Consistent treatment of charged systems within periodic boundary conditions: The projector augmented-wave and pseudopotential methods revisited

Fabien Bruneval and Jean-Paul Crocombette

CEA, DEN, Service de Recherches de Métallurgie Physique, F-91191 Gif-sur-Yvette, France

Xavier Gonze

European Theoretical Spectroscopy Facility, Institute of Condensed Matter and Nanosciences, Université catholique de Louvain, Chemin des étoiles 8, bte L07.03.01, B-1348 Louvain-la-neuve, Belgium

Boris Dorado, Marc Torrent, and François Jollet

CEA, DAM, DIF, F-91297 Arpajon, France

The *ab initio* calculation of charged defect properties in solids is not straightforward because of the delicate interplay between the long-range Coulomb interaction and the periodic boundary conditions. We derive the projector augmented-wave (PAW) energy and Hamiltonian with special care taken on the potentials from the Coulomb interaction. By explicitly treating the background compensation charge, we find additional terms in the total energy of the charged cells and in the potential. We show that these background terms are needed to accurately reproduce all-electron calculations of the formation energy of a charged defect. In particular, the previous PAW expressions were spuriously sensitive to the pseudization conditions and this artifact is removed by the background term. This PAW derivation also provides insights into the norm-conserving pseudopotential framework. We propose then an alternative definition for the total energy of charged cells and for the Kohn-Sham potential within this framework that better approximates the all-electron results.

PACS number(s): 71.15.Dx, 71.55.−i

I. INTRODUCTION

In order to calculate the properties of charged defects in semiconductors, of polarons in insulators, or of isolated ions, it is often required to consider charged systems in *ab initio* calculations. The combined use of charged simulation cells and of periodic boundary conditions leads to intricacies that require a lot of care [1–4]. First, a truly charged periodic system would have an infinite energy. This problem is circumvented by adding a compensating background charge to restore the global charge neutrality. Second, even with a compensating background, the electrostatic potential is still not uniquely defined. Indeed, the electrostatic potential induced by a lattice of point charges is a conditionally convergent series. This complicated mathematical behavior unfortunately leads to many delicate consequences in solid-state physics. One famous example is the dependence of the Madelung constant upon the shape of the truncation of the Coulomb series [1]. Another occurrence of this phenomenon is the well-known dependence of the work function upon the surface type [5].

In practical *ab initio* implementations, the subtleties related to the definition of the electrostatic potential are hidden deeply, owing to the choice of the Ewald summation technique together with the convention of zero average potentials [3,6]. This is of course an arbitrary choice. However, once this convention has been chosen, it has to be consistently propagated in the different electrostatic terms of the Hamiltonian: the ion-ion repulsion, the electron-ion attraction, and the electron-electron repulsion. For the all-electron (AE) methods that consider straightforwardly the physical nucleus attraction potential Z/r (in atomic units), this is not much of a problem. The situation is more complicated for the plane-wave methods using pseudopotentials. The valence electrons do not experience the bare ionic potential, but rather a smooth pseudopotential that induces an extra term in the total energy, namely, the difference between the average bare potential and the average pseudopotential [1,7]. This is the origin of the so-called "$Z\alpha$" term, first derived by Ihm and co-workers [8].

The situation becomes even more complex when turning to the projector augmented-wave (PAW) method. The PAW method, introduced by Blöchl [9], is an improvement over the pseudopotential approach. Owing to the PAW transformation, the pseudo-wave functions are mapping the true AE wave functions. The PAW bears many similarities with pseudopotentials, as demonstrated a few years later by Kresse and Joubert [10]. Most noticeably for our discussion, the pseudo-wave functions experience a pseudopotential, which requires a subtle treatment of the compensating background.

In this paper, we demonstrate that the current PAW total energy and Hamiltonian do not incorporate the compensating background contribution in a consistent manner. Extra terms have to be added to the potential of any system and to the energy of charged systems. It may appear counterintuitive that the origin of the potentials may have an effect on the physical properties. However, we show that the formation energy of a charged defect is indeed affected by an inconsistent treatment of the background. Although potential alignment techniques have been devised to circumvent the problem [11–14], a unanimous agreement in the literature about a unique definition that would work whatever the nature of the charged defect is still lacking. Only with these terms properly included could the PAW results be independent from the details of the PAW pseudopotential and could they adequately reproduce the reference AE calculations. As a by-product, we also propose

a modification of the total energy in the norm-conserving pseudopotential framework.

The paper is organized as follows: In Sec. II, we review the peculiarities of the Coulomb interaction in periodic systems. In Sec. III, we derive the PAW equations with a proper account of the compensating background density. Section IV provides the applications: the validation of the additional potential and energy terms by comparing to AE calculations and an application to a highly charged defect. Section V is devoted to the extension to the norm-conserving framework.

We will work in atomic units of length (1 bohr = 1), energy (1 hartree = 1), and action ($\hbar = 1$). However, two conventions for the atomic unit of charge are possible. While the common choice is to select a negative sign for the electronic charge, so that $e = -1$ (e.g., a charged Li vacancy in LiH is negatively charged), on the contrary, in the PAW literature an electron is given a positive charge, so that $e = +1$ [see, e.g., Ref. [10], shortly after Eq. (9)]. Such a choice does not affect the quantities in which two charges are multiplied by each other, namely, all contributions to the energy, as detailed below. However, it does have an influence on the sign of the electrostatic potential. Still, the potential felt by the electrons (e.g., the one present in the Schrödinger equation), obtained by multiplying the electrostatic potential by the electronic charge, is free of such a convention problem.

In Sec. IV, dealing with applications, we rely on the usual convention ($e = -1$). For the other sections, we avoid the problem of convention, either because the relevant quantities are invariant upon a charge sign change, or because we refer to the potential felt by the electrons (electronic potential) instead of the electrostatic potential.

II. COULOMB INTERACTIONS IN SOLIDS

To highlight the role of the different Coulomb interactions in a solid, the total energy E of a unit cell of solid can be grouped as different contributions to

$$E = T + E_{\text{Coul}} + E_{\text{xc}}, \qquad (1)$$

where T is the kinetic energy, E_{Coul} the Coulomb energy, and E_{xc} the exchange-correlation energy. In the present paper, we focus on the Coulomb term; the details of the other two terms will not be discussed any further. With these notations, the Coulomb energy gathers all the electrostatic interactions in the solid: the electron-electron interaction (also named the Hartree energy), the nucleus-nucleus interaction, and the electron-nucleus interaction [also referred to as the external potential in the density-functional theory (DFT) language].

A. Coulomb interaction

In the following, an extensive use of Coulomb integrals, potential, and energy will be necessary. Let us introduce some useful notations here.

The Coulomb interaction between charge densities n_1 and n_2 is defined as

$$\langle n_1, n_2 \rangle = \iint d\mathbf{r}_1 d\mathbf{r}_2 n_1(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} n_2(\mathbf{r}_2), \qquad (2)$$

where the integrals run over the complete space. The Coulomb interaction is linear, symmetric, and positive definite. It is then a scalar product.

The potential created by a charge density $n(\mathbf{r})$ is obtained from the Poisson equation, which reads

$$v_H[n](\mathbf{r}) = \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \qquad (3)$$

The potential is obviously linear with respect to its argument n.

The Coulomb self-energy $E[n]$ of charge distribution n reads

$$E[n] = \tfrac{1}{2} \langle n, n \rangle. \qquad (4)$$

The factor $\frac{1}{2}$ comes from the double counting of the interactions. This is the energy of the entire system. Note that this Coulomb self-energy could also be obtained from the potential

$$E[n] = \frac{1}{2} \int d\mathbf{r} n(\mathbf{r}) v_H[n](\mathbf{r}). \qquad (5)$$

When we turn to periodic system, it is more convenient to work with the energy per unit cell, $E[n]/N_{\mathbf{R}}$, where $N_{\mathbf{R}}$ stands for the number of unit cells contained in the full solid.

As explained in the Introduction, the *ab initio* implementations in periodic solids generally rely on the Ewald technique [4,6,15], which presents several subtle points due to the long-ranged Coulomb interaction $1/r$.

The Ewald technique proposes to evaluate the potential of a lattice of point charges with a compensating background:

$$\sum_{\mathbf{R}} \delta'_{\mathbf{R}}(\mathbf{r}) = \left[\sum_{\mathbf{R}} \delta(\mathbf{r} - \mathbf{R}) \right] - \frac{1}{\Omega}. \qquad (6)$$

$\mathbf{R}$ stands for the lattice vectors and Ω for the unit cell volume. Here and below, we emphasize that a charge distribution is charge neutral by adding a prime. The direct solution of the Poisson equation for this charge density is impossible, since the lattice of point charges would induce a semiconvergent series, the value of which is undefined.

