridization, respectively).

V. APPLICATIONS TO DEFECTS IN DIFFERENT MATERIALS

A. Computational details

In order to test the introduced IIC, we studied three example vacancy defects using DFT. All these simulations were performed within the plane-wave *ab inito* density functional theory code CASTEP.[36] We used a modified version of release 16.1.1 that correctly calculated the non-coulombic pseudopotential energy of systems with net charge.[20] As a result of this error, the total energy of charged systems in CASTEP

was calculated incorrectly in versions before 16.1.2. All ions were represented with norm-conserving pseudopotentials. We used the following Castep on-the-fly pseudopotential definitions: For diamond, C 2|1.4|10|12|13|20N:21N(qc = 7). For MgO, O 1|1.2|23|26|31|20N:21L(qc = 9), Mg 1|1.8|3|4|4|30N: 31L:32N. For SnO$_2$, O 2|1.1|14|16|19|20N:21N(tm,nonlcc), Sn 3|2.2|6|8|8|50N:51N:42N(tm,nonlcc). Geometry optimizations were performed with the BFGS algorithm.[37]

The dielectric constants used in this paper were calculated using CASTEP's implementation of the density functional perturbation theory (DFPT).[38]

For diamond, we used high quality cutoffs, in light of difficulties encountered by previous studies of carbon vacancies in this material. In particular, previous work has demonstrated the importance of **k**-point sampling in calculating accurate formation energies in this material.[39] Brillouin zone sampling was achieved using the Monkhorst-Pack (MP) scheme.[40] The **k**-point grids used were chosen to maintain a constant sampling of $\Delta k = 0.0234$ Å^{-1} across the sizes of the supercell used. We also used a planewave cutoff of 2000 eV. This was equivalent to a $24 \times 24 \times 24$ sampling in the conventional unit cell of 8 atoms. Additionally, the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional was used,[41] in order to compare directly with previous studies.[13,39]

Using the above cutoffs, the structural properties of diamond were well reproduced. The lattice constant was calculated as 3.56 Å, 0.3% smaller than the experimental value of 3.57 Å. As is typical, the bandstructure was not as well replicated, with a bandgap of 4.26 eV. This is a 22% underestimation of the experimental value of 5.49 eV.

For MgO, a more coarse parameterization was used, as this ionic material is computationally cheaper to accurately simulate. The PBEsol exchange correlation functional was used,[42] as it led to improved agreement with MgO's experimental properties. The planewave cutoff was 843.55 eV, and a $7 \times 7 \times 7$ **k**-point grid was used to sample the unit cell.

The experimental lattice constant of MgO is 4.21 Å, compared to the value of 4.24 Å calculated here.[43] This represents a 0.6% overestimate. This leads to a bandgap of 4.5 eV, a 43% underestimation compared to the experiential gap of 7.7 eV.[44]

For SnO$_2$, a planewave basis set with a cutoff energy of 2100 eV was used, with a $4 \times 4 \times 8$ **k**-point sampling for the unit cell. We used the PBE functional[41] to directly compare with a previous study. We calculated lattice constants of $a = 4.78$ Å, $b = 3.20$ Å compared to the experimental values of $a = 4.74$ Å, $b = 3.19$ Å.[45] These were 0.98% and 0.39% overestimates, respectively.

The bandgap was calculated as 2.04 eV, which was a 43% underestimate of the experimental value of 3.59 eV.[46]

We note that the semi-local functionals used in this study do not produce accurate formation energies,[47] but do allow the use of larger supercells, for which size dependence can be seen more clearly. The use of hybrid functionals enables an improved description of defect level positions relative to the bulk bandstructure, but does not strongly affect the electronic charge densities which form the focus of this paper.

B. Supercell size convergence in diamond

In diamond, the dielectric screening is isotropic and principally due to electronic contributions. In fact, ionic relaxation makes a negligible contribution to screening, and the static and high frequency dielectric constants calculated using DFPT are identical ($\varepsilon_0 = \varepsilon_\infty = 5.76$). We used a carbon chemical potential calculated from bulk diamond, as in the previous studies. Previously the formation energy of the V_C^0 defect was reported as 7.31 eV,[13] but in good agreement, we calculated it as 7.18 eV.

First, we looked at the unrelaxed carbon vacancy, as charge corrections have been studied in detail for this defect.[13,22] The V_C defect supports both positive and negative charge states. Shim *et al.*[39] considered several correction methods and reported that the MP method appeared to perform reliably for negatively charged carbon vacancies, but seemed to fail for positively charged vacancies. They suggested that this behavior was due to different screening responses for the two types of defect. Later work by Freysoldt *et al.*[22] attempted to capture this physics using the FNV method, with exponentially decaying Gaussians as the model charge. This motivated the present study—changes in the screening response of the material should be naturally included in our method. We present the internal polarization ratios f we calculated using various image corrections in Sec. S1 of the supplementary material.

