

**Electrostatics of solvated systems in periodic boundary conditions**Oliviero Andreussi<sup>1,2,\*</sup> and Nicola Marzari<sup>1,†</sup><sup>1</sup>*Theory and Simulations of Materials (THEOS), and National Center for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne, Station 12, 1015 Lausanne, Switzerland*<sup>2</sup>*Department of Chemistry, University of Pisa, Via Moruzzi 3, 56124 Pisa, Italy*

(Received 27 June 2014; revised manuscript received 28 October 2014; published 1 December 2014)

Continuum solvation methods can provide an accurate and inexpensive embedding of quantum simulations in liquid or complex dielectric environments. Notwithstanding a long history and manifold applications to isolated systems in open boundary conditions, their extension to materials simulations, typically entailing periodic boundary conditions, is very recent, and special care is needed to address correctly the electrostatic terms. We discuss here how periodic boundary corrections developed for systems in vacuum should be modified to take into account solvent effects, using as a general framework the self-consistent continuum solvation model developed within plane-wave density-functional theory [O. Andreussi *et al.*, *J. Chem. Phys.* **136**, 064102 (2012)]. A comprehensive discussion of real- and reciprocal-space corrective approaches is presented, together with an assessment of their ability to remove electrostatic interactions between periodic replicas. Numerical results for zero- and two-dimensional charged systems highlight the effectiveness of the different suggestions, and underline the importance of a proper treatment of electrostatic interactions in first-principles studies of charged systems in solution.

DOI: [10.1103/PhysRevB.90.245101](https://doi.org/10.1103/PhysRevB.90.245101)

PACS number(s): 31.15.-p, 31.70.Dk, 71.15.Dx, 71.15.Mb

**I. INTRODUCTION**

Computer simulations of materials have been significantly progressing in recent years due to the many improvements in both computational tools and underlying algorithms. In particular, density-functional theory (DFT) has become a very valuable tool to model complex systems with high accuracy. Even though a large effort in the field has been devoted to advancing the accuracy of the algorithms beyond the level of DFT, these improvements usually come with a substantial increase of the computational costs, therefore imposing some serious limitations on the system sizes that can be handled. For this reason, hierarchical algorithms have been developed, which allow us to treat different parts of the systems with different degrees of accuracy, without compromising the description of the important atomistic features that need to be characterized.

Among hierarchical methods, a fundamental role has been played by continuum dielectric models, which combined with *ab initio* and DFT atomistic calculations have been shown to be very effective in modeling solvents and complex environments in an inexpensive and accurate way [1–4]. Although most of the continuum dielectric models have been developed in the chemistry community and applied to study isolated systems, a large effort has been spent in recent years to extend these models to the boundary between condensed matter physics and chemistry [5–12]. In particular, the possibility of reducing the computational complexity of solvated or electrified interfaces would allow the extensive modeling of a large range of fundamental processes, such as those taking place in heterogeneous catalysis, electrochemistry, and photochemistry.

We recently proposed a self-consistent continuum solvation (SCCS) model [5,7,13] that combines a highly flexible definition of the dielectric, defined in terms of a minimal set of parameters, together with an implementation in a plane-wave pseudopotential DFT framework that is perfectly suited to model periodic solid-state systems. The model was tested thoroughly and showed not only an impressive agreement with similar models in the literature, but also very good performance in reproducing the experimental solvation free energies of neutral compounds [13] and charged species [14]. By taking advantage of fast Fourier transform (FFT) techniques to compute the electrostatic potential and its gradient in reciprocal space, the overall computational cost of SCCS is small, and its scaling with system size makes its impact negligible for large-scale calculations.

Nonetheless, solving for electrostatic potentials in reciprocal space is straightforward only when neutral fully periodic systems are considered. In the other cases, instead, FFT approaches can give rise to serious errors and strong system-size dependence, in vacuum as well as in a continuum dielectric. In particular, periodic boundary conditions are not compatible with charged systems, and so charged systems are modeled as if they were immersed in a neutralizing charge background (labeled NCB in the following). Moreover, when periodic boundary conditions are used to model heterogeneous or nonuniform systems, one needs to carefully monitor the size of the periodic cell chosen to avoid spurious interactions with the periodic replicas. For these reasons, it is well known that simulations of charged systems in the solid state (e.g., charged defects in semiconductors [15–26]) or in explicit solvents solutions (e.g., for solvation energies of charged ions [27,28]) need to deal with serious artifacts due to the size and periodicity of the simulation cell. This is particularly important when modeling systems of reduced dimensionality; nevertheless, the problem is intrinsically easier to handle than in the three-dimensional case. A wide variety of approaches has been proposed in the literature to remove the artifacts due

\*oliviero.andreussi@unipi.it

†nicola.marzari@epfl.ch

to the presence of fictitious replicas [19,28–44]. One class of methods (labeled here “non-self-consistent” or NSC) aims at correcting only the electrostatic energy of the systems, while keeping the degrees of freedom of the system frozen in the presence of periodic boundary conditions. This is the approach, e.g., of the Makov-Payne method [31], that is one of the most widespread methodologies to take care of PBC errors for zero-dimensional (0D) systems.

In order to fully remove the effects of periodic boundary conditions on partially periodic systems, other approaches (labeled as “self-consistent” or SC, in the following) have been developed that correct the electrostatic potential. This correction enters directly into the electrostatic energy, Kohn-Sham potential, and interatomic forces, such that the electrostatic energy has no spurious contributions from the periodic replica, but also all the degrees of freedom of the system are optimized in the correct electrostatic environment. These fully self-consistent correction schemes can be further divided in two classes, depending on whether the correction to the electrostatic potential is computed in real space (R-space) [40,41] or in reciprocal space (G-space) [33,35,36,42]. For both classes, correction for two-dimensional (2D), one-dimensional (1D), and 0D systems have been proposed and implemented.

In this work, some of the existing PBC correction schemes developed for partially periodic systems in vacuum are extended in order to take into account the presence of a continuum dielectric medium in the system. In the following, the three general classes of corrections, i.e., NSC, SC R-space, and SC G-space, are analyzed and the modifications of the algorithms needed to include a continuum dielectric are outlined. Equations for the most important cases are derived and the proposed approaches are implemented and tested.

The paper is organized as follows: In Sec. II A, we introduce the notation and the main electrostatic equations used throughout the paper; in Sec. II B, we review the main equations describing electrostatic interactions in periodic systems, highlighting the limitations of standard approaches; in Sec. II C, we summarize the equations behind the SCCS model, as derived in Ref. [13], underlining the effects of periodic boundary conditions; in Sec. II D, we describe the Makov-Payne approach [31] (NSC, 0D) and appropriately modify it in order to combine it with the SCCS model; in Sec. II E, the point-countercharge (PCC) correction scheme [40,41] is analyzed and extended to take into account of the complex dielectric environment, and its application to the case of slab geometries is presented (SC, R-space, 2D); in Sec. II F, the Martyna-Tuckerman method [33] is discussed and its modifications are derived and implemented for the case of isolated systems (SC, G-space, 0D); in Sec. III, we present detailed numerical results for the 0D and 2D cases; eventually, in Sec. IV we draw our conclusions.