Using the following short-range/long-range decomposition,

$$\frac{1}{r} = \frac{\text{erfc}(\eta r)}{r} + \frac{\text{erf}(\eta r)}{r}, \qquad (7)$$

the potential can be split into two contributions. The value of η does not influence the final result and can be tuned for numerical convenience. After some algebra, the potential created by the charge density of Eq. (6) can be written as the sum of two absolutely converging series up to a constant A:

$$v_H\left[\sum_{\mathbf{R}} \delta'_{\mathbf{R}} \right](\mathbf{r}) = \sum_{\mathbf{R}} \frac{\text{erfc}(\eta|\mathbf{r} - \mathbf{R}|)}{|\mathbf{r} - \mathbf{R}|}$$
$$+ \frac{4\pi}{\Omega} \sum_{\mathbf{G} \neq 0} \frac{e^{-G^2/4\eta}}{G^2} e^{i\mathbf{G} \cdot \mathbf{r}} + A. \qquad (8)$$

The first term in real space arises from the short-range interaction, whereas the second term arises from the long-range part and is conveniently evaluated in reciprocal space. In Eq. (8), there should be an additional dipole term which

is disregarded in the Ewald sums. In other words, this corresponds to immersing the solid inside a metallic cavity that would perfectly compensate the global dipole of the considered solid. Furthermore, the choice for the value of A is purely conventional. In general, the average value of v_H is set to zero and consequently $A = -\frac{\pi}{\eta^2 \Omega}$.

Within these conventions, the energies and potentials of neutral and charged solids are completely fixed. We insist that the choice of A is conventional. However, once the conventions are settled, the calculated energies and potentials should not depend on the *ab initio* calculation technique: In practice, PAW and AE should produce the same results. Furthermore, for charged systems it is customary to postprocess the results with so-called charge corrections [11–14,16,17]. It should be noted that the correction schemes also use the Ewald technique with a fixed convention for the value of A. This convention, however, rarely explicitly stated, is in all practical cases the zero average convention. It is then important that the same convention is used for both the electronic structure calculation and the postprocessing scheme.

B. The physical densities and the background compensated densities

In a solid, there are two charged particles: the electrons and the protons. The electrons will be treated quantum mechanically, whereas the protons will simply act as point charges. Furthermore, it is also convenient to distinguish the core electrons from the valence electrons.

In a solid, all the densities are periodic: They are unchanged by a translation of any lattice vector $\mathbf{R}$. Then, the valence electron density n integrates to N_v per unit cell,

$$\frac{1}{N_{\mathbf{R}}} \int d\mathbf{r} \, n(\mathbf{r}) = N_v. \tag{9}$$

All the integrals in this paper run over the whole solid. The quantities per unit cell are readily obtained by dividing by $N_{\mathbf{R}}$.

The core electron density n_c can be written as a sum over atomic sites:

$$n_c(\mathbf{r}) = \sum_{\mathbf{R}a} n_c^{\mathbf{R}a}(\mathbf{r}) \tag{10}$$

$$= \sum_{\mathbf{R}a} n_c^a(\mathbf{r} - \mathbf{R} - \tau_a), \tag{11}$$

where τ_a is the position of atom a in the unit cell. Here and consistently in the following, the densities with superscript $\mathbf{R}a$ are referred in the solid coordinates, whereas the densities with superscript a have the origin in the position of atom a. The core density is most commonly kept frozen in the atomic configuration within the PAW framework. The atomic core densities n_c^a are then considered as spherical, $n_c^a(r)$, and are obtained from the atomic data file in general. The core density integrates to N_c electrons per unit cell, so that the total electronic density $n + n_c$ integrates to $N_v + N_c = N$ in a unit cell.

The charge density of the nuclei is a sum of point charges:

$$q_Z(\mathbf{r}) = \sum_{\mathbf{R}a} q_Z^{\mathbf{R}a}(\mathbf{r}) \tag{12}$$

$$= \sum_{\mathbf{R}a} -eZ^a \delta(\mathbf{r} - \mathbf{R} - \tau_a). \tag{13}$$

Note that the latter definition is independent of the choice of sign for e, as discussed at the end of the Introduction. In order to link with the PAW literature, we introduce

$$n_Z(\mathbf{r}) = q_Z(\mathbf{r})/e. \tag{14}$$

n_Z integrates to $-Z$ in a unit cell, with

$$Z = \sum_a Z^a. \tag{15}$$

For convenience, the frozen densities that can be decomposed as a sum over atomic sites are often treated together, as an "ionic density" n_{Zc}:

$$n_{Zc}(\mathbf{r}) = n_Z(\mathbf{r}) + n_c(\mathbf{r}) \tag{16}$$

$$= \sum_{\mathbf{R}a} n_{Zc}^a(|\mathbf{r} - \mathbf{R} - \tau_a|). \tag{17}$$

The ionic density integrates to $-Z_{\text{ion}} = N_c - Z$.

Finally, the total charge density en_T which contains contributions from all charges (electrons and protons) is computed from

$$n_T = n_Z + n_c + n = n_{Zc} + n. \tag{18}$$

Some physical properties require one to calculate charged unit cells. Nonzero charges q are obtained whenever the number of protons is not balanced by the number of electrons in the cell:

$$q = Z - N = Z_{\text{ion}} - N_v. \tag{19}$$

However, the potential obtained from such an unbalanced density would diverge. In practice, a compensating background is added in order to ensure the global charge neutrality:

$$n_T'(\mathbf{r}) = n_T(\mathbf{r}) + \frac{q}{\Omega}. \tag{20}$$

We remind that we introduced the prime notation for charge compensated densities that average to zero.

With these definitions, the total Coulomb energy per unit cell in a solid can be explicitly written

$$E_{\text{Coul}} = \frac{1}{2N_{\mathbf{R}}} \langle n_T', n_T' \rangle - \frac{1}{2N_{\mathbf{R}}} \sum_{\mathbf{R}a} \langle n_Z^{\mathbf{R}a}, n_Z^{\mathbf{R}a} \rangle. \tag{21}$$

This is the Coulomb self-energy of n_T' with explicit removal of the nuclei self-interaction. The self-interaction energy of a point charge is infinite and therefore each term in the previous equations is infinite. Fortunately, the difference between the two terms remains finite. Although not mathematically correct, this way of writing the equations is extremely practical. The mathematical correctness would be recovered by considering Gaussian shaped nuclei instead of point nuclei and then performing the limit to vanishing Gaussian widths. This would unfortunately make the equations less readable. Equation (21)

is valid not only for neutral systems, but also for charged systems owing to the use of the background compensated n'_T.

III. PAW BACKGROUND TERMS

The PAW method allows one to reconstruct AE wave functions, AE densities, and AE expectation values of operators out of pseudoquantities. The technique has many conceptual and numerical advantages, but they are obtained at the expense of introducing intermediate densities and potentials. This makes the derivation of the equation less straightforward. Indeed, we show in this section that background terms in the potential and in the energy have been omitted so far.

A. PAW charge densities

The PAW transformation maps the physical valence densities n using smooth densities $\tilde{n}$ [9]. The physical densities have a full nodal structure in the vicinity of the atoms, whereas the smooth densities do not. The smooth density deviates from the physical density only inside spheres centered on atoms. In order to transform smooth densities into physical densities, just on-site corrections to the densities are necessary.

Then the physical valence density is written

$$n(\mathbf{r}) = \tilde{n}(\mathbf{r}) - \tilde{n}^1(\mathbf{r}) + n^1(\mathbf{r}), \tag{22}$$

where $\tilde{n}^1(\mathbf{r})$ is the spurious smooth density in the PAW spheres and $n^1(\mathbf{r})$ is the physical density in the spheres. We follow the standard notations [10,18,19]: The smooth quantities have a tilde and the on-site quantities have an exponent 1.

The PAW technique requires also the introduction of the smooth density $\tilde{n}_{Zc}(\mathbf{r})$ to mimic the ionic core density (core electrons plus protons) that we introduced in the previous section, $n_{Zc}(\mathbf{r})$. The potential created by $\tilde{n}_{Zc}$ plays the role of a pseudopotential for the smooth valence density $\tilde{n}$. $v_H[\tilde{n}_{Zc}]$ can be thought of as the local component in the pseudopotential scheme. It is the sum over atomic contribution

$$\tilde{n}_{Zc}(\mathbf{r}) = \sum_{\mathbf{R}a} \tilde{n}^{\mathbf{R}a}_{Zc}(\mathbf{r}) \tag{23}$$

$$= \sum_{\mathbf{R}a} \tilde{n}^a_{Zc}(|\mathbf{r} - \mathbf{R} - \tau_a|). \tag{24}$$

As consequence, the total density n_T including the core and valence electrons and the protons can be recast into three terms following in the same spirit as for Eq. (22):

$$n_T = \tilde{n}_T - \tilde{n}^1_T + n^1_T, \tag{25}$$

where

$$\tilde{n}_T = \tilde{n} + \hat{n} + \tilde{n}_{Zc}, \tag{26}$$

$$\tilde{n}^1_T = \tilde{n}^1 + \hat{n} + \tilde{n}_{Zc}, \tag{27}$$

$$n^1_T = n^1 + n_{Zc}. \tag{28}$$

The technical compensation charge $\hat{n}$ has been further added and subtracted in the total density. This last density is chosen

so that the moments in the multipole expansion of $n^1_T - \tilde{n}^1_T$ are zero. This is necessary to eliminate electrostatic interactions between PAW spheres.