Using the procedure defined in Eq. (12), charge models were constructed for the investigated defects. The positive and negative charge states of the defect distribute the charge in very similar ways, as shown by the comparison in Fig. 12. This leads to similar interaction corrections for both states and suggests that the polarization response of diamond does not strongly depend on the sign of the charge.

This result does not explain the behavior observed by the previous Shim *et al.* study[13] and suggests that we need to find another reason to understand the difference between the correction of the positive and negative states they reported.
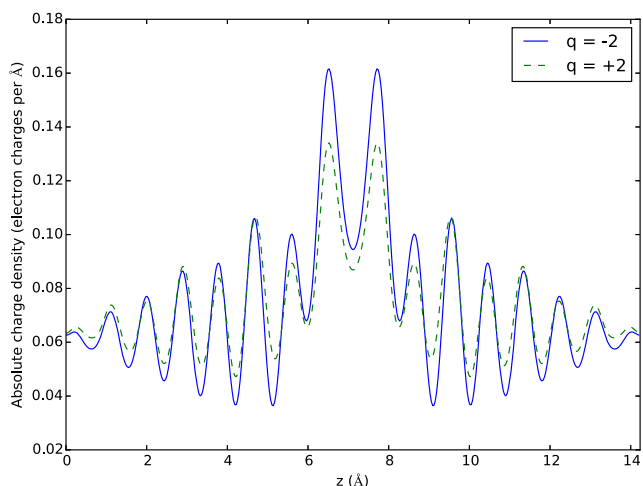
The supercell size convergence of two of these charge states is shown in Fig. 13. The $1 \times 1 \times 1$ cells are excluded, as these cells provide a poor description of the defect. We avoid these small cells in all the convergence plots shown in this paper.

The first general point to make is that applying a ΔV_{AB} alignment correction does make a significant contribution to the total corrections calculated. However, this alignment increases the formation energy of negative states and lowers the formation energy of positive states. This is in contrast to the IIC, which always increases the defect formation energy. This is a general result for cubic cells, where the jellium background decreases the formation energy more than the repulsive interaction with images which increases it. It can be seen that the unaligned corrections (dashed lines) show a stronger size dependence. Introducing alignment makes the solid lines flatter and makes the calculated formation energy
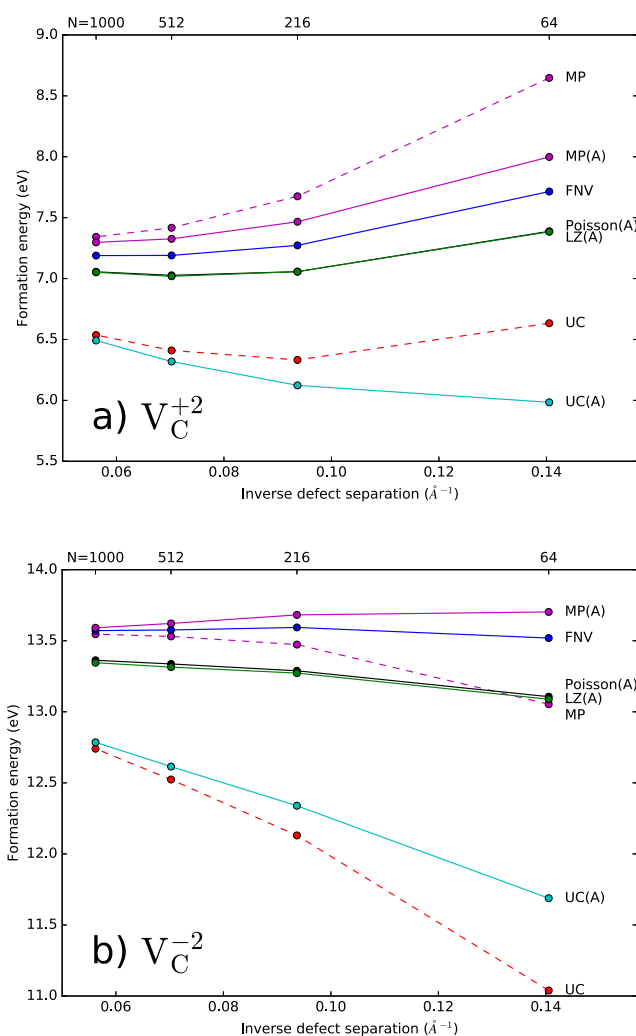

FIG. 13. Supercell size convergence of the different charge states of the unrelaxed carbon vacancy, when ionic relaxation is not allowed, with various IICs, both with ΔV_{AB} potential alignment [solid lines denoted by (A)] and without it (dashed lines). Number of atoms N in the bulk supercell shown on the top axis. Interaction uncorrected (UC), Makov-Payne corrected (MP), Lany-Zunger corrected (LZ), FNV corrected, and our Poisson correction are shown. Note that the FNV method includes its own potential alignment scheme, which uses $\Delta V_{q/b}$ potential alignment.