## II. METHODS

### A. Electrostatics in periodic boundary conditions

In order to establish a consistent notation, we report here the main electrostatic equations, as reported in many standard textbooks but with a specific focus on their form in periodic systems. Electrostatic interactions are governed by Maxwell’s

equations, which relate electric field  $\mathbf{E}(\mathbf{r})$  and charge density  $\rho(\mathbf{r})$ :

$$\nabla \cdot \mathbf{E}(\mathbf{r}) = 4\pi\rho(\mathbf{r}), \quad (1)$$

$$\nabla \times \mathbf{E}(\mathbf{r}) = 0. \quad (2)$$

Due to the irrotational nature of the electrostatic field, it is often convenient to express it in terms of the gradient of a scalar potential, i.e., the electrostatic potential, as

$$\mathbf{E}(\mathbf{r}) = -\nabla v(\mathbf{r}) \quad (3)$$

and Eqs. (1) and (2) are recast into a single second-order differential equation, i.e., the Poisson equation

$$\nabla^2 v(\mathbf{r}) = -4\pi\rho(\mathbf{r}). \quad (4)$$

Once a proper set of boundary conditions is imposed, the above differential equation can be solved exactly. In particular, in a closed volume of space it is sufficient to specify the potential (Dirichlet boundary conditions) or the normal component of the field (von Neumann boundary conditions) at the boundary in order to have a unique solution of the electrostatic problem. Also, it is customary to recast Eq. (4) in an integral formulation by the use of Green’s functions, namely,

$$v[\rho](\mathbf{r}) \equiv v(\mathbf{r}) = \int_B G(\mathbf{r} - \mathbf{r}')\rho(\mathbf{r}')d\mathbf{r}', \quad (5)$$

where the integration is performed over the arbitrary bounded region  $B$ . In the above equation and in the following, we decided to make explicit the functional dependence of the potential on the density that generates it.

Given the definitions above, the electrostatic energy of a charge distribution can then be expressed as

$$E[\rho] = \frac{1}{8\pi} \int_B |\mathbf{E}|^2 d\mathbf{r}. \quad (6)$$

For an isolated charge density in vacuum, it is customary to impose homogeneous Dirichlet or von Neumann conditions at infinity, such that

$$E[\rho] = \frac{1}{2} \int_B \rho(\mathbf{r}) v[\rho](\mathbf{r}) d\mathbf{r} \quad (7)$$

and

$$G(\mathbf{r} - \mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|}. \quad (8)$$

For this class of systems, both the potential and the energy can be easily computed by exploiting Eq. (8) and by setting the integrand limit in Eqs. (7) and (5) to an arbitrary cell size  $D$  large enough to contain the entire charge density of the system

$$v[\rho](\mathbf{r}) = \int_D G(\mathbf{r} - \mathbf{r}')\rho(\mathbf{r}')d\mathbf{r}', \quad (9)$$

$$E[\rho] = \frac{1}{2} \int_D \rho(\mathbf{r}) v[\rho](\mathbf{r}) d\mathbf{r}. \quad (10)$$

Nonetheless, the characteristic  $1/r$  behavior of the electrostatic potential can be the source of two specular problems: the divergence at short distances and the slow decay at large distances make the electrostatic potential difficult to handle,

introducing issues with the self-interaction of charges and of conditionally convergent calculations of the field.

In periodic systems, the fundamental electrostatic equations, e.g., Eqs. (9) and (10), may be written in the same form reported above, whereas it is intended that the integration domain corresponds to the periodic unit cell, typically chosen as the primitive one, and the physical quantities entering the equations (density, potential, Green's function, etc.) refer to such infinitely periodic systems. In order to avoid confusion on which kind of system is considered, in all the equations in the following sections, we decided to use special typographic characters ( $\varrho$ ,  $\mathbf{E}$ ,  $v$ ,  $\Xi$ ,  $G$ , and  $D$ ) to identify quantities referring to infinite periodic systems, while keeping the standard labels ( $\rho$ ,  $E$ ,  $v$ ,  $\mathbf{E}$ ,  $G$ , and  $D$ ) for localized isolated systems.

In a periodic system, the entire, infinite, charge density  $\varrho(\mathbf{r})$  will contribute to the potential  $v(\varrho, \mathbf{r})$ . Nonetheless, such a potential can still be expressed univocally with an integral confined to the unit cell  $D$  of the periodic system, by exploiting in Eq. (5) the Green's function  $G(\mathbf{r} - \mathbf{r}')$  appropriate for periodic boundary conditions

$$v[\varrho](\mathbf{r}) = \int_{\infty} G(\mathbf{r} - \mathbf{r}')\varrho(\mathbf{r}')d\mathbf{r}' = \int_D G(\mathbf{r} - \mathbf{r}')\varrho(\mathbf{r}')d\mathbf{r}'. \quad (11)$$

Similarly, Eq. (7) can also be used as is in order to compute the electrostatic energy per unit cell of a periodic system  $E[\varrho]$ , provided that the integration is over the unit cell  $D$  of the periodic system

$$E[\varrho] = \frac{1}{2} \int_D \varrho(\mathbf{r}) v[\varrho](\mathbf{r}) d\mathbf{r}. \quad (12)$$

### B. Periodic electrostatic potential

When dealing with periodic systems, it is natural to recast the electrostatic equations in reciprocal space, in order to exploit the simple form of the Fourier-transformed differential operator

$$\nabla f(\mathbf{r}) \rightarrow \widetilde{\nabla} f(\mathbf{k}) = i\mathbf{k} \widetilde{f}(\mathbf{k}), \quad (13)$$

$$\nabla \cdot \mathbf{F}(\mathbf{r}) \rightarrow \widetilde{\nabla} \cdot \widetilde{\mathbf{F}}(\mathbf{k}) = i\mathbf{k} \cdot \widetilde{\mathbf{F}}(\mathbf{k}), \quad (14)$$

where the overwritten tilde identifies Fourier-transformed functions. By applying the above relations to Eqs. (1) and (3), the general solution of the electrostatic field and potential in a periodic system can be written as

$$\Xi(\mathbf{k}) = -4\pi \frac{i\mathbf{k}\varrho(\mathbf{k})}{|\mathbf{k}|^2} \text{ for } \mathbf{k} \neq 0 \quad (15)$$

and

$$v[\varrho](\mathbf{k}) = \frac{i\mathbf{k} \cdot \Xi(\mathbf{k})}{|\mathbf{k}|^2} = 4\pi \frac{\varrho(\mathbf{k})}{|\mathbf{k}|^2} \text{ for } \mathbf{k} \neq 0. \quad (16)$$

For  $\mathbf{k} = \mathbf{0}$ , the electrostatic equations need to be handled with care. Indeed, special forms of the divergence theorem impose that a periodic solution for the electrostatic field and potential is only possible provided that the right-hand side of Eq. (1) and the left-hand side of (3), once transformed in Fourier space, are zero for  $\mathbf{k} = \mathbf{0}$ . In particular, in order to obtain a periodic solution for the electrostatic field, the total

charge of the system has to be zero:

$$\varrho(\mathbf{k} = \mathbf{0}) \equiv \langle \varrho \rangle = \frac{1}{V} \int_D \varrho(\mathbf{r}) d\mathbf{r} = 0. \quad (17)$$

Similarly, a periodic solution of the electrostatic potential will only be possible for a zero average electrostatic field

$$\Xi(\mathbf{k} = \mathbf{0}) = \frac{1}{V} \int_D \Xi(\mathbf{r}) d\mathbf{r} = 0. \quad (18)$$