Since n_{Zc} and $\tilde{n}_{Zc}$ charge distributions are monopole, carrying the same charge $-Z_{\text{ion}}$, $\hat{n}$ makes the moments of $n^1 - \tilde{n}^1 - \hat{n}$ vanish. Although the smooth density does not necessarily conserve the number of electrons, the sum $\tilde{n} + \hat{n}$ does.

B. Coulomb energy within PAW

The electrostatic energy per unit cell in the PAW framework can be written as

$$E_{\text{Coul}} = \frac{1}{2N_{\mathbf{R}}} \langle n'_T, n'_T \rangle - \frac{1}{2N_{\mathbf{R}}} \sum_{\mathbf{R}a} \langle n^{\mathbf{R}a}_{Zc}, n^{\mathbf{R}a}_{Zc} \rangle. \tag{29}$$

This expression is similar to Eq. (21) except that the core-core and core-nucleus interactions have been removed, since they only account for a change of origin in the total energies. We stress that the expression for the Coulomb self-energy of the charge distribution departs from the usual expression [10], as we explicitly introduced the compensating background in the n'_T densities. This has no consequence for the energy of neutral systems. However, it has one for charged systems, as we will show in the following.

The compensating background is homogeneous in the solid. It is practical then to include it in the smooth density,

$$n'_T = \tilde{n}'_T - \tilde{n}^1_T + n^1_T. \tag{30}$$

Then, we transform

$$\langle n'_T, n'_T \rangle = \langle \tilde{n}'_T, \tilde{n}'_T \rangle + 2\langle n^1_T - \tilde{n}^1_T, \tilde{n}'_T \rangle$$

$$+ \langle n^1_T - \tilde{n}^1_T, n^1_T - \tilde{n}^1_T \rangle. \tag{31}$$

This follows the usual PAW derivation except for the account of the compensating background in the smooth density. The last two terms can be evaluated on-site, since the charge distribution $n^1_T - \tilde{n}^1_T$ has vanishing moments outside the PAW spheres. Compared to the standard derivation, only two terms need to be explicitly derived: $\langle \tilde{n}'_T, \tilde{n}'_T \rangle$ and $\langle n^1_T - \tilde{n}^1_T, \tilde{n}'_T \rangle$. The following focuses on these two modified terms. We refer the reader to Ref. [10] for the usual terms, which are not detailed here.

1. The smooth density Coulomb self-energy

In fact, the explicit introduction of the compensating background in the term $\langle \tilde{n}'_T, \tilde{n}'_T \rangle$ does not yield any change compared to the standard implementations, since the average values of the Hartree potential and of the ionic pseudopotential are usually set to zero manually. Let us demonstrate this equivalence here.

The total smooth density $\tilde{n}'_T$ can be recast into two charge-neutral terms:

$$\tilde{n}'_T = \left(\tilde{n} + \hat{n} - \frac{N_v}{\Omega} \right) + \left(\tilde{n}_{Zc} + \frac{Z_{\text{ion}}}{\Omega} \right) \tag{32}$$

$$= (\tilde{n} + \hat{n})' + \tilde{n}'_{Zc}. \tag{33}$$

When inserted in $\langle \tilde{n}'_T, \tilde{n}'_T \rangle$, this decomposition turns to the familiar sum of Hartree energy, local pseudopotential energy, and ion-ion repulsion energy:

$$\frac{1}{2N_{\mathbf{R}}} \langle \tilde{n}'_T, \tilde{n}'_T \rangle = \frac{1}{2N_{\mathbf{R}}} \langle (\tilde{n} + \hat{n})', (\tilde{n} + \hat{n})' \rangle$$
$$+ \frac{1}{N_{\mathbf{R}}} \langle (\tilde{n} + \hat{n})', \tilde{n}'_{Zc} \rangle + \frac{1}{2N_{\mathbf{R}}} \langle \tilde{n}'_{Zc}, \tilde{n}'_{Zc} \rangle. \quad (34)$$

Owing to the explicit introduction of the compensating background charge, we immediately recognize that the Hartree energy in the previous equation is half the integral of the background compensated valence electron density $(\tilde{n} + \hat{n})'$ times the Hartree potential induced by the same charge density $v_H[(\tilde{n} + \hat{n})']$. Note that the average value of the Hartree potential is explicitly set to zero due to the vanishing average of $(\tilde{n} + \hat{n})'$ in the argument of v_H. The same remarks hold for the local pseudopotential energy (the second term on the right-hand side of the previous equation), which arises from the zero-averaged pseudopotential $v_H[\tilde{n}'_{Zc}]$. We have then identified the Hartree energy and local pseudopotential energy as they are usually calculated in practical codes.

The last term, together with the removal of the self-interaction [last sum in Eq. (29)], is the pseudo-ion/pseudo-ion repulsion with the convention of a vanishing average potential. It reduces to the usual point-charge/point-charge repulsion, usually named E_{Ewald}, plus the so-called $Z\alpha$ term [8]

$$E_{Z\alpha} = \frac{Z_{\text{ion}}}{\Omega} \sum_a \alpha^a, \quad (35)$$

where the integral α^a,

$$\alpha^a = \int d\mathbf{r} \left\{ v_H[\tilde{n}^a_{Zc}](\mathbf{r}) + \frac{Z^a_{\text{ion}}}{|\mathbf{r}|} \right\}, \quad (36)$$

measures the deviation in average potential between the pseudodensity $v_H[\tilde{n}_{Zc}]$ and a point charge $-Z_{\text{ion}}\delta(\mathbf{r})$

In the Appendix, we provide the full derivation of these two terms, since several expressions exist in the literature. The energy $E_{Z\alpha}$ is sometimes written with a factor Z_{ion} [8,20] or with a factor N_v [1,7]. As long as neutral systems are considered, the choice does not matter. However, for charged systems $N_v \neq Z_{\text{ion}}$, the total energy depends on the particular expression implemented. The Appendix demonstrates that the consistent expression should employ the factor Z_{ion}.

2. The background terms in the Coulomb energy

In Eq. (31), there is another occurrence of background charge density from the term $\langle n^1_T - \tilde{n}^1_T, \tilde{n}'_T \rangle$. We show in the following that this term adds extra terms in the total energy of a charged system.

In the usual derivation of the PAW energies, the density $\tilde{n}_T$ is replaced by its on-site projection $\tilde{n}^1_T$, since the integral in $\langle n^1_T - \tilde{n}^1_T, \tilde{n}_T \rangle$ does not have any contribution from outside the sphere, due to the vanishing moments of $n^1_T - \tilde{n}^1_T$. This transformation is approximate but believed to be very accurate [9,10]. It would be exact in the completeness limit of the projectors inside the PAW sphere.

When the background is also included, the transformation reads

$$\tilde{n}'_T \approx \tilde{n}^1_T + \frac{q}{\Omega}. \quad (37)$$

Whereas the term $\langle n^1_T - \tilde{n}^1_T, \tilde{n}^1_T \rangle$ is treated in the existing PAW derivations and will not be discussed further here, the background term

$$E_{\text{PAW bg}} = \frac{1}{N_{\mathbf{R}}} \left\langle n^1_T - \tilde{n}^1_T, \frac{q}{\Omega} \right\rangle \quad (38)$$

has not been explored so far.

This correcting term can be split for numerical convenience into

$$E^{(1)}_{\text{PAW bg}} = \frac{1}{N_{\mathbf{R}}} \left\langle n_{Zc} - \tilde{n}_{Zc}, \frac{q}{\Omega} \right\rangle, \quad (39)$$

$$E^{(2)}_{\text{PAW bg}} = \frac{1}{N_{\mathbf{R}}} \left\langle n^1 - \tilde{n}^1 - \hat{n}, \frac{q}{\Omega} \right\rangle. \quad (40)$$

These expressions can be decomposed on the PAW spheres, due to the vanishing moments of the left-hand side arguments of the Coulomb integrals. The first term bears striking similarities to the $Z\alpha$ term in the pseudo-ion/pseudo-ion part:

$$E^{(1)}_{\text{PAW bg}} = \frac{q}{\Omega} \sum_a \beta^a, \quad (41)$$

where the integral

$$\beta^a = \int d\mathbf{r} \left\{ v_H[n^a_{Zc}](\mathbf{r}) - v_H[\tilde{n}^a_{Zc}](\mathbf{r}) \right\} \quad (42)$$

can be precalculated from the PAW atomic data. The meaning of β^a is clear: It measures the difference between the physical potential due to nucleus and the core electrons $v_H[n^a_{Zc}]$ and the pseudopotential $v_H[\tilde{n}^a_{Zc}]$.