FIG. 12. Comparison of the planar charge density of the charge models $q(\mathbf{r})$, constructed for the +2 and −2 charge states of the unrelaxed carbon vacancy in a $4 \times 4 \times 4$ supercell.
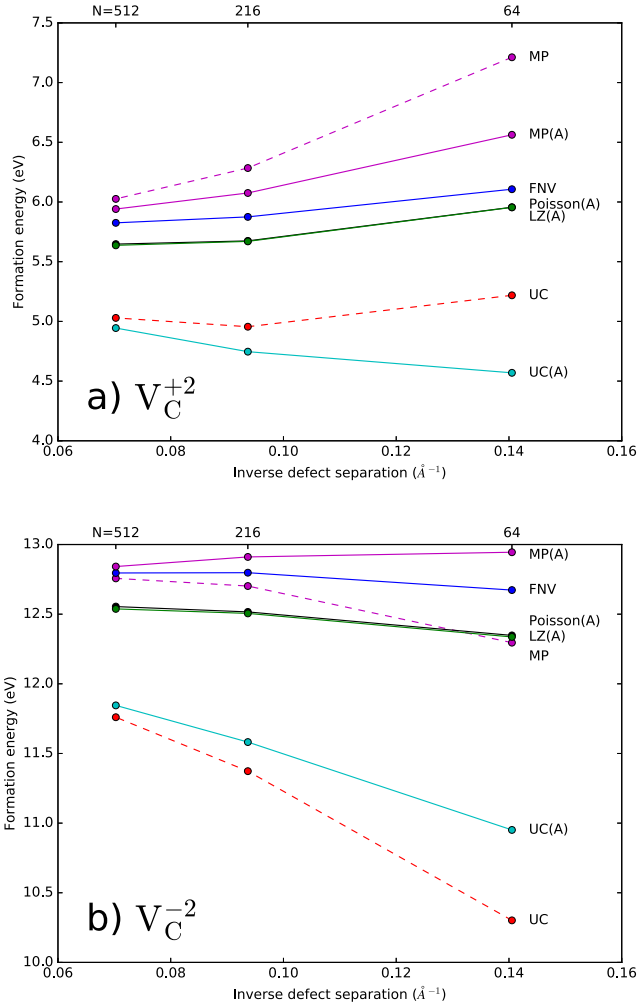
FIG. 14. The same notation as Fig. 13. Supercell size convergence of the different charge states of the relaxed carbon vacancy, where the ionic structure of the charged defect has been geometry optimized.

TABLE III. Formation energies for the different charge states of the carbon vacancy in diamond, referenced to the chemical potential of a carbon atom in diamond. All energies are in eV.

q=	0	+1	+2	−1	−2
Unrelaxed	7.97	6.87	7.06	8.80	13.36
Relaxed	7.18	5.86	5.65	8.19	12.57

is unrelated to the image interactions present in E_{IIC}. If the positive alignment correction is neglected for the negative vacancies, then an IIC that is too large appears more accurate. In the opposite way, weaker IICs appear to be more effective when the positive states do not include the negative V_{AB} alignment correction.

When ΔV_{AB} potential alignment is included in the PC corrections [yielding MP(A) and LZ(A), respectively], we can see more clearly that the MP IIC includes stronger interaction than the Poisson IIC, just as the LZ method contains a slightly weaker interaction.

We suggest that the above discussion helps understand the results reported by Shim *et al.*[13,39] The MP method appears to yield an accurate correction for the negative charge states because the MP method's overcorrection is compensated by the missing potential alignment term. For the positive states, this favorable error cancellation does not occur.

Finally, the FNV correction falls between the corrections of MP(A) and Poisson(A). The exact ratio is dependent on the charge state in question. There is freedom in the construction of the FNV correction, which allows other models to return different ratios. For these diamond calculations, we have used the default parameters. Diamond vacancy specific values were suggested in a previous publication,[22] but we find that these modifications do not affect the overall trends observed.