As this latter condition univocally fixes the constant value of the electrostatic field, the only undefined quantity for  $\mathbf{k} = \mathbf{0}$  is the potential: given that the system is neutral, such component has no effects on the final electrostatic energy

$$\frac{1}{2} \int_D v[\varrho](\mathbf{k} = \mathbf{0}) \varrho(\mathbf{k} = \mathbf{0}) d\mathbf{r} = 0. \quad (19)$$

Even if  $\varrho$  is defined to be non-neutral inside the unit cell, Eqs. (15) and (16) can still be used exactly as written, together with the choice  $v(\mathbf{k} = \mathbf{0}) = 0$ , but the quantities obtained will actually correspond to a periodic system where the original charge density has been compensated by a homogeneous background (NCB)

$$\varrho \rightarrow \varrho - \langle \varrho \rangle. \quad (20)$$

The specific choice  $v(\mathbf{k} = \mathbf{0}) = 0$  is made so that the NCB density does not appear explicitly in the formulas since its only contribution to the energy, i.e., the term for  $\mathbf{k} = \mathbf{0}$ , cancels out in Eq. (19). Nonetheless, for the sake of correctly identifying the physical system under consideration, in the following we will explicitly write the dependence of the potential on the compensated charge density of the system, namely,  $v[\varrho - \langle \varrho \rangle](\mathbf{k})$ .

It has to be noted that the above equations have been derived for ideally infinite periodic systems, but it could be convenient to take a different, real-space, perspective and to think of a periodic system as generated by an increasingly larger number of unit cells. In such a picture, while the reciprocal-space approach can still be used to look for periodic solutions of the electrostatic field and potential, it is physically acceptable to have an additional nonperiodic, but linear, component for the electrostatic potential. In other words, an additional linear potential of the form  $\Xi_0 \cdot \mathbf{r}$  would still preserve the periodic solution for the electrostatic field, and thus a physically acceptable solution for the energy of the periodic system. Moreover, for the same reasons, the  $\mathbf{k} = \mathbf{0}$  component of the potential will not have any effect on the total energy of a neutral system.

As the  $\mathbf{k} = \mathbf{0}$  components of the electrostatic field and potential cannot be univocally determined by the electrostatic differential equations, they can only be determined by the boundary conditions imposed on the system. Exploiting Eq. (11), the general solution for the electrostatic potential of a periodic system can be written as

$$v[\varrho](\mathbf{r}) = \frac{4\pi}{V} \sum_{\mathbf{k} \neq 0} \frac{\varrho(\mathbf{k})}{|\mathbf{k}|^2} e^{i\mathbf{k} \cdot \mathbf{r}} + \Xi_0 \cdot \mathbf{r} + v_0, \quad (21)$$

where the last two terms are usually referred in the literature as the extrinsic potential [27,28,45] to distinguish them from the intrinsic part, which can be solved independently of the

boundary conditions. In most reciprocal-space approaches to the electrostatic potential, only the intrinsic part of the potential is computed, while the extrinsic contributions are assumed to be equal to zero. This choice corresponds to a specific assumption on the boundary conditions of the electrostatic problem (spherical surface and tin-foil boundary conditions, as discussed in the following) and it can introduce artifacts in periodic calculations of partially periodic and nonperiodic systems.

In order to further analyze the expression of the extrinsic contributions, we can follow the derivation of de Leeuwen *et al.* [46–49] and treat the infinite periodic system as a limiting case of a spherical ensemble of unit cells embedded in a vacuumlike dielectric. Such a choice univocally determines the electrostatic equations and corresponds to the usual boundary conditions from which Eq. (8) was derived. Thus, the potential can be expressed as

$$\begin{aligned} v[\varrho](\mathbf{r}) &= \int_{\infty} G(\mathbf{r} - \mathbf{r}') \varrho(\mathbf{r}') d\mathbf{r}' \\ &= \sum_{\mathbf{R}} \int_{\mathcal{D}} G(\mathbf{r} - \mathbf{r}' + \mathbf{R}) \varrho(\mathbf{r}' + \mathbf{R}) d\mathbf{r}' \\ &= \int_{\mathcal{D}} \left[ \sum_{\mathbf{R}} G(\mathbf{r} - \mathbf{r}' + \mathbf{R}) \right] \varrho(\mathbf{r}') d\mathbf{r}' \end{aligned} \quad (22)$$

from which, comparing with Eq. (5), the periodic Green's function can be defined as

$$G(\mathbf{r} - \mathbf{r}') = \sum_{\mathbf{R}} G(\mathbf{r} - \mathbf{r}' + \mathbf{R}) = \sum_{\mathbf{R}} \frac{1}{|\mathbf{r} - \mathbf{r}' + \mathbf{R}|}, \quad (23)$$

where the sum over lattice vectors  $\mathbf{R}$  is supposed to be performed over spherical shells around the origin. As thoroughly discussed by Makov and Payne [31], the contribution of the terms in the periodic sum that determines the electrostatic potential vanishes as

$$\frac{q^{(n)}}{l^{n+1}}, \quad (24)$$

where  $q^{(n)}$  is the  $n$ th multipole moment of  $\varrho(\mathbf{r})$  and  $l$  is the distance of the shell from the origin. Similarly, the contribution of each shell of the periodic system to the electrostatic field in the original cell will vanish as the inverse  $n + 2$  power of  $l$ . For a three-dimensional system, the periodic sum that determines the potential (field) is divergent for a charge distribution with nonzero dipole (monopole) moment. This behavior corresponds to the impossibility, shown above, of obtaining a periodic solution for the potential (field) in reciprocal space for a system with nonzero electric field (total charge). Moreover, the periodic sum that determines the potential (field) is conditionally convergent for a charge distribution with nonzero quadrupole (dipole) moment, while it is absolutely convergent for higher multipole moments. Conditional convergence implies that the results will depend on the order over which the sum is performed and on the boundary conditions applied. This can be thought as the result of the fact that a periodic ensemble of quadrupole moments (dipoles) generates a nonzero surface distribution of dipoles (charges), which in turn will give rise to a nonzero

average electrostatic potential (field) inside the system. The magnitude of these quantities will depend on the geometry of the surface of the system and on the dielectric properties of the embedding medium. For the assumptions made above (spherical system embedded in vacuum), the expression for the extrinsic contributions to the potential, first derived by de Leeuw *et al.* [47,48], reads as

$$\Xi_0 = \frac{4\pi}{3} \frac{1}{V} \int_{\mathcal{D}} \mathbf{r} \varrho(\mathbf{r}) d\mathbf{r} \equiv \frac{4\pi}{3} \frac{\mathbf{d}}{V}, \quad (25)$$

$$v_0 = \frac{2\pi}{3} \frac{1}{V} \int_{\mathcal{D}} \mathbf{r}^2 \varrho(\mathbf{r}) d\mathbf{r} \equiv \frac{2\pi}{3} \frac{Q}{V}. \quad (26)$$

The above expressions have been recently rederived, for the same system shape and boundary conditions, by Hunenberger *et al.* by following a different approach [27,28,45]. In particular, it is important to notice that the constant electric field that appears in Eq. (25) is nothing but the electrostatic field generated by a constant polarization density  $\mathbf{P} = \mathbf{d}/V$ .