Figure 1 shows the potential used for the calculation of β^a for silicon and carbon. These two elements have the same number of valence electrons, however, the value of β^a depends much on the cutoff radius and on the pseudization scheme. In these examples, within these pseudization conditions, silicon has $\beta^a = -13.1$ hartree, whereas carbon has $\beta^a = -4.4$ hartree. This shows the wide range of possible values for β^a. Note that for carbon, the core electrons are few and therefore there is only a very small difference between the point-charge potential $-Z^a_{\text{ion}}/r$ and the physical potential $v_H[n^a_{Zc}]$. As a consequence, the integral β^a is close the opposite of integral $\alpha^a = 4.0$ hartree. For silicon, the core electrons are more widely spread, as can be appreciated in Fig. 1, and therefore the deviation of β^a from the opposite of $\alpha^a = 7.5$ hartree is noticeable.

Let us turn now to the second background term $E^{(2)}_{\text{PAW bg}}$. This term is slightly more complicated, since it explicitly depends on the valence density. It cannot be precalculated. However, it can be expanded on coefficients that can be precalculated and stored.
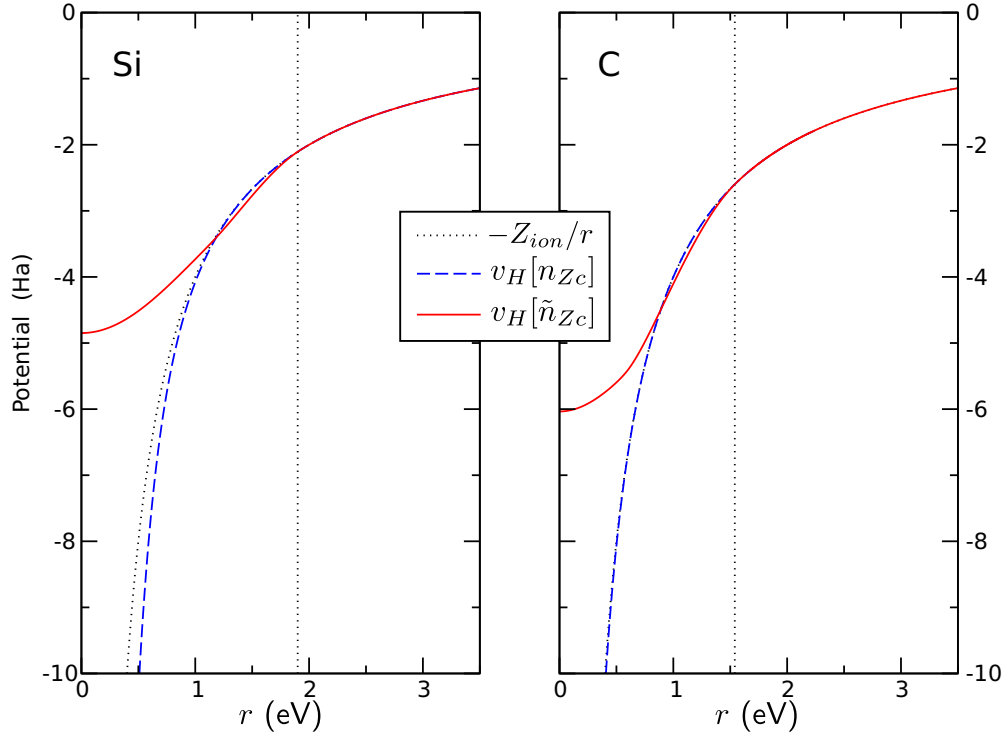
FIG. 1. (Color online) Spherical potentials involved in the integrals α^a and β^a exemplified for silicon in the left-hand panel and carbon in the right-hand panel. The point-charge potential is the dotted black line, the physical nucleus plus core electron potential $v_H[n_{Zc}]$ is the dashed blue line, and the pseudopotential $v_H[\tilde{n}_{Zc}]$ is the solid red line. Both pseudopotentials were generated with the Vanderbilt technique [21,22] using the cutoff radius symbolized with the vertical dotted line.

Indeed, the three densities in Eq. (40) can be expanded on the projectors inside each PAW sphere as

$$n^1(\mathbf{r}) = \sum_{\mathbf{R}aij} \rho_{ij}^a \phi_i^{a*}(\mathbf{r}-\mathbf{R}-\tau_a)\phi_j^a(\mathbf{r}-\mathbf{R}-\tau_a), \quad (43)$$

$$\tilde{n}^1(\mathbf{r}) = \sum_{\mathbf{R}aij} \rho_{ij}^a \tilde{\phi}_i^{a*}(\mathbf{r}-\mathbf{R}-\tau_a)\tilde{\phi}_j^a(\mathbf{r}-\mathbf{R}-\tau_a), \quad (44)$$

$$\hat{n}(\mathbf{r}) = \sum_{\mathbf{R}aij} \rho_{ij}^a \sum_{LM} \hat{Q}_{ij}^{aLM}(\mathbf{r}-\mathbf{R}-\tau_a). \quad (45)$$

The details of the algebra can be found, for instance, in Ref. [18]. Index a runs over atomic sites. Index i (and j) is a composite index for projector number n_i, and angular momenta l_i and m_i. ρ_{ij}^a is the density matrix in the basis of the projectors of site a. ϕ_i^a and $\tilde{\phi}_i^a$ are respectively the AE and the pseudo-wave functions for atom a and projector i. $\hat{Q}_{ij}^{aLM}$ are the coefficients of the multipole expansion of the compensation charge, with L and M being the angular momentum indexes.

Besides the density matrix ρ_{ij}^a, all these coefficients can be precalculated at the beginning of a PAW run. Let us gather these coefficients under the name γ_{ij}^a, so that the second background term can be written as

$$E_{\text{PAW bg}}^{(2)} = \frac{q}{\Omega} \sum_{aij} \rho_{ij}^a \gamma_{ij}^a. \quad (46)$$

The γ_{ij}^a are the average of the following potentials inside the sphere a,

$$\gamma_{ij}^a = \int d\mathbf{r} \, v_H \left[\phi_i^{a*}\phi_j^a - \tilde{\phi}_i^{a*}\tilde{\phi}_j^a - \sum_{LM} \hat{Q}_{ij}^{aLM} \right](\mathbf{r}). \quad (47)$$

Note that the multipole expansion of the density in the argument of v_H has zero moments, since the charge distribution $n^1 - \tilde{n}^1 + \hat{n}$ has a vanishing multipole expansion by construction of $\hat{n}$. However, this does not imply that the induced potential vanishes inside the sphere. It just vanishes outside the sphere.

All the terms in the argument of v_H in Eq. (47) need not be calculated. Indeed, performing the average in a sphere only selects the monopole of the potential and, as the Coulomb interaction $1/|\mathbf{r}-\mathbf{r}'|$ is diagonal in a multipole expansion, only the monopole of the charge distribution yields a nonvanishing contribution. Using the definition of $\hat{Q}_{ij}^{aLM}$ (see, e.g., Ref. [18]) and after some algebra, the only contribution in the argument of v_H in Eq. (47) that needs to be calculated, R_{ij}^a, reads

$$R_{ij}^a(r) = \frac{\delta_{l_i,l_j}\delta_{m_i,m_j}}{4\pi} \left\{ \frac{\phi_{n_il_i}^a(r)\phi_{n_jl_j}^a(r) - \tilde{\phi}_{n_il_i}^a(r)\tilde{\phi}_{n_jl_j}^a(r)}{r^2} \right.$$
$$\left. - g_0(r)\int dr' \left[\phi_{n_il_i}^a(r')\phi_{n_jl_j}^a(r') - \tilde{\phi}_{n_il_i}^a(r')\tilde{\phi}_{n_jl_j}^a(r') \right] \right\},$$
$$(48)$$

with $\phi_{n_il_i}^a(r)$ and $\tilde{\phi}_{n_il_i}^a(r)$ the radial AE and pseudo-wave functions, and $g_0(r)$ a shape function for angular momentum $l = 0$ [18].

The final expression of γ_{ij}^a is simply

$$\gamma_{ij}^a = \int dr 4\pi r^2 v_H[R_{ij}^a](r). \qquad (49)$$

As the first background term, the origin of the second background term is due to the introduction of a working quantity that modifies the smooth density compared to the AE density. This second background term is related to the existence of the compensation charge $\hat{n}$. In a norm-conserving (NC) framework, the density $\tilde{n}$ would integrate to N_v and no working density $\hat{n}$ would be required. The magnitude of this second term is not easily appreciated from its analytic expression. In the following section, we will show in a practical case that this term, though smaller than the first one, is indeed not negligible.

The total PAW background energy can be written as

$$E_{\text{PAW bg}} = \frac{q}{\Omega} \sum_a \left(\beta^a + \sum_{ij} \rho_{ij}^a \gamma_{ij}^a \right), \qquad (50)$$

where the coefficients β^a and γ_{ij}^a can all be calculated from the PAW atomic data at the beginning of a solid-state calculation. This extra term has to be added to the usual PAW total energy. It is zero for charge-neutral cells, however, it will modify the charged cell total energy.