When we allow full atomic relaxation, the results are quite similar to those shown previously. As shown in Fig. 14, the FNV method is the most strongly affected—as discussed in the literature, and it is sensitive to the changes in potential alignment caused by atomic relaxation.[9,11] Atomic relaxation increases the volume in the cell that is perturbed by the defect and hence makes it harder to define a region far from the defect. As we use a ΔV_{AB} alignment correction based on the unrelaxed cells, this complication can be completely avoided in the other results we present. Additionally, the corrections calculated using the Poisson method change only very
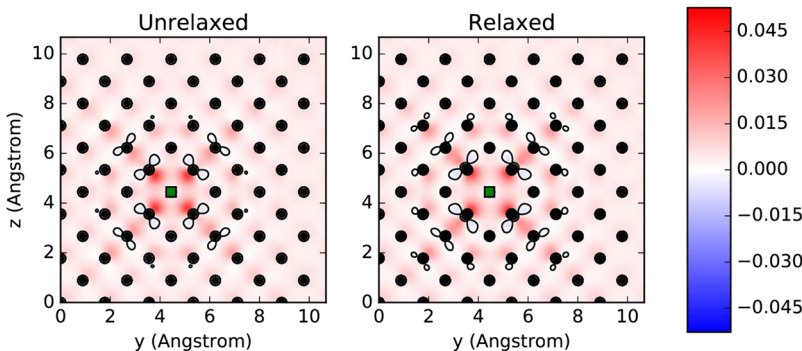
less dependent on the supercell size. But also, only including alignment without the IIC appears to make the formation energy of positive states worse. This changes the way the IIC methods compare for smaller cells (≤ 216 ions), when the ΔV_{AB} potential alignment is neglected. Note that the finite-size scaling of results with only the MP correction appears to be improved for negative defects over positive defects.

This shows how potential alignment can directly cloud comparison of IICs. The potential alignment component ΔV_{AB}

FIG. 15. Diamond vacancy defect charge models [defined via Eq. (12)], for the $3 \times 3 \times 3$ supercell containing the V_C^{+2} defect. The $q = 0$ isosurface is shown. The position of the vacancy is denoted by a green square and carbon ions are denoted by black circles. As a covalent material, most of the defect charge is present on bonds, and relaxation does not significantly affect the charge model.
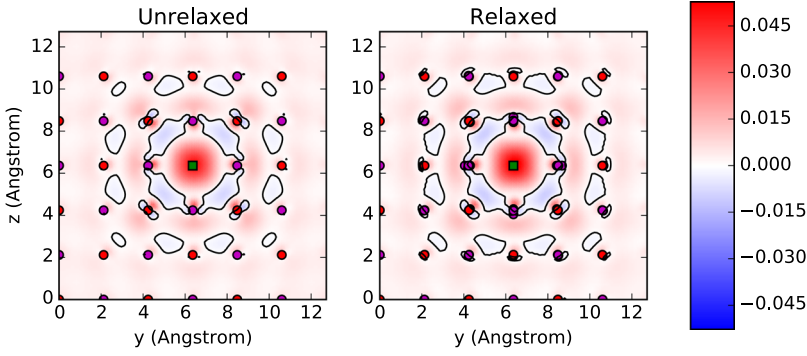
FIG. 16. MgO vacancy defect charge models [defined via Eq. (12)] for the $3 \times 3 \times 3$ supercell containing the V_O^{+2} defect. The $q = 0$ isosurface is shown. The position of the vacancy is denoted by a green square, Mg ions are denoted by magenta circles, and O ions are denoted by red circles. The ionic material polarizes and shows alternating regions of positive and negative charge. Relaxation of the ions makes a minor change to the calculated charge model.

slightly. As the static and high frequency dielectric constants are identical, the MP and LZ corrections do not change on relaxation.

Some examples of the charge models used in our correction are shown in Fig. 15, which shows the minor role of relaxation in this material.

Our final defect formation energies are summarized in Table III.

C. Supercell size convergence in MgO

In MgO, electronic and ionic contributions to the dielectric response are both important and hence allowing relaxation triples the dielectric constant ($\varepsilon_\infty = 3.19$, $\varepsilon_0 = 9.41$). For the exchange of oxygen atoms, we used a chemical potential set to an isolated O atom in its triplet ground state.

We looked at the oxygen vacancy defect in MgO, which supports a + and +2 charge state. The supercell size convergence of the relaxed defect is shown in Fig. 17. These results show very similar trends to those we presented for diamond.

TABLE IV. Summary of calculated formation energies for defects in MgO and SnO$_2$, referenced to the chemical potentials of isolated atoms. Formation energies are for the largest cells calculated, with Poisson IIC and ΔV potential alignment.