The extrinsic contributions to the electrostatic potential can be further extended to the case of a system embedded in a dielectric medium with arbitrary dielectric permittivity  $\epsilon$ , while still keeping the assumption of a spherical geometry. In this case, the Onsager model of solvation [50] analytically reduces the effects of the embedding medium to an additional reaction field that, for the case of a dipolar system, is again constant inside the system. The classical expression for the Onsager reaction field [50] gives

$$\mathbf{E}_R = -\frac{4\pi}{3} \frac{2(\epsilon - 1)}{2\epsilon + 1} \frac{\mathbf{d}}{V}, \quad (27)$$

which summed to the constant field obtained in vacuum gives the final result of

$$\Xi_0^\epsilon = \frac{4\pi}{3} \frac{\mathbf{d}}{V} - \frac{4\pi}{3} \frac{2(\epsilon - 1)}{2\epsilon + 1} \frac{\mathbf{d}}{V} = \frac{4\pi}{2\epsilon + 1} \frac{\mathbf{d}}{V}. \quad (28)$$

This expression reduces to the case in vacuum for  $\epsilon = 1$ , while vanishing when the periodic system is immersed in a perfect conductor (tin-foil boundary conditions, i.e.,  $\epsilon = \infty$ ).

To summarize, when dealing with the electrostatic equations in periodic systems, two main limitations occur. First, the total charge of the system needs to be zero, in order to provide a nondiverging solution for the electrostatic field and the energy of the system. Charged unit cells can still be treated using Eq. (21), but the potential obtained will be the one of the charge density considered plus a neutralizing homogeneous background charge (NCB). Second, by using the standard reciprocal-space approach for the calculation of the potential of a periodic system and by neglecting the extrinsic contributions to the potential [Eq. (21)], a well-defined choice on the boundary conditions of the problem is made, which can introduce spurious contribution to the energy.

In addition to the issues alluded to above, the long-range decay of the potential represents a serious drawback for simulations where periodic boundary conditions are only used as an approximation to model heterogeneous systems without introducing surface effects. When studying charged defects in crystals, or solvation energies of ions and biomolecules in explicit solvents, the electrostatic interactions coupled with the fictitious periodicity of the cell introduce artifacts in the simulations that are challenging to handle.

The problem is even more evident, although easier to solve, when one wants to model systems of reduced periodicity, being them slabs (2D), linear systems (1D), or isolated compounds (0D). The problem in these cases is twofold: first, the electrostatic potential of the ideal isolated system would not usually show the same periodicity of the simulation unit cell, thus, it cannot be obtained as a solution of a Poisson equation that obeys periodic boundary conditions; second, it is usually computationally convenient to exploit the Fourier-transform approach of perfectly periodic systems as derived in Eq. (21), thus automatically introducing spurious interactions with periodic replicas of the unit cell.

The two shortcomings discussed above can be solved independently. In particular, auxiliary-function methods are able to screen in reciprocal space the long-range part of the electrostatic potential. Thus, interactions with spurious periodic replicas are removed, even though the computed potential still retains the (incorrect) periodicity of the simulation cell. On the other hand, since the system is anyway confined in a restricted part of the simulation cell, in order to have a correct estimate of the electrostatic energy it is not necessary to have the electrostatic potential described accurately everywhere in the unit cell, but it is only important to have the correct potential in the region where the source charges are located. For this reason, the isolated system of interest is usually treated inside large supercells, in such a way that deformations of the potential due to the boundary of the cell do not affect the calculation of the electrostatic energy of the system. We note in passing that an alternative real-space approach has been recently proposed that is able to recover the ideal potential of the system in a computationally effective way by using a multigrid method to correct the 3D FFT-based potential [40,41].

### C. Electrostatics in dielectric environments and periodic boundary conditions

We summarize here the main equations behind continuum solvation, and in particular as embodied in the SCCS model [13]. The quantum-mechanical system of interest is immersed in a dielectric medium characterized by a density-dependent dielectric constant. A dielectric function is defined in order to ensure that the dielectric constant is equal to one in the interior of the solute, where the electronic density is high, and smoothly acquires the value of the bulk dielectric permittivity of the solvent  $\epsilon_0$ , where the electronic density goes to zero. An optimal definition of the dielectric function was provided in Ref. [13] in terms of only two tunable thresholds. For the sake of simplicity, in our notation in the following we will not highlight the specific functional definition of the dielectric function  $\epsilon[\rho^{el}(\mathbf{r})]$ , and only consider it as a continuous function  $\epsilon(\mathbf{r})$  defined everywhere in the simulation cell.

In the presence of a dielectric continuum, the electrostatic potential will be the solution of the generalized Poisson equation

$$\nabla \cdot \epsilon(\mathbf{r}) \nabla v^\epsilon[\rho^{\text{solute}}](\mathbf{r}) = -4\pi\rho^{\text{solute}}(\mathbf{r}), \quad (29)$$

where the superscript  $\epsilon$  has been added to distinguish the potential from the one computed in vacuum. By introducing a

polarization charge density

$$\begin{aligned} \rho^{\text{pol}}(\mathbf{r}) &= \nabla \cdot \left( \frac{\epsilon(\mathbf{r}) - 1}{4\pi} \nabla v^\epsilon[\rho^{\text{solute}}](\mathbf{r}) \right) \\ &= \frac{1}{4\pi} \nabla \ln \epsilon(\mathbf{r}) \cdot \nabla v^\epsilon[\rho^{\text{solute}}](\mathbf{r}) - \frac{\epsilon(\mathbf{r}) - 1}{\epsilon(\mathbf{r})} \rho^{\text{solute}}(\mathbf{r}), \end{aligned} \quad (30)$$

the generalized Poisson equation in a dielectric can be recast into a vacuumlike Poisson equation

$$\begin{aligned} \nabla^2 v^\epsilon[\rho^{\text{solute}}](\mathbf{r}) &= -4\pi[\rho^{\text{solute}}(\mathbf{r}) + \rho^{\text{pol}}(\mathbf{r})] \\ &= -4\pi\rho^{\text{tot}}(\mathbf{r}), \end{aligned} \quad (31)$$

that depends self-consistently on the polarization charge density (and thus on  $v^\epsilon$  itself), where the electrostatic potential  $v^\epsilon$  can be expressed as a vacuum potential depending on both the source and polarization charge densities

$$\begin{aligned} v^\epsilon[\rho^{\text{solute}}](\mathbf{r}) &= v[\rho^{\text{solute}} + \rho^{\text{pol}}](\mathbf{r}) \\ &= v[\rho^{\text{tot}}](\mathbf{r}) = v[\rho^{\text{solute}}](\mathbf{r}) + v[\rho^{\text{pol}}](\mathbf{r}). \end{aligned} \quad (32)$$