C. Extra contribution to the PAW stress tensor

Since Eq. (50) depends on the volume of the cell Ω, there will be an additional term in the diagonal of the stress tensor of charged cells. The stress tensor needs to be corrected with the addition of $\sigma_{\text{PAW bg}}^{xx'}$ (x and x' indexes over the Cartesian axis):

$$\sigma_{\text{PAW bg}}^{xx'} = -\delta_{xx'} \frac{q}{\Omega^2} \sum_a \left(\beta^a + \sum_{ij} \rho_{ij}^a \gamma_{ij}^a \right). \qquad (51)$$

D. Extra contribution to the PAW potential

A less obvious consequence of the additional background energy is its influence on the PAW potential for both charged and neutral systems. Indeed, the Kohn-Sham potential is defined as a functional derivative with respect to the (physical) density [23]. In the PAW framework, this implies to differentiate with respect to the pseudodensity operator [9,10]. The extra term in the energy gives rise to a contribution to the potential named $v_{\text{PAW bg}}$.

The energy $E_{\text{PAW bg}}$ has an obvious dependence on the density matrix ρ_{ij}^a. However, it also has a dependence with respect to n through the factor $q = Z_{\text{ion}} - N_v$. Indeed, the number of valence electrons is a functional of the density

$$N_v = \int d\mathbf{r}\, n(\mathbf{r}) = \int d\mathbf{r}[\tilde{n}(\mathbf{r}) + \hat{n}(\mathbf{r})]. \qquad (52)$$

Taking the derivative of N_v with respect to $\tilde{n}$ and to ρ_{ij}^a (contained in $\hat{n}$) introduces the overlap operator $\hat{S}$ [10].

Therefore, the new background contribution to the nonlocal PAW potential is

$$\hat{v}_{\text{PAW bg}} = -\frac{\hat{S}}{\Omega} \sum_a \left(\beta^a + \sum_{ij} \rho_{ij}^a \gamma_{ij}^a \right)$$
$$+ \frac{q}{\Omega} \sum_{aij} |\tilde{p}_i^a\rangle \gamma_{ij}^a \langle \tilde{p}_j^a|, \qquad (53)$$

where $\tilde{p}_i^a$ are the PAW projectors.

The striking result is the existence of a correcting term in the potential even for neutral systems. Even though the energy correction of the neutral system is zero, its derivative with respect to the density is nonvanishing.

The changes introduced in the absolute value of the potential would affect all the eigenvalues with the same rigid shift. For instance, when referring the position of the band edges to the position of core states or to the average position of the electrostatic potential, the difference would remain unchanged and, consequently, the calculations of band offsets would remain unaffected [24,25]. However, when the composition of the solid is changed with the introduction of defects, the extra terms in the potential have a finite effect, as we will show in the following section.

IV. PAW APPLICATIONS TO CHARGED SYSTEMS

We have derived additional terms in the PAW energy, potential, and stress. This section provides practical examples for the influence of the extra contributions. The additional terms have been implemented in the PAW code ABINIT [26].

A. Lattice of protons

Our first example is a gedanken experiment that will not require any numerical calculation. Let us consider a lattice of protons, let us say, one proton per unit cell to fix the ideas, with no electrons. Of course, a neutralizing background is required to keep the total energy finite.

In this simplistic system, all the components of the total energy related to electrons are zero. In the conventional derivations of PAW, two terms remain: the Ewald point-charge/point-charge repulsion energy and the $Z\alpha$ energy. But one of these is actually spurious. Indeed, the Ewald repulsion energy is precisely the electrostatic self-energy of the charge distribution of the point-charge protons with their compensating background. However, the energy $E_{Z\alpha}$ should not be present, since it adds a contribution that depends on the local pseudopotential $v_H[\tilde{n}_{Zc}]$.

We demonstrate now that adding the background terms fixes the problem. The coefficients γ_{ij}^a have no effect since the density matrix ρ_{ij}^a vanishes. The physical core plus nucleus potential $v_H[n_{Zc}]$ reduces to $-Z_{\text{ion}}/r$, as there is no core electron either. As a consequence, comparing Eqs. (36) and (42), $\beta^a = -\alpha^a$ and thus $E_{\text{PAW bg}}^{(1)} = -E_{Z\alpha}$.

Owing to the background energy term $E_{\text{PAW bg}}^{(1)}$, the spurious $E_{Z\alpha}$ contribution is eliminated from the total energy. The total energy of the lattice of a proton with no electrons is then independent of the pseudization details, as we expect.

B. Charged vacancy in LiH: Benchmark against AE results

In order to check the validity of the PAW derivation and the magnitude of the additional terms in the energy and potential, it would be desirable to have a valid reference calculation for a charged system. We choose here to focus on the charged lithium vacancy in rocksalt LiH, V_{Li}^-. This particular system was selected for the small charge of the nuclei, so that AE calculations within a plane-wave basis set were tractable.

In practice, we employ model pseudopotentials with just a local component $-Z_a \, \mathrm{erf}(r/r_c)/r$ instead of the full potential $-Z_a/r$. We converge the calculation with respect to both the radius r_c and the plane-wave basis cutoff energy E_{cut}. Using this type of pseudopotential is equivalent to considering that the nuclei are Gaussian charge distributions with a spread r_c. Although the absolute energies are impossible to converge, the total energy differences and the potentials show a much smoother behavior with respect to r_c and E_{cut}. This procedure allows us to extract unambiguous AE data with parameters $r_c = 0.0025$ bohrs and $E_{cut} = 2000$ hartree, which still keeps the calculation cost low enough even for the eight-atom supercells.

The formation energy of the lithium vacancy $E_f(V_{Li}^-)$ is evaluated through

$$E_f(V_{Li}^-) = E(Li_3H_4^-) - E(Li_4H_4) + E(Li) - \epsilon_{VBM}(LiH), \tag{54}$$

where $E(Li_3H_4^-)$ is the total energy of a supercell, $E(Li)$ is the total energy of the isolated Li atom, and $\epsilon_{VBM}(LiH)$ is the valence band maximum of bulk LiH. This is the usual formation energy [27] with $q = -1$, with the Fermi level set to the top valence band, and with the chemical potential of Li fixed to the atom energy. Of course, supercells with only seven and eight atoms do not give a proper evaluation of the true formation energy, but our purpose is simply to compare the *ab initio* methods. For the same reason, we do not include any charge correction nor potential alignment [11–14,16,17]. The PAW and the AE quantities should in principle match on an absolute scale, since the same convention has been retained for the evaluation of the Coulomb potentials without the need for a correcting post-treatment.

In practice, the supercell consists of seven or eight unrelaxed atoms in a cubic supercell with an edge 7.60 bohrs, and the k-point sampling is a Γ-centered $4 \times 4 \times 4$ grid, within the local density approximation (LDA). The isolated lithium atom is placed in a supercell with the same dimensions, using a Γ-only sampling. The corresponding PAW calculations use an over converged cutoff energy of 30 hartree for the wave functions and 60 hartree for the dense grid. In Fig. 2, the PAW total energies and valence band maximum are compared to the AE results. The first series shows the total energy difference $E(Li_3H_4^-) - E(Li_4H_4) + E(Li)$ with or without inclusion of the background terms. The second series shows the valence band maximum of the bulk $\epsilon_{VBM}(LiH)$. The last series is the formation energy, i.e., the difference between the two previous series. Including the additional background terms has a sizable effect on the total energies and on the top valence energy. Even if these changes cancel each other out to some extent, the final physical quantity E_f is modified by the inclusion of the background terms. The two additional terms arising from
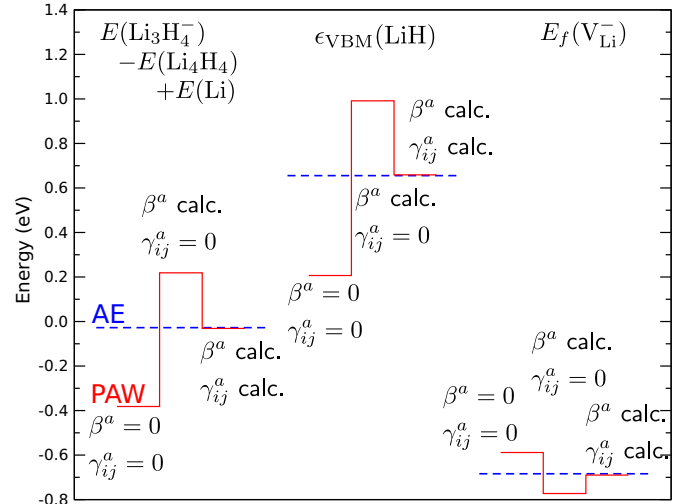

FIG. 2. (Color online) Contributions to the formation energy of a charged lithium vacancy V_{Li}^-, calculated with different PAW terms (solid red lines), compared to an AE reference (dashed blue lines). The first PAW calculations do not include β^a nor γ_{ij}^a, the second calculation does include β^a but not γ_{ij}^a, and the third calculation includes both terms. The formation energy is evaluated for the Fermi level fixed at the valence band maximum.

integrals β^a and γ_{ij}^a both have a visible effect. Only when the two are properly included could the PAW calculations reproduce the AE results.