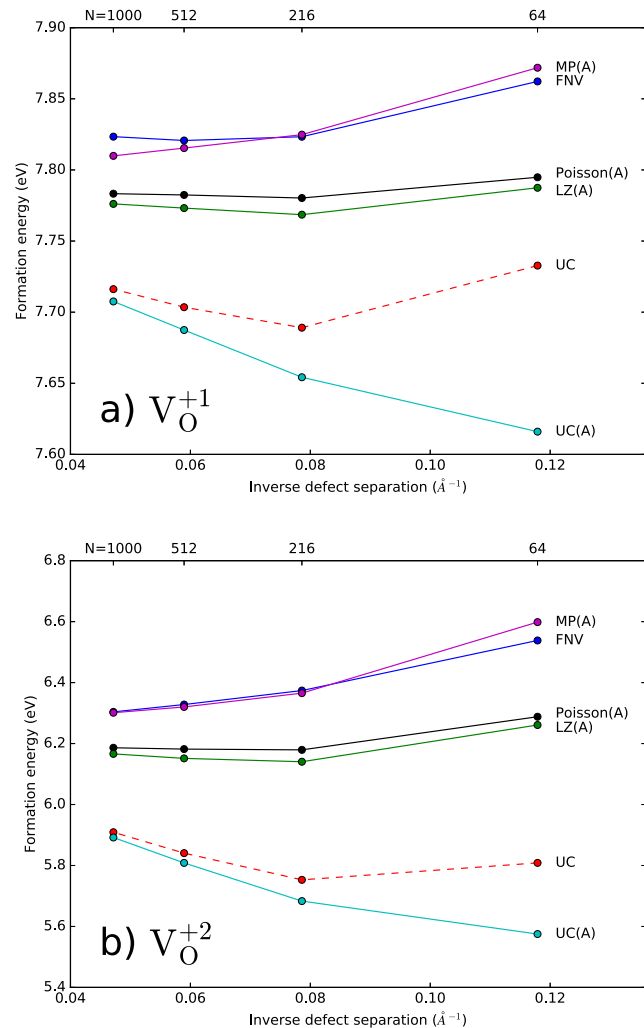| Host material	Defect	Formation energy (eV)
MgO	V_O^0	10.27
	V_O^+	7.78
	V_O^{2+}	6.19
SnO$_2$	V_{Sn}^0	15.56
	V_{Sn}^{-4}	18.75

FIG. 17. The same notation as Fig. 13. Supercell size convergence of the different charge states of the relaxed oxygen vacancy in MgO. Neutral potential alignment and several IICs are shown.
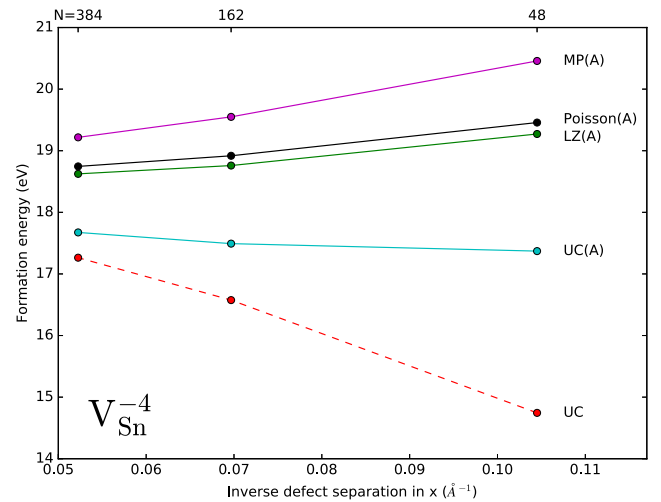

FIG. 18. The same notation as Fig. 13. Supercell size convergence of -4 charge state of the relaxed Sn vacancy in SnO$_2$. The symbol (A) denotes the ΔV_{AB} potential alignment. Both image charge uncorrected (UC) and several charge correction schemes are shown.
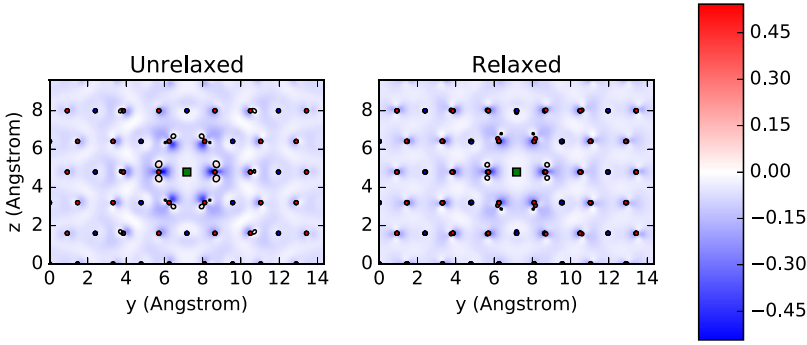
FIG. 19. SnO$_2$ vacancy defect charge models [defined via Eq. (12)], for the $3 \times 3 \times 3$ supercell containing the V_{Sn}^{-4} defect. The $q = 0$ isosurface is shown. The position of the vacancy is denoted by a green square, Sn ions are denoted by blue circles, and O ions are denoted by red circles. This material shows more complex polarization structure, which is not strongly affected by relaxation.