From the knowledge of the electrostatic field, one can derive in a straightforward way the Kohn-Sham potential, the electrostatic energy, and the forces acting on the nuclei, as shown in Ref. [13]. In particular, the total electrostatic energy of the system can be separated into two contributions

$$\begin{aligned} E^\epsilon[\rho^{\text{solute}}] &= \frac{1}{2} \int_D \rho^{\text{solute}}(\mathbf{r}) v^\epsilon[\rho^{\text{solute}}](\mathbf{r}) d\mathbf{r} \\ &= \frac{1}{2} \int_D \rho^{\text{solute}}(\mathbf{r}) v[\rho^{\text{solute}}](\mathbf{r}) d\mathbf{r} \\ &\quad + \frac{1}{2} \int_D \rho^{\text{solute}}(\mathbf{r}) v[\rho^{\text{pol}}](\mathbf{r}) d\mathbf{r} \\ &= E[\rho^{\text{solute}}] + E^{\text{pol}}[\rho^{\text{solute}}, \rho^{\text{pol}}], \end{aligned} \quad (33)$$

where we decided to indicate explicitly the dependence of the second contribution on the polarization charge density

$$\begin{aligned} E^{\text{pol}}[\rho^{\text{solute}}, \rho^{\text{pol}}] &= \frac{1}{2} \int_D \rho^{\text{solute}}(\mathbf{r}) v[\rho^{\text{pol}}](\mathbf{r}) d\mathbf{r} \\ &= \frac{1}{2} \int_D \rho^{\text{pol}}(\mathbf{r}) v[\rho^{\text{solute}}](\mathbf{r}) d\mathbf{r}. \end{aligned} \quad (34)$$

For isolated systems, the Poisson equation should be solved together with boundary conditions of vanishing potential at long distances. Nonetheless, most of the approaches proposed in the literature in order to solve Eq. (29) or (31) introduce some approximations on the boundary conditions, in order to simplify or speed up the calculation. In particular, in the original formulation of Fattebert and Gygi [5,6] and in some of its following implementations [7–9], a multigrid method was used to solve for the electrostatic potential, together with an arbitrary homogeneous zeroing of the potential at the boundary of the simulation cell (Dirichlet boundary conditions). In the recently developed SCCS method, instead, an iterative approach has been proposed, coupled with standard FFTs and which relies on periodic boundary conditions.

In particular, one can approximate the isolated potential  $v[\rho^{\text{tot}}](\mathbf{r})$  by the periodic potential  $v[\rho^{\text{tot}}](\mathbf{r})$ , which can be

computed in reciprocal space by exploiting Eq. (11) as

$$v[\varrho^{\text{tot}}](\mathbf{r}) = \sum_{\mathbf{g} \neq 0} \frac{4\pi}{g^2} \tilde{\varrho}^{\text{tot}}(\mathbf{g}) e^{i\mathbf{g}\cdot\mathbf{r}}, \quad (35)$$

where the total charge density  $\varrho^{\text{tot}}(\mathbf{r})$  is also different from the ideal isolated one  $\rho^{\text{tot}}(\mathbf{r})$ , due to its periodicity and of being optimized in the presence of periodic boundary conditions. While the effect of periodicity on the optimization of the nuclear and ionic degrees of freedom of a system can be considered to be negligible [31], periodic boundary conditions enter directly in the definition of the polarization charge density, due to its dependence on the gradient of the electrostatic field

$$\nabla v[\varrho^{\text{tot}}](\mathbf{r}) = \sum_{\mathbf{g}} \frac{4\pi i \mathbf{g}}{g^2} \tilde{\varrho}^{\text{tot}}(\mathbf{g}) e^{i\mathbf{g}\cdot\mathbf{r}}. \quad (36)$$

Moreover, when charged solutes are treated, i.e., when

$$\int_D \varrho^{\text{solute}}(\mathbf{r}) d\mathbf{r} = q^{\text{solute}} \neq 0, \quad (37)$$

the presence of the compensating NCB background should be explicitly accounted for in using the approximation in Eq. (35). The polarization charge in the most general case of a charged system in its periodic approximation is thus given by

$$\begin{aligned} \varrho^{\text{pol}}(\mathbf{r}) = & \frac{1}{4\pi} \nabla \ln \epsilon(\mathbf{r}) \cdot \nabla v[\varrho^{\text{tot}} - \langle \varrho^{\text{tot}} \rangle](\mathbf{r}) \\ & - \frac{\epsilon(\mathbf{r}) - 1}{\epsilon(\mathbf{r})} \varrho^{\text{solute}}(\mathbf{r}) + \frac{\epsilon(\mathbf{r}) - 1}{\epsilon(\mathbf{r})} \langle \varrho^{\text{solute}} \rangle. \end{aligned} \quad (38)$$

Similarly to the case of a polarization charge density in vacuum, the first two terms of  $\varrho^{\text{pol}}$  are localized in the narrow transition region at the boundary of the solute, as explained in Ref. [13]. On the contrary, the last contribution appearing in Eq. (38) is defined everywhere in the simulation cell, except for the vacuum region inside the solute, where  $\epsilon(\mathbf{r}) = 1$ .

It is important to notice that, even though for an isolated charged solute Gauss's law would require the total polarization charge to fulfill the following sum rule:

$$\int_D \rho^{\text{pol}}(\mathbf{r}) d\mathbf{r} = -\frac{\epsilon_0 - 1}{\epsilon_0} q^{\text{solute}}, \quad (39)$$

the total polarization charge of a system in periodic boundary conditions will sum up to zero

$$\int_D \varrho^{\text{pol}}(\mathbf{r}) d\mathbf{r} = 0, \quad (40)$$

due to the PBC-imposed neutrality of the source charge density.

Provided that the full Eq. (38) is used to compute the polarization density, all the equations derived in Ref. [13] apply straightforwardly. For neutral solutes immersed in solvents with high dielectric permittivity and reasonably large cell sizes, the effect of PBC was already shown to be negligible (see Fig. 17 of Ref. [13]). Nonetheless, charged systems immersed in solvents with low dielectric permittivities may present a substantial dependence on the size of the simulation cells.

#### D. Makov-Payne correction in dielectric environments

To summarize the previous discussion, when approximating an isolated system with its periodic counterpart in a quantum-mechanical calculation, one is actually performing two different approximations:

(i) First,

$$\varrho(\mathbf{r}) \neq \rho(\mathbf{r}), \quad (41)$$

i.e., the charge density that one is optimizing with PBC will in general converge to a different final state from the ideal isolated case, due to the interaction with the periodic images and the neutralizing charge background (NCB).

(ii) Second,

$$v[\rho](\mathbf{r}) \neq v[\varrho](\mathbf{r}), \quad (42)$$

i.e., even assuming that we are dealing with a neutral system and that the effects of periodicity on its optimized charge density are negligible, the periodic potential will be different from the isolated case due to the contributions arising from the periodic images and, possibly, due to the different boundary conditions used to solve the problem.