As mentioned earlier, the PAW results without the background terms show a spurious dependence on the pseudization procedure. We evaluate this effect in Fig. 3 by varying the cutoff radius r_c^{loc} for the generation of the local pseudopotential $v_H[\tilde{n}_{Zc}]$ using the Vanderbilt procedure [21,22]. The corrected PAW results including the background terms are much more stable with respect to a change of pseudopotential than the uncorrected PAW data. The statement is not only true for the intermediate components such as the total energy difference or the top valence band energy [Figs. 3(a) and 3(b)], but also holds for the physical formation energy $E_f(V_{Li}^-)$ [Fig. 3(c)].

C. Highly charged interstitial SiC, Si_{TC}^{4+}: Benchmarking different codes

The magnitude of the background terms is proportional to the charge q of the defect. We now turn to a well-documented [28,29] charged defect of cubic silicon carbide, the silicon interstitial tetrahedrically coordinated to carbon atoms, Si_{TC}^{4+}.

In Fig. 4, we compare the Perdew-Burke-Ernzerhof (PBE) [30] formation energy of Si_{TC}^{4+} from three different PAW codes: VASP [31], QUANTUM ESPRESSO [32], and ABINIT [26]. In ABINIT, we have switched on and off the background correcting terms for the calculation of the formation energy. Note that for consistency QUANTUM ESPRESSO and ABINIT use the same PAW atomic data. This has not been possible for VASP, unfortunately, so that the VASP curve has been shifted up quite arbitrarily.

Carefully looking at the PAW implementation in the different codes, VASP and ABINIT without background have the same convention of setting the average smooth potential to zero, $\langle v_H[\tilde{n}_{Zc}]\rangle = 0$. Indeed, the slope of the corresponding
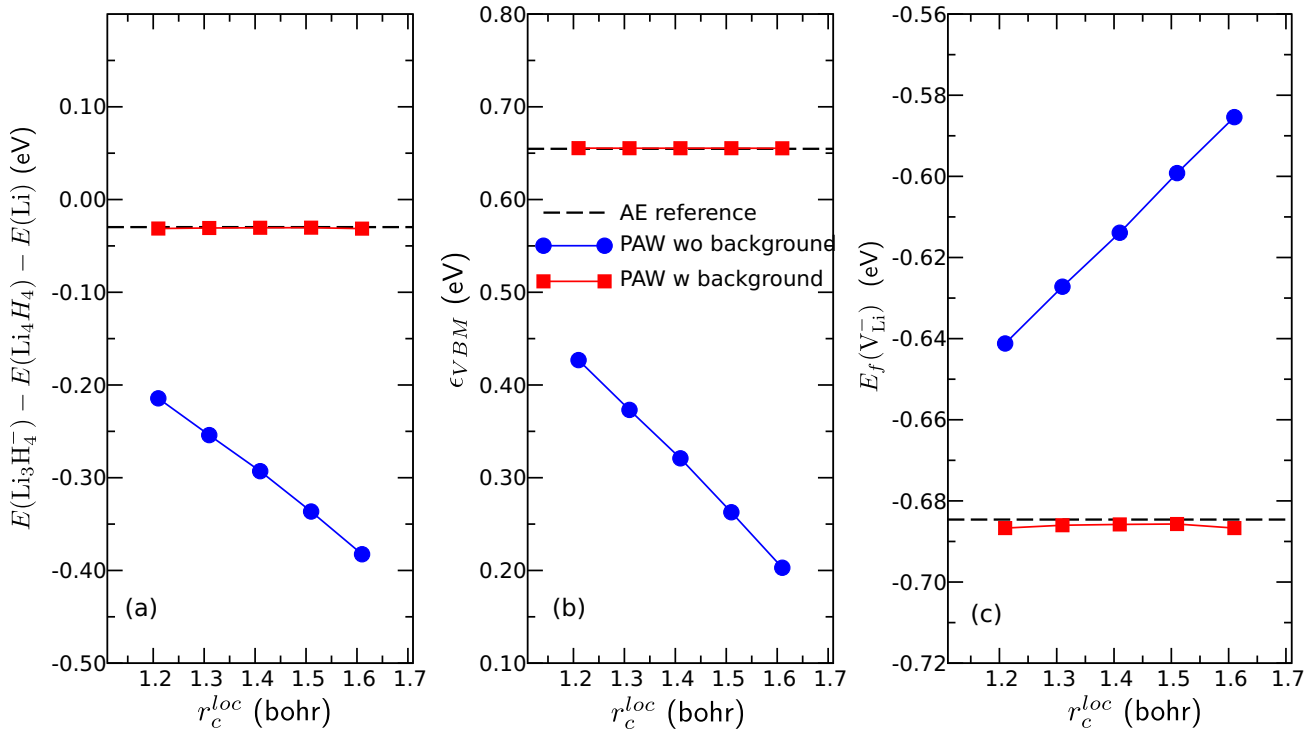
FIG. 3. (Color online) Dependence of uncorrected (blue circles) and corrected (red squares) PAW results with respect to the local potential cutoff radius r_c^{loc} in LiH. (a) represents the total energy difference in Eq. (54), i.e., the three first terms. (b) shows the valence band maximum of bulk LiH, i.e., the last term in Eq. (54). (c) shows the formation energy of the negatively charged Li vacancy with the Fermi level fixed at the valence band maximum.
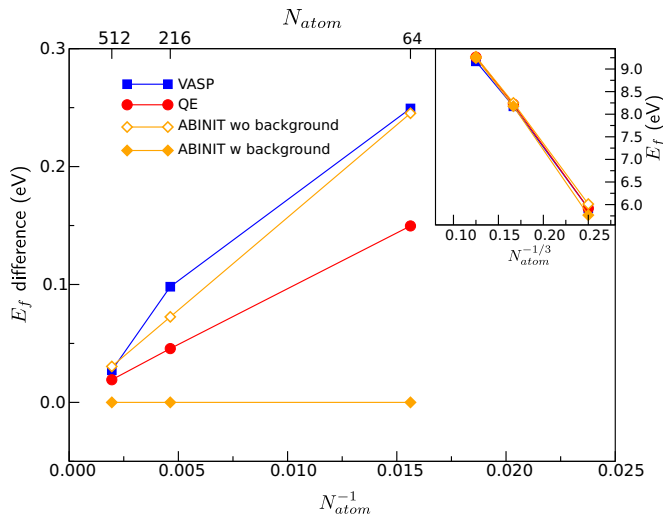

FIG. 4. (Color online) PAW formation energy E_f deviation for the silicon interstitial Si_{TC}^{4+} in cubic SiC as a function of supercell size within PBE with the Fermi level fixed at the valence band maximum and in silicon-rich conditions. Results have been obtained from VASP (squares), QUANTUM ESPRESSO (circles), ABINIT without the background terms (open diamonds), and ABINIT including the background terms (solid diamonds). This last curve has been chosen as a zero, so as to highlight the difference between the implementations. The absolute formation energy E_f is given in the inset, where the spurious charge-charge interaction dominates [16] (convergence as $N_{atom}^{-1/3}$). The QUANTUM ESPRESSO and ABINIT calculations use the same PAW atomic data, whereas it was necessary to shift up the VASP results by 100 meV.

curves matches. The QUANTUM ESPRESSO convention sets the average smooth potential to $\langle v_H[\tilde{n}_{Zc}] \rangle = \sum_a \beta^a / \Omega$. This produces a different slope. The meaning of this choice will be discussed in detail in the next section. Finally, the results from ABINIT with background terms are built to match the AE formalism using a total electrostatic potential that averages to zero $\langle v_H[n_{Zc}] \rangle = 0$.

For the interstitial Si_{TC}^{4+}, the difference between the conventions implemented in the codes quite significantly impacts the formation energy. For the 64-atom supercell, the difference can be as large as 0.25 eV and it is still 0.10 eV for the 216-atom supercell. Fortunately, the background terms are proportional to $1/\Omega$ and their effect should vanish with increasing supercell sizes.

V. CONSEQUENCES FOR THE PSEUDOPOTENTIAL METHOD

So far, we have stressed the importance of having a consistent convention for the potentials in the PAW method. The derivation for PAW in Sec. III highlights the role of different potentials: the pseudopotential, written $v_H[\tilde{n}_{Zc}]$ in the PAW language, and the true physical core electron plus nucleus potential, labeled $v_H[n_{Zc}]$. In an AE calculation, the situation is clear: The average value of the physical potential $v_H[n_{Zc}]$ is set to zero. However, within the pseudopotential framework, different choices can be found in the available implementations.