Previous work found that the formation energy of the V_O^0 defect with respect to the chemical potential of an isolated oxygen atom is 10.08 eV.[48] In good agreement, we calculated a formation energy of 10.27 eV, using this chemical potential (the majority of this energy difference is due to our use of the PBEsol functional—using PW91, we calculate a formation energy of 10.05 eV).

The calculated defect charge models are shown in Fig. 16. These plots show clear fringes of positive and negative polarization in our charge model, as expected from an ionic crystal.

Corrected formation energies for the relaxed MgO cells are shown in Fig. 17. The FNV corrections also become larger than the MP corrections when relaxation is allowed. The difficulty in applying the FNV method to materials that display large atomic relaxations has already been reported, as the method can become very sensitive to the region in which potential alignment is performed over.[9]

Our final defect formation energies for MgO are summarized in Table IV.

D. Supercell size convergence in SnO$_2$

In SnO$_2$, the ionic contribution to screening is stronger than the electronic contribution. Additionally, the unit cell is a rectangular cuboid and the dielectric constants are anisotropic. The calculated dielectric responses were

$$\varepsilon_\infty = \begin{pmatrix} 4.457 & 0 & 0 \\ 0 & 4.457 & 0 \\ 0 & 0 & 4.874 \end{pmatrix},$$

$$\varepsilon_0 = \begin{pmatrix} 13.112 & 0 & 0 \\ 0 & 13.112 & 0 \\ 0 & 0 & 10.718 \end{pmatrix},$$

where ε_∞ is almost isotropic and ε_0 is anisotropic. We used a chemical potential for Sn calculated from an isolated atom in its triplet ground state.

The defect formation energy of the relaxed V_{Sn}^0 defect has been reported as 16.0 eV.[49] This is higher than the 15.56 eV we calculated. Additionally, we find that a net magnetic moment of $4\mu_B$ on the neutral defect is energetically favorable, in agreement with Ref. 50 and disagreement with the value of $3.63\mu_B$ previously reported.[49] A more detailed study of the V_{Sn} defect will follow in a future publication.

The new method we have outlined can still be applied to materials with an anisotropic dielectric. Additionally, the MP method has already been extended to materials with anisotropic dielectrics.[21] However, the LZ method has not been extended. Instead, we apply f_{LZ} (calculated using the knowledge of the non-cubic cell, but not the anisotropic dielectric) to the anisotropic MP correction. We have not applied the FNV method, although Kumagai and Oba's[9] work generalizes the FNV method to anisotropic media.

The corrections calculated in Fig. 18 are similar to those for the other materials. The anisotropic dielectric further increases the screening factor f, calculated using our method, by approximately 0.008. Hence, the anisotropic dielectric only has a minor effect on defect formation energies in this material. Potentially, this could be because the anisotropy in the cell vectors and dielectric is balanced against each other. A stronger effect could be seen in materials where this is not the case.

Our final defect formation energies are summarized in Table IV. The calculated defect charge models are shown in Fig. 19.

VI. DISCUSSION AND CONCLUSIONS

In this paper, we have developed an IIC which directly calculates the electrostatic image interaction introduced by the additional charge associated with a charged defect. When applying this method to cubic bulk crystals, we found strong agreement with the Lany-Zunger[6,7] IIC. We also demonstrated the extension of the developed method to SnO$_2$, as an example of a material with an anisotropic dielectric constant, beyond the scope of the LZ method.

However, even with this new correction, we find that an appropriately constructed IIC still requires a potential alignment procedure. We introduced a decomposition of potential alignment into three components that can, in principle, contribute to the failure of the bulk potential to be recovered far from the defect. We find that the first component ΔV_{AB} is due to a change in the atom number. This can introduce significant energy errors when the atoms added (or removed) have a large atomic radius. Second, the ΔV_{BC} component is introduced by ionic relaxation. In cases where we can still find a good reference region that reproduces bulk properties inside the defect cell, this potential shift is near zero. For examples where the relaxation around the charged defect leads to atomic displacements across the entire supercell, it remains a challenge to evaluate this potential alignment term. However, we

present indirect evidence that it remains close to zero. Finally, we find that the ΔV_{CD} component also introduces a large alignment shift, but that this is already included in the image charge interaction E_{IIC}.

Although the arguments look rather different, the conclusions we have reached readily support what is normally termed the traditional potential alignment technique,[11] where alignment is calculated between the bulk material and a neutral defect. The atomic relaxation present in the traditional approach can lead to differences to the ΔV_{AB} component, but these modifications are very minor for many examples (see Sec. S3 of the supplementary material). This correspondence could be expected to fail when ionic relaxations around the neutral defect are long range.

We have presented a plausible argument that ΔV_{BC} alignment shifts are small, but we have not been able to prove this. If a defect can be identified where this alignment term can both be unambiguously determined and is significant, this would provide an example where the methods endorsed by this paper would not be reliable. In such an example, the FNV method could be anticipated to be more effective—as the inclusion of a ΔV_{BC} could become beneficial rather than harmful. The application of the potential alignment decomposition introduced in this paper to a wide range of defects and materials would help to further understand the role of the ΔV_{BC} component.