Both approximations will contribute to an error in the calculation of the total energy, i.e.,

$$E[\varrho] \neq E[\rho] \neq E[\rho]. \quad (43)$$

Nonetheless, a simple analytical expression for the leading contributions to the difference between the above energies can be derived in the special case of a cubic simulation cell. The first derivation of such an expression is due to Makov and Payne [31] and provides an approximation of  $E[\rho]$  whose system-size dependence is at worst of the order of  $L^{-5}$ , where  $L$  is the size of the cubic cell. Namely, the exact electrostatic energy of the isolated system can be written in terms of its periodic approximation as

$$\begin{aligned} E^{\text{solute}}[\rho^{\text{solute}}] &= E^{\text{solute}}[\varrho^{\text{solute}} - \langle \varrho^{\text{solute}} \rangle] + \frac{(q^{\text{solute}})^2 \alpha_0}{2L} \\ &- \frac{2\pi}{3L^3} [q^{\text{solute}} Q^{\text{solute}} - (d^{\text{solute}})^2] + O(L^{-5}), \end{aligned} \quad (44)$$

where, with respect to Eq. (15) of Ref. [31], the second contribution has the correct sign and is expressed explicitly in terms of the isolated solute multipole moments. Moreover, the Makov-Payne derivation correctly assumes that charge relaxation due to the artificial periodicity of the system only contributes to the correction of the energy at higher orders. Thus, the multipole moments that enter Eq. (44) are computed from the periodic density in the unit cell without including the eventual NCB density

$$q^{\text{solute}} \approx \int_D \varrho^{\text{solute}}(\mathbf{r}) d\mathbf{r} \equiv q^{\text{solute-NCB}}, \quad (45)$$

$$\mathbf{d}^{\text{solute}} \approx \int_D \varrho^{\text{solute}}(\mathbf{r}) \mathbf{r} d\mathbf{r} \equiv \mathbf{d}^{\text{solute-NCB}}, \quad (46)$$

$$Q^{\text{solute}} \approx \int_D \varrho^{\text{solute}}(\mathbf{r}) r^2 d\mathbf{r} \equiv Q^{\text{solute-NCB}}. \quad (47)$$

The first contribution in Eq. (44) is due to the interaction energy of the NCB-neutralized monopole moment in the periodic system interacting with its replicas and is easily expressed in terms of the Madelung constant of a cubic lattice  $\alpha_0$ . Dipole-dipole and quadrupole-monopole interactions with periodic replicas are canceled in the periodic energy due to the cubic symmetry of the lattice, while the contributions due to quadrupole-quadrupole and higher multipoles interactions decay at worst as  $L^{-5}$ . The origin of the second contribution in Eq. (44) is due to the tin-foil boundary conditions that are implicitly assumed in a periodic boundary calculation. These boundary conditions artificially impose that the average electrostatic field and potential in the cell are zero. As a consequence, the interaction energies of the multipole moments of the system with themselves (specifically the dipole-dipole and the monopole-quadrupole interactions) are modified with respect to the isolated case due to these arbitrary shifts. In particular, the energy due to the dipole moment

$$E[\mathbf{d}^{\text{solute}}] = -\frac{1}{2}\mathbf{d}^{\text{solute}} \cdot \mathbf{\Xi}(0) \quad (48)$$

lacks the contribution

$$\begin{aligned} \Delta E &= E[\mathbf{d}^{\text{solute}}] - E[\mathbf{d}^{\text{solute}}] \\ &= -\frac{1}{2}\mathbf{d}^{\text{solute}} \cdot \langle \mathbf{E} \rangle = -\frac{1}{2}\mathbf{d}^{\text{solute}} \cdot \frac{4\pi}{3} \frac{\mathbf{d}^{\text{solute}}}{L^3}. \end{aligned} \quad (49)$$

Similarly, the energy due to the monopole-quadrupole interaction

$$E = q^{\text{solute}} v^{\text{Q}^{\text{solute}}}(0) \quad (50)$$

has to be corrected due to the shift of the potential with respect to the ideal isolated system in vacuum, namely,

$$E^{mq,corr} = q^{\text{solute}} \langle v^{\text{Q}^{\text{solute}}} \rangle = q^{\text{solute}} \frac{2\pi}{3} \frac{Q^{\text{solute}}}{L^3}. \quad (51)$$

In the above equations, we used the fact that, as shown in Ref. [28] and reported in Eqs. (25) and (26), only the dipole and quadrupole moments contribute to the average values of the electrostatic field and potential, respectively, i.e.,

$$\langle \mathbf{\Xi}^{\rho^{\text{solute}}} \rangle = \langle \mathbf{\Xi}^{\mathbf{d}^{\text{solute}}} \rangle = \mathbf{\Xi}_0, \quad (52)$$

$$\langle v^{\rho^{\text{solute}}} \rangle = \langle v^{\text{Q}^{\text{solute}}} \rangle = v_0. \quad (53)$$

Makov and Payne also derived a simplified expression for a system in a condensed phase by adopting the approach of Leslie and Gillan and rescaling the potential by the dielectric constant  $\epsilon_0$  of the system. The result

$$\begin{aligned} E^{\text{solute}}[\rho^{\text{solute}}] &= E^{\text{solute}}[q^{\text{solute}}] + \frac{(q^{\text{solute}})^2 \alpha_0}{2L\epsilon_0} \\ &\quad - \frac{2\pi}{3L^3\epsilon_0} [q^{\text{solute}} Q^{\text{solute}} - (d^{\text{solute}})^2] \\ &\quad + O(L^{-5}) \end{aligned} \quad (54)$$

assumes a uniform homogeneous dielectric everywhere in space. Such an assumption does not take into account the variations of the dielectric constant in the different regions of the system and, in particular, is not correct for the SCCS model, where a solute is immersed in a medium whose dielectric constant varies from one (vacuum) to the bulk dielectric

constant of the solvent. Nonetheless, an approach similar to the one of Makov and Payne can be used to derive the correction to the electrostatic energy in the SCCS framework up to terms of the order of  $L^{-3}$ .

Contrary to what is generally assumed for the energy contribution due to the polarization of the solute charge density due to periodic images, the periodic solution of the polarization charge density has a significant effect on the polarization contribution to the electrostatic energy of a solvated system. The problem is twofold, and is partly related to the fact that the neutralizing background induces a small polarization which is diffused all over the unit cell, and partly due to the fact that periodic images can induce a non-negligible polarization charge density in the region of space close to the solute charge density. These spurious polarization charges affect the multipole moments of the polarization charge

$$q^{\text{pol}} \neq q^{\text{pol}}, \quad (55)$$

$$\mathbf{d}^{\text{pol}} \neq \mathbf{d}^{\text{pol}}, \quad (56)$$

$$Q^{\text{pol}} \neq Q^{\text{pol}} \quad (57)$$

and need to be taken care of explicitly, before a scheme analogous to the one of Makov and Payne can be adopted.

In particular, the difference  $\Delta\rho(\mathbf{r})$  between the periodic [Eq. (38)] and the isolated [Eq. (30)] polarization charges can be written as

$$\begin{aligned} \Delta\rho^{\text{pol}}(\mathbf{r}) &\equiv \rho^{\text{pol}}(\mathbf{r}) - \rho^{\text{pol}}(\mathbf{r}) \\ &= \frac{1}{4\pi} \nabla \ln \epsilon(\mathbf{r}) \cdot [\nabla v[\rho^{\text{tot}} - \langle \rho^{\text{tot}} \rangle](\mathbf{r}) \\ &\quad - \nabla v[\rho^{\text{tot}}](\mathbf{r})] + \frac{\epsilon(\mathbf{r}) - 1}{\epsilon(\mathbf{r})} \langle \rho^{\text{solute}} \rangle \\ &\approx \frac{1}{4\pi} \nabla \ln \epsilon(\mathbf{r}) \cdot \nabla v[\Delta\rho^{\text{pol}}](\mathbf{r}) + \frac{\epsilon(\mathbf{r}) - 1}{\epsilon(\mathbf{r})} \langle \rho^{\text{solute}} \rangle \\ &\quad + \frac{1}{4\pi} \nabla \ln \epsilon(\mathbf{r}) \cdot \nabla \Delta v[\rho^{\text{tot}}](\mathbf{r}). \end{aligned} \quad (58)$$