For instance, prior to version 7.5, the ABINIT code [26] uses the $Z\alpha$ energy term [8] and consequently sets the average

pseudopotential $v_H[\tilde{n}_{Zc}]$ to zero [33]. If the total energy expression uses instead a factor N_v in the $Z\alpha$ term [1,7], then the potential will be shifted accordingly. The potential is obtained as a functional derivative with respect to the electronic density and the number of valence electrons N_v is indeed a functional of the density. In QUANTUM ESPRESSO [32], for instance, this choice is made and the pseudopotential $v_H[\tilde{n}_{Zc}]$ averages to $\sum_a \alpha^a/\Omega$. This corresponds to a specific choice of the constant A introduced by the Ewald convention for the calculation of the potential $v_H[\tilde{n}_{Zc}]$. It could seem surprising to choose an inconsistent definition for the constant A.

Indeed, a consistent choice of the constant A that induces zero average for all the electrostatic potentials naturally leads to the absolute values as obtained with the $Z\alpha$ energy term, as derived in Ref. [8]. However, this choice induces a dependence of the energies and potentials on the pseudization details. It would be appreciated to devise a scheme which is independent of the pseudopotential and, even better, which reproduces as far as possible the absolute AE results. This could be achieved in practice by introducing the frozen core density, as we demonstrate in this section. This also gives an *a posteriori* justification for the convention using N_v in the $Z\alpha$ term.

A. Accounting for the physical core density in norm-conserving pseudopotentials

It is straightforward to adapt the PAW derivation of Sec. III to the simpler case of NC pseudopotentials. First of all, there is no equivalent to the integrals γ_{ij}^a in the pseudopotential framework, since there is no on-site representation of the charge density. But the integrals β^a, which measure the difference between the physical core electron+nucleus potential and the pseudopotential, still exist. Therefore, setting the physical potential to a zero average introduces an extra term in the total energy. The core electron and nucleus electrostatic energy reads

$$E_{\text{Coul}}^{Zc} = E_{\text{Ewald}} + E_{Z\alpha} + E_{\text{PAW bg}}^{(1)}, \tag{55}$$

where $E_{\text{PAW bg}}^{(1)}$ has been defined in Eq. (39), and detailed in Eq. (41). The origin of the usual terms E_{Ewald} and of $E_{Z\alpha}$ is recapitulated in the Appendix.

The addition of $E_{\text{PAW bg}}^{(1)}$ to the total energy does not modify the total energy of a charge-neutral cell. However, the potential as obtained from a functional derivative of the energy with respect to the density is affected, since $E_{\text{PAW bg}}^{(1)}$ has a dependence on N_v. The additional contribution to the potential $v_{\text{NC bg}}$ is a constant:

$$v_{\text{NC bg}} = -\frac{1}{\Omega} \sum_a \beta^a. \tag{56}$$

The calculation of the integrals β^a is straightforward: It just requires the knowledge of the physical frozen core electron density. This piece of information is available during the pseudopotential generation, but it is unfortunately generally not stored in the pseudopotential files. It would be direct to include it also.

Imagine now that the core density is much localized around the nucleus. In this case, the core+nucleus potential would only
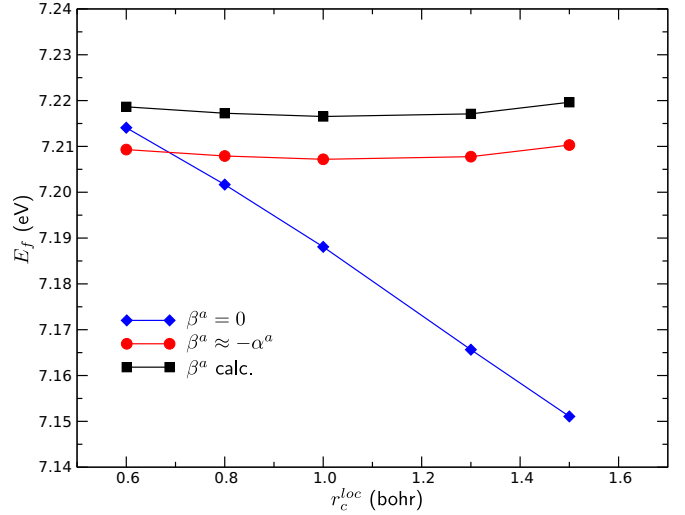

FIG. 5. (Color online) NC formation energy E_f within LDA of the unrelaxed doubly charged vacancy in diamond V_C^{2+} with the Fermi level fixed at the valence band maximum as a function of the cutoff radius of the local pseudopotential ($l = 2$), using $\beta^a = 0$ (blue diamonds), $\beta^a \approx -\alpha^a$ (red circles), or explicitly calculating β^a (black squares).

slightly deviate from the point-charge potential Z_{ion}/r, as we observed in the case of carbon in right-hand panel of Fig. 1. Then the β^a integrals would be very similar to α^a integrals:

$$\beta^a \approx -\alpha^a. \tag{57}$$

In this approximation, the following simplification occurs:

$$E_{Z\alpha} + E_{\text{PAW bg}}^{(1)} \approx \frac{N_v}{\Omega} \sum_a \alpha^a. \tag{58}$$

The present derivation gives an *a posteriori* justification for the total energy and potential formulas, which are written in some textbooks [1,7] and used in some codes, such as QUANTUM ESPRESSO [32].

B. Charged defect examples

We now test the effect of modifying the total energy and potential expressions in NC calculations of the formation energy of charged defects. Within the NC framework, it would be perfection to hope to obtain the same results as with reference AE calculations. However, it would be more realistic to have a weak dependence of the physical properties upon the pseudopotential details.

We consider in Fig. 5 a doubly charged vacancy V_C^{2+} in a 64-atom cubic supercell of diamond. For simplicity, the atoms are not relaxed, the lattice constant is set to 6.75 bohrs, and the cutoff energy is set to a very large value of 150 hartree. The carbon pseudopotential that has been generated with the Troullier-Martins technique [34] with a local component is $l = 2$. As the carbon valence electrons have mainly a sp character, a change in the d component of the pseudopotential should only indirectly impact the physical properties. It is interesting to focus on this particular defect which has been recently shown to be out of reach of the usual correction schemes due to its delocalized nature [14]. In Fig. 5, we
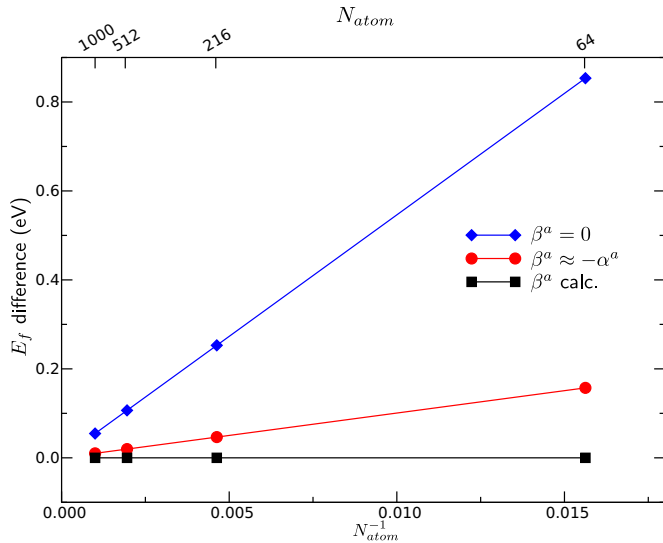
FIG. 6. (Color online) NC formation energy E_f deviation within LDA of the unrelaxed silicon interstitial Si_{TC}^{4+} in cubic SiC with the Fermi level fixed at the valence band maximum and in silicon-rich conditions as a function of supercell size, using $\beta^a = 0$ (blue diamonds), $\beta^a \approx -\alpha^a$ (red circles), or explicitly calculating β^a (black squares). This last curve has been chosen as a zero, so as to highlight the difference between the implementations. The absolute formation energy E_f shows the same behavior as in Fig. 4.

show the behavior of the formation energy of the charged vacancy as a function of the pseudopotential cutoff radius r_c^{loc}. The usual expression, which sets the average pseudopotential to zero ($\beta^a = 0$), presents a strong dependence on r_c^{loc}. If the zero of the potentials is defined with point-charge nuclei and core electrons ($\beta^a \approx -\alpha^a$), then the formation energy is remarkably stable. Furthermore, it deviates only slightly from the final results, which consistently set the origin of the potential through the true core electron plus nuclei density (actual calculation of β^a). It could be argued that all these changes have a small magnitude, however, the dependence on the details of the local component of the pseudopotential is clearly pathological.

Furthermore, when the defects involve adding or removing atoms with a wider core electron density, the effects can be significantly larger. Turning back to the silicon interstitial Si_{TC}^{4+} we used in the previous section, the added silicon atom has a larger core consisting of ten electrons. Figure 6 shows the difference in formation energy within LDA for the three choices of the potential origin. For clarity, the formation energy with an explicit calculation of β^a has been set to zero. The difference between the three schemes decreases as Ω^{-1}, however, for relatively large supercells (64–216 atoms), the difference can be as large as 0.25–0.80 eV. Except for the smallest supercell size, the approximation $\beta^a \approx -\alpha^a$ is a very decent approximation.