The above results lead us to the conclusion that the ΔV_{AB} component we define, caused by the removal (or addition) of atoms from the supercell alone, is the dominant remaining finite-size error once the E_{IIC} correction has been applied. As a consequence, we expect that potential alignment corrections are not required for calculations on self-trapped polarons. An advantage of this conclusion is that determination of the potential alignment correction is no longer hindered by relaxation of the defect geometry.

It is already known that IICs are positive for cubic cells, as the presence of a jellium background over stabilizes defects.[12] However, in this work, we show that the sign of the potential alignment correction introduced by the exchange of atoms (the dominant V_{AB} component) can also be predicted, as in Table V. Generally, the total finite-size corrections required are smaller for positive vacancies and negative interstitials, as the IICs and potential alignment corrections compete and partly cancel—in the opposite case, negative vacancies and positive interstitials require larger corrections as the two sources of finite size error add together and reinforce the size dependence. This behavior has already been noted in the literature,[10,13] but our results provide a strong justification for this observed behavior. (We note that the FNV correction does not always obey these rules for the sign of the potential alignment, as the ΔV_{AB} component is not always the dominant component of the FNV alignment scheme.)

TABLE V. Prediction of the sign of the alignment correction $q\Delta V$ caused by the exchange of atoms.

	Positive charge state	Negative charge state
Vacancy defect	−	+
Interstitial defect	+	−

All of the above conclusions combine to justify the origin of the empirical scaling rule suggested in Ref. 15, which predicts finite size dependence that scales inversely with supercell length (L^{-1}) and supercell volume (L^{-3}). The L^{-1} scaling is found to be introduced for cubic supercells by the defect electrostatics of a charge interacting with its images in a jellium background (the IIC). This behavior has already been predicted in Refs. 6 and 7, but we maintain the same result without using an approximate multipole expansion. We find that the second volume dependent L^{-3} term is caused by a potential alignment effect (the ΔV_{AB} component), due to the removal (or addition) of ions effectively modifying the average potential in a periodic DFT calculation. The ΔV_{AB} is found to be approximated by calculations on neutral atoms, which are simple enough for an analytic L^{-3} dependence to be derived. This term is predicted to be larger when atoms with larger atomic radii are exchanged to form a defect.

Taken together, these developments enable robust corrections across various materials and defect charge states. For cubic systems with isotropic dielectrics, the LZ IIC combined with a potential alignment between the bulk material (the A configuration) and with atoms added (or removed, the B configuration) can provide reliable corrections. Our conclusions on potential alignment are fully transferable to anisotropic dielectrics and non-cubic cells, where the V_{AB} component will always produce an inverse volume scaling finite size error. For the case of IICs, we present a method that can be applied to anisotropic dielectrics and non-cubic cells, where an inverse-length scaling finite-size dependence can no longer be expected.

Finally we note that the IIC introduced in this paper can be applied to more complex dielectric environments, such as systems containing multiple interfaces between materials, and such systems will be considered in a future paper. It would also be possible to use a similar solver to apply self-consistent corrections to DFT calculations, as in the DCC method.[24] The relationship discovered in Eq. (17) is an important ingredient of such a scheme.

SUPPLEMENTARY MATERIAL

See supplementary material for the following additional data: Internal polarization ratios f calculated using Eq. (6), the various correction methods and materials, and additional fits to the atomic descriptors we considered for the alignment constants of neutral atoms. We introduce the agreement between traditional alignment constants and the ΔV_{AB} alignment.

ACKNOWLEDGMENTS

[1]C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, Rev. Mod. Phys. **86**, 253 (2014).

[2]M. Leslie and N. Gillan, J. Phys. C: Solid State Phys. **18**, 973 (1985).

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

[4]D. B. Laks, C. G. Van de Walle, G. F. Neumark, P. E. Blöchl, and S. T. Pantelides, Phys. Rev. B **45**, 10965 (1992).

[5]K. Burke, J. Chem. Phys. **136**, 150901 (2012).

[6]S. Lany and A. Zunger, Phys. Rev. B **78**, 235104 (2008).

[7]S. Lany and A. Zunger, Modell. Simul. Mater. Sci. Eng. **17**, 084002 (2009).

[8]C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, Phys. Rev. Lett. **102**, 016402 (2009).

[9]Y. Kumagai and F. Oba, Phys. Rev. B **89**, 195205 (2014).

[10]S. E. Taylor and F. Bruneval, Phys. Rev. B **84**, 075155 (2011).

[11]H.-P. Komsa, T. T. Rantala, and A. Pasquarello, Phys. Rev. B **86**, 045112 (2012).