The corrective potential  $\Delta v[\rho](\mathbf{r})$  is defined following Dabo *et al.* [40,41], i.e., as the difference between the ideal isolated potential in vacuum and its periodic counterpart computed using tin-foil boundary conditions

$$\begin{aligned} \Delta v[\rho](\mathbf{r}) &= v[\rho - \langle \rho \rangle](\mathbf{r}) - v[\rho](\mathbf{r}) \\ &\approx v[\rho - \langle \rho \rangle](\mathbf{r}) - v[\rho](\mathbf{r}). \end{aligned} \quad (59)$$

The correction to the polarization charge can be computed iteratively with the same approach used for the periodic polarization charge, provided that an expression for the corrective potential is available. The last two contributions to the polarization in Eq. (58) do not change during the iteration cycles, thus they can be considered as two separate sources and the corrective polarization can be separated into two contributions, one due to the NCB density and the other due to the corrective potential. Namely,

$$\begin{aligned} \Delta\rho^{\text{pol,NCB}}(\mathbf{r}) &= \frac{1}{4\pi} \nabla \ln \epsilon(\mathbf{r}) \cdot \nabla v[\Delta\rho^{\text{pol,NCB}}](\mathbf{r}) \\ &\quad + \frac{\epsilon(\mathbf{r}) - 1}{\epsilon(\mathbf{r})} \langle \rho^{\text{solute}} \rangle \end{aligned} \quad (60)$$

and

$$\begin{aligned} \Delta\rho^{\text{pol,periodic}}(\mathbf{r}) &= \frac{1}{4\pi} \nabla \ln \epsilon(\mathbf{r}) \cdot \nabla v[\Delta\rho^{\text{pol,periodic}}](\mathbf{r}) \\ &+ \frac{1}{4\pi} \nabla \ln \epsilon(\mathbf{r}) \cdot \nabla \Delta v[\rho^{\text{tot}}](\mathbf{r}). \end{aligned} \quad (61)$$

By exploiting the derivation of Refs. [40,41] for the point-charge approximation of the corrective potential (see following section), the gradient in the second term of the difference between periodic and isolated polarization can be approximated as

$$\nabla \Delta v[\rho](\mathbf{r}) \approx \frac{4\pi}{3L^3} (\mathbf{d} - \mathbf{r}). \quad (62)$$

It is important to note that the above approximation is correct only close to the origin of the system charge distribution and it becomes more and more accurate as the cell size increases. Both the periodic and the NCB contribution to the corrective polarization charge are proportional to  $L^{-3}$ . While the periodic polarization is defined only in the small region around the solute, where the dielectric is allowed to vary, the NCB polarization is defined everywhere in space. Nonetheless, its value in the bulk of the solvent is constant and given by

$$\Delta\rho^{\text{pol,NCB,bulk}} = \frac{(\epsilon_0 - 1)}{\epsilon_0} \langle \rho^{\text{solute}} \rangle. \quad (63)$$

The bulk constant charge can be removed from the corrective polarization so that

$$\begin{aligned} \Delta\rho^{\text{pol,NCB,confined}} &= \frac{1}{4\pi} \nabla \ln \epsilon(\mathbf{r}) \cdot \nabla v[\Delta\rho^{\text{pol,NCB}}](\mathbf{r}) \\ &+ \left( \frac{1}{\epsilon_0} - \frac{1}{\epsilon(\mathbf{r})} \right) \frac{q^{\text{solute}}}{L^3} \end{aligned} \quad (64)$$

is a quantity confined in a well-defined region of space, which does not depend on cell size. With this choice, the energy contributions due to the corrective polarizations are

$$\begin{aligned} &\int_{\text{D}} \Delta\rho^{\text{pol}}(\mathbf{r}) v[\rho^{\text{solute}} - \langle \rho^{\text{solute}} \rangle](\mathbf{r}) d\mathbf{r} \\ &= \int_{\text{D}} \Delta\rho^{\text{pol,periodic}}(\mathbf{r}) v[\rho^{\text{solute}} - \langle \rho^{\text{solute}} \rangle](\mathbf{r}) d\mathbf{r} \\ &+ \int_{\text{D}} \Delta\rho^{\text{pol,NCB,confined}}(\mathbf{r}) v[\rho^{\text{solute}} - \langle \rho^{\text{solute}} \rangle](\mathbf{r}) d\mathbf{r} \\ &+ \int_{\text{D}} \Delta\rho^{\text{pol,NCB,bulk}}(\mathbf{r}) v[\rho^{\text{solute}} - \langle \rho^{\text{solute}} \rangle](\mathbf{r}) d\mathbf{r} \\ &= \int_{\text{D}} \Delta\rho^{\text{pol,periodic}}(\mathbf{r}) v[\rho^{\text{solute}} - \langle \rho^{\text{solute}} \rangle](\mathbf{r}) d\mathbf{r} \\ &+ \int_{\text{D}} \Delta\rho^{\text{pol,NCB,confined}}(\mathbf{r}) v[\rho^{\text{solute}} - \langle \rho^{\text{solute}} \rangle](\mathbf{r}) d\mathbf{r}, \end{aligned} \quad (65)$$

where the NCB bulk contribution vanished, as a constant charge density does not contribute to the periodic energy. Both corrective contributions in Eq. (65) will scale as  $L^{-3}$  since the charge densities are confined in a region of space which is not dependent on the cell size. Thus, when trying to remove the system's size dependence from the calculation, both terms should be subtracted from the periodic polarization energy

computed in the SCCS framework

$$E[\rho^{\text{solute}}, \rho^{\text{pol}}] = E[\rho^{\text{solute}}, \rho^{\text{pol}}] - E^{\text{pol}}[\rho^{\text{solute}}, \Delta\rho^{\text{pol}}]. \quad (66)$$

Eventually, we are left with the periodic energy of the solute in vacuum  $E^{\text{solute}}[\rho^{\text{solute}}]$  (whose ideally isolated counterpart  $E^{\text{solute}}[\rho^{\text{solute}}]$  can be recovered through the Makov-Payne expression) plus the periodic energy of interaction between the solute charge density and a polarization optimized as if the system were isolated:

$$\begin{aligned} E^{\text{pol}}[\rho^{\text{solute}}, \rho^{\text{pol}}] &= \frac{1}{2} \int \rho^{\text{pol}}(\mathbf{r}) v[\rho^{\text{solute}}](\mathbf{r}) d\mathbf{r} \\ &= \frac{1}{2} \int \rho^{\text{solute}}(\mathbf{r}) v[\rho^{\text{pol}}](\mathbf{r}) d\mathbf{r}. \end{aligned} \quad (67)$$