From these numerical applications, we conclude that it is important to include the background term in the total energy and in the potential. If the exact calculation of integrals β^a, though simple to perform, is not available, the approximation $\beta^a \approx -\alpha^a$ also yields reasonable results.

VI. CONCLUSIONS

In this paper, we derived two additional terms in the PAW energy of charged systems and in the PAW potential of all systems in order to reconcile AE calculations and PAW framework. These two terms [see Eq. (50)] arise from the proper treatment of the compensating background density, which is required by the use of periodic boundary calculations.

They are of a different nature. The first term measures the difference between the average smooth pseudopotential and the true physical nucleus plus the core electron potential. This contribution is usually the largest. Though smaller, the second term can also have a visible influence. The second term measures the difference in potential induced by the introduction of the compensation density $\hat{n}$.

The correct inclusion of these two terms has two positive consequences: It makes the PAW results directly comparable with AE calculations, and it makes the PAW results less sensitive to the PAW atomic data. We would like to stress that the proper treatment of the background terms not only affects the absolute energy values, but also impacts the physical quantities that are extracted, such as the formation energy or the relaxation volume of charged defects. The formation energy of the charged defects in small supercells or with a high charge state can be modified by several tenths of eV. Though these differences could also be reconciled by a potential alignment correction, a universal definition of the potential alignment is still missing.

For consistency with AE calculations, the first background term should also be introduced in the NC framework. This explains why different plane-wave codes could produce different physical results with the same pseudopotential data. The inclusion of the background term in Eq. (55) yields a total energy expression, which best approximates the AE results. However, its impact is even larger than experienced in the PAW framework. For the highly charged defect Si_{TC}^{4+} in SiC, the formation energy is still changed by 0.25 eV, even for the 216-atom supercell.

ACKNOWLEDGMENTS

APPENDIX: PSEUDO-ION/PSEUDO-ION INTERACTION

The expression for the pseudo-ion/pseudo-ion interaction varies in the literature. This Appendix is meant to fix this point.

The pseudo-ion/pseudo-ion repulsion energy is defined as

$$E_{Coul}^{Zc} = \frac{1}{2N_{\mathbf{R}}}\langle \tilde{n}'_{Zc}, \tilde{n}'_{Zc} \rangle - \frac{1}{2}\sum_a \langle \tilde{n}_{Zc}^a, \tilde{n}_{Zc}^a \rangle, \quad (A1)$$

where $\tilde{n}_{Zc}$ and $\tilde{n}_{Zc}^a$ have been defined in Eqs. (23) and (24).

Indeed, it is sometimes written as the Ewald point-charge/point-charge repulsion E_{Ewald} plus the celebrated $Z\alpha$ term [8,20]. However, from some other sources, the $Z\alpha$ term is replaced by an $N\alpha$ term, with N_v replacing the Z_{ion} in Eq. (35) [1,7]. We have stressed in Sec. II that once the Ewald

convention has been chosen for the electrostatic interactions, there should only be one expression for the total energy (the unknown constant is fixed and the global dipole is assumed to be zero).

Let us demonstrate here that the correct expression is indeed the $Z\alpha$ with factor Z_{ion}. We have to bridge the difference between the original charge distribution n_{Zc}, which is a sum over atomic site contributions, and the point-charge distribution used in the Ewald energy. The point-charge distribution n_{pc}, which reads

$$n_{pc}(\mathbf{r}) = \sum_{\mathbf{R}a} n_{pc}^a(\mathbf{r} - \mathbf{R} - \tau_a) \tag{A2}$$

$$= \sum_{\mathbf{R}a} -Z_{ion}^a \delta(\mathbf{r} - \mathbf{R} - \tau_a), \tag{A3}$$

has the same multipole expansion as the original distribution n_{Zc}, under the mild assumption that the PAW spheres are nonoverlapping spheres. However, we have to then introduce the compensating background both in n'_{Zc} and n'_{pc}.

Then transforming the term $\langle \tilde{n}'_{Zc}, \tilde{n}'_{Zc} \rangle$, we write

$$\langle \tilde{n}'_{Zc}, \tilde{n}'_{Zc} \rangle = \langle n'_{pc}, n'_{pc} \rangle + \langle \tilde{n}'_{Zc} - n'_{pc}, \tilde{n}'_{Zc} + n'_{pc} \rangle. \tag{A4}$$

In the last term, the backgrounds in $\tilde{n}'_{Zc} - n'_{pc}$ compensate and therefore the primes can be dropped there. As $\tilde{n}_{Zc}$ and

n'_{pc} are contained in the sphere and have the same multipole expansion, only the on-site terms remain:

$$\frac{1}{2N_{\mathbf{R}}} \langle \tilde{n}_{Zc} - n_{pc}, \tilde{n}'_{Zc} + n'_{pc} \rangle = \frac{1}{2} \sum_a \langle \tilde{n}_{Zc}^a - n_{pc}^a, \tilde{n}_{Zc}^a + n_{pc}^a \rangle$$
$$+ \sum_a \left\langle \tilde{n}_{Zc}^a - n_{pc}^a, \frac{Z_{ion}}{\Omega} \right\rangle, \tag{A5}$$

where the backgrounds have been written explicitly. The last term is precisely the $Z\alpha$ energy of Eq. (35).

Inserting the last equation in Eq. (A4), we obtain

$$\frac{1}{2N_{\mathbf{R}}} \langle \tilde{n}'_{Zc}, \tilde{n}'_{Zc} \rangle = \frac{1}{2N_{\mathbf{R}}} \langle n'_{pc}, n'_{pc} \rangle - \frac{1}{2} \sum_a \langle n_{pc}^a, n_{pc}^a \rangle$$
$$+ \frac{1}{2} \sum_a \langle n_{Zc}^a, n_{Zc}^a \rangle + E_{Z\alpha}. \tag{A6}$$

Reordering the terms in the last equation finally proves the announced result:

$$E_{Coul}^{Zc} = E_{Ewald} + E_{Z\alpha}. \tag{A7}$$

The result that the factor in energy $E_{Z\alpha}$ is not N_v could have been anticipated, since there is no reason to introduce the number of electrons in the energy E_{Coul}^{Zc} that only depends on the pseudopotential quantities.

---

[1] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, U.K., 2004), Vol. 1.

[2] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford Science, New York, 1987).

[3] R. A. Coldwell-Horsfall and A. A. Maradudin, J. Math. Phys. **1**, 395 (1960).

[4] S. W. de Leeuw, J. W. Perram, and E. R. Smith, Proc. R. Soc. London, Ser. A **373**, 27 (1980).

[5] A. W. Dweydari and C. H. B. Mee, Phys. Status Solidi A **27**, 223 (1975).

[6] P. P. Ewald, Ann. Phys. **369**, 253 (1921).

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

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

[9] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).

[10] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).

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

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

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

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

[15] L. M. Fraser, W. M. C. Foulkes, G. Rajagopal, R. J. Needs, S. D. Kenny, and A. J. Williamson, Phys. Rev. B **53**, 1814 (1996).

[16] M. Leslie and M. J. Gillan, J. Phys. C **18**, 973 (1985).

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

[18] M. Torrent, F. Jollet, F. Bottin, G. Zérah, and X. Gonze, Comput. Mater. Sci. **42**, 337 (2008).

[19] M. Torrent, N. Holzwarth, F. Jollet, D. Harris, N. Lepley, and X. Xu, Comput. Phys. Commun. **181**, 1862 (2010).

[20] X. Gonze, Phys. Rev. B **55**, 10337 (1997).

[21] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[22] N. Holzwarth, A. Tackett, and G. Matthews, Comput. Phys. Commun. **135**, 329 (2001).

[23] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).

[24] A. Baldereschi, S. Baroni, and R. Resta, Phys. Rev. Lett. **61**, 734 (1988).

[25] H. P. Komsa, E. Arola, E. Larkins, and T. T. Rantala, J. Phys.: Condens. Matter **20**, 315004 (2008).

[26] X. Gonze, B. Amadon, P. M. Anglade, J. M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Cote, T. Deutsch, L. Genovese, P. Ghosez, M. Giantomassi, S. Goedecker, D. R. Hamann, P. Hermet, F. Jollet, G. Jomard, S. Leroux, M. Mancini, S. Mazevet, M. J. T. Oliveira, G. Onida, Y. Pouillon, T. Rangel, G. M. Rignanese, D. Sangalli, R. Shaltaf, M. Torrent, M. J. Verstraete, G. Zerah, and J. W. Zwanziger, Comput. Phys. Commun. **180**, 2582 (2009).

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

[28] M. Bockstedte, A. Mattausch, and O. Pankratov, Phys. Rev. B **68**, 205201 (2003).

[29] G. Roma, F. Bruneval, T. Liao, O. N. B. Martinez, and J.-P. Crocombette, Defect Diffus. Forum **323–325**, 11 (2012).

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

[31] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[32] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys.: Condens. Matter **21**, 395502 (2009).

[33] Starting with ABINIT v7.5, the user has the possibility to change the definition of the average value of the Coulomb potential.

[34] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).