[12]N. D. M. Hine, K. Frensch, W. M. C. Foulkes, and M. W. Finnis, Phys. Rev. B **79**, 024112 (2009).

[13]J. Shim, E.-K. Lee, Y. J. Lee, and R. M. Nieminen, Phys. Rev. B **71**, 035206 (2005).

[14]Y.-N. Wu, X.-G. Zhang, and S. T. Pantelides, Phys. Rev. Lett. **119**, 105501 (2017).

[15]C. W. M. Castleton, A. Höglund, and S. Mirbt, Phys. Rev. B **73**, 035215 (2006).

[16]S. B. Zhang and J. E. Northrup, Phys. Rev. Lett. **67**, 2339 (1991).

[17]C. Varvenne, F. Bruneval, M.-C. Marinica, and E. Clouet, Phys. Rev. B **88**, 134102 (2013).

[18]J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C: Solid State Phys. **12**, 4409 (1979).

[19]L. Kantorovich, *Quantum Theory of the Solid State: An Introduction* (Springer Science & Business Media, 2004), Vol. 136.

[20]R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, 2004).

[21]S. T. Murphy and N. D. M. Hine, Phys. Rev. B **87**, 094111 (2013).

[22]C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, Phys. Status Solidi B **248**, 1067 (2011).

[23]H.-P. Komsa and A. Pasquarello, Phys. Rev. Lett. **110**, 095505 (2013).

[24]I. Dabo, B. Kozinsky, N. E. Singh-Miller, and N. Marzari, Phys. Rev. B **77**, 115139 (2008).

[25]L. Anton *et al.*, computer code DL_MG, The Numerical Algorithms Group Ltd., 2013.

[26]L. Anton, J. Dziedzic, C.-K. Skylaris, and M. Probert, Multigrid solver module for onetep, castep and other codes, 2013.

[27]R. D. Budiardja and C. Y. Cardall, Comput. Phys. Commun. **182**, 2265 (2011).

[28]U. Trottenberg, C. W. Oosterlee, and A. Schuller, *Multigrid* (Academic Press, 2000).

[29]J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, Comput. Phys. Commun. **167**, 103 (2005).

[30]J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **4**, 15 (2014).

[31]J. VandeVondele and J. Hutter, J. Chem. Phys. **127**, 114105 (2007).

[32]M. Krack, Theor. Chem. Acc. **114**, 145 (2005).

[33]L. Genovese, T. Deutsch, A. Neelov, S. Goedecker, and G. Beylkin, J. Chem. Phys. **125**, 074105 (2006).

[34]L. Genovese, T. Deutsch, and S. Goedecker, J. Chem. Phys. **127**, 054704 (2007).

[35]D. C. Ghosh and R. Biswas, Int. J. Mol. Sci. **3**, 87 (2002).

[36]S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. Payne, Z. Kristallogr.—Cryst. Mater. **220**, 567 (2005).

[37]B. G. Pfrommer, M. Cote, S. G. Louie, and M. L. Cohen, J. Comput. Phys. **131**, 233 (1997).

[38]K. Refson, P. R. Tulip, and S. J. Clark, Phys. Rev. B **73**, 155114 (2006).

[39]J. Shim, E.-K. Lee, Y. J. Lee, and R. M. Nieminen, Phys. Rev. B **71**, 245204 (2005).

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

[41]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[42]J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Phys. Rev. Lett. **100**, 136406 (2008).

[43]V. Tsirelson, A. Avilov, Y. A. Abramov, E. Belokoneva, R. Kitaneh, and D. Feil, Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. **54**, 8 (1998).

[44]T. Lindner, H. Sauer, W. Engel, and K. Kambe, Phys. Rev. B **33**, 22 (1986).

[45]W. H. Baur and A. A. Khan, Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. **27**, 2133 (1971).

[46]K. Reimann and M. Steube, Solid State Commun. **105**, 649 (1998).

[47]H.-P. Komsa, P. Broqvist, and A. Pasquarello, Phys. Rev. B **81**, 205118 (2010).

[48]J. Carrasco, N. Lopez, and F. Illas, Phys. Rev. Lett. **93**, 225502 (2004).

[49]A. Sarkar, D. Sanyal, P. Nath, M. Chakrabarti, S. Pal, S. Chattopadhyay, D. Jana, and K. Asokan, RSC Adv. **5**, 1148 (2015).

[50]G. Rahman, V. M. García-Suárez, and S. C. Hong, Phys. Rev. B **78**, 184404 (2008).

[51]J. D. Hunter, Comput. Sci. Eng. **9**, 90 (2007).

[52]J. D. Jackson, *Classical Electrodynamics* (Wiley, 1999).

[53]R. K. Wangsness, *Electromagnetic Fields* (Wiley, 1986).