A Makov-Payne-type expression for this latter term can be derived by assuming, as in Ref. [31], that  $\rho^{\text{solute}}(\mathbf{r}) = \rho^{\text{solute}}(\mathbf{r})$  inside the unit cell and by considering the Makov-Payne corrections for the electrostatic energy of the system composed by the total charge density

$$\rho^{\text{tot}}(\mathbf{r}) = \rho^{\text{sol}}(\mathbf{r}) + \rho^{\text{pol}}(\mathbf{r}) \quad (68)$$

and the one of a system composed solely by the polarization charge. Namely, by rewriting Eq. (44) for the total charge distribution and by performing some simple algebraic manipulation, one obtains

$$\begin{aligned} &E^{\text{pol}}[\rho^{\text{solute}}, \rho^{\text{pol}}] \\ &= \frac{1}{2} (E^{\text{solute}}[\rho^{\text{tot}}] - E^{\text{solute}}[\rho^{\text{solute}}] - E^{\text{solute}}[\rho^{\text{pol}}]) \\ &= E^{\text{pol}}[\rho^{\text{solute}}, \rho^{\text{pol}}] + \frac{(q^{\text{solute}})^2 \alpha_0}{2L} \left( -1 + \frac{1}{\epsilon_0} \right) \\ &- \frac{\pi q^{\text{solute}}}{3L^3} \left( -\frac{\epsilon_0 - 1}{\epsilon_0} Q^{\text{solute}} + Q^{\text{pol}} \right) \\ &+ \frac{2\pi}{3L^3} (\mathbf{d}^{\text{solute}} \cdot \mathbf{d}^{\text{pol}}) + O(L^{-5}), \end{aligned} \quad (69)$$

where the relation

$$q^{\text{pol}} = -\frac{\epsilon_0 - 1}{\epsilon_0} q^{\text{solute}} \quad (70)$$

has been exploited between the total polarization charge in the isolated system and the solute charge. When summing the correction to the polarization energy to the one of the electrostatic energy of the system in vacuum [Eq. (44)], the final expression for the energy of the solvated system becomes

$$\begin{aligned} E[\rho^{\text{solute}}, \rho^{\text{pol}}] &= E[\rho^{\text{solute}}, \rho^{\text{pol}}] - E^{\text{pol}}[\rho^{\text{solute}}, \Delta\rho^{\text{pol}}] \\ &+ \left( \frac{1}{\epsilon_0} \right) \frac{(q^{\text{solute}})^2 \alpha_0}{2L} \\ &- \frac{2\pi q^{\text{solute}}}{3L^3} \left( \frac{Q^{\text{solute}}}{2\epsilon_0} + \frac{Q^{\text{solute}} + Q^{\text{pol}}}{2} \right) \\ &+ \frac{2\pi}{3L^3} [(d^{\text{solute}})^2 + \mathbf{d}^{\text{solute}} \cdot \mathbf{d}^{\text{pol}}] + O(L^{-5}). \end{aligned} \quad (71)$$

Compared to the result derived by Makov and Payne for aperiodic systems in a condensed phase, we see that the



monopole contribution is identical, reflecting the fact that it is an interaction energy between systems in neighboring cells and is thus exactly rescaled by the presence of the dielectric continuum. On the other hand, a more complex expression for the other terms has now been obtained in Eq. (71), and this is one of the main results of this paper.

If we consider the simple case of a uniform dielectric extending over the whole space, the polarization charge would be simply given by

$$\rho^{\text{pol}}(\mathbf{r}) = -\frac{\epsilon_0 - 1}{\epsilon_0} \rho^{\text{solute}}(\mathbf{r}), \quad (72)$$

which translates into dipole and quadrupole moments

$$\mathbf{d}^{\text{pol}} = -\frac{\epsilon_0 - 1}{\epsilon_0} \mathbf{d}^{\text{solute}}, \quad (73)$$

$$Q^{\text{pol}} = -\frac{\epsilon_0 - 1}{\epsilon_0} Q^{\text{solute}}, \quad (74)$$

which, inserted in Eq. (71), give back the result proposed by Makov and Payne [Eq. (54)]. In the case of a dipolar solute in a spherical cavity, the polarization dipole is analytically obtained from the Onsager model as

$$\mathbf{d}^{\text{pol,Onsager}} = -\frac{2(\epsilon_0 - 1)}{2\epsilon_0 + 1} \mathbf{d}^{\text{solute}}, \quad (75)$$

which gives a term proportional to

$$\frac{2\pi}{(2\epsilon_0 + 1)L^3} (d^{\text{solute}})^2, \quad (76)$$

consistent with the expression of the extrinsic field of a periodic system immersed in a dielectric  $\mathbf{E}_0^\epsilon$  in Eq. (28). In general, for arbitrary, molecular-shaped cavities, the dipole and quadrupole contributions are not analytic functions of the solute multipole moments and need to be computed explicitly from the integral of the polarization charge density.

### E. Countercharge corrections in dielectric environments

Several schemes have been proposed along the lines of the Makov-Payne correction, but which self-consistently correct the electrostatic potential rather than just the final electrostatic energy. The general framework is to recover the electrostatic potential of the isolated system by adding to the periodic boundary potential a corrective term. The correction can then be analytically computed for specific approximations on the charge distribution of the system, or the exact problem can be solved numerically via multigrid techniques. The different schemes have been recently classified into three categories, depending on the different level of approximations used to treat the charge density of the system: following Refs. [40,41], they are labeled as point-countercharge (PCC), Gaussian-countercharge (GCC), and density-countercharge (DCC) methods. Here, we will discuss the PCC correction scheme. For a pointlike unit charge in a cubic cell, the corrective potential

$$\Delta v[\rho](\mathbf{r}) = \frac{\alpha_0}{L} - \frac{2\pi}{3L^3} r^2 + O(|\mathbf{r}|^4) \quad (77)$$

can be recovered by exploiting symmetry and the Poisson Eq. (4), as shown in Refs. [40,41]. The resulting parabolic potential

is accurate only close to where the charge is located. For an arbitrary charge distribution, the corrective potential can be expressed in terms of the corrective potential of a collection of point charges that matches the system's multipole moments. Due to the quadratic nature of the correction, only the potential generated by multipoles up to the quadrupole can be corrected. The final PCC expression for the corrective potential reads as

$$\Delta v[\rho](\mathbf{r}) = \frac{\alpha_0}{L} q - \frac{2\pi q}{3L^3} r^2 + \frac{4\pi}{3L^3} \mathbf{d} \cdot \mathbf{r} - \frac{2\pi}{3L^3} Q. \quad (78)$$

The correction to the energy

$$\begin{aligned} \Delta E &= E^{\text{pol}}[\rho] - E[\rho] \\ &= \frac{1}{2} \int_{\text{D}} \rho \Delta v[\rho](\mathbf{r}) \\ &= \frac{\alpha_0}{2L} q^2 - \frac{2\pi}{3L^3} (qQ - d^2) \end{aligned} \quad (79)$$

reduces correctly to the Makov-Payne expression, with the only difference that the molecular charge distribution is now optimized in the presence of the corrected potential, i.e., the approximations in Eqs. (45), (46), and (47) are not needed.

When a continuum dielectric is present in the system, the electrostatic energy is that of Eq. (33), and the potential that needs to be corrected is the one arising from the total charge distribution

$$v^\epsilon[\rho^{\text{solute}}](\mathbf{r}) = v[\rho^{\text{solute}} + \rho^{\text{pol}}](\mathbf{r}) = v[\rho^{\text{tot}}](\mathbf{r}), \quad (80)$$

including the polarization charge. This means that Eq. (78) can be simply modified as

$$\begin{aligned} \Delta v[\rho^{\text{tot}}](\mathbf{r}) &= \frac{\alpha_0}{L} q^{\text{tot}} - \frac{2\pi q^{\text{tot}}}{3L^3} r^2 \\ &\quad + \frac{4\pi}{3L^3} \mathbf{d}^{\text{tot}} \cdot \mathbf{r} - \frac{2\pi}{3L^3} Q^{\text{tot}}, \end{aligned} \quad (81)$$

where

$$q^{\text{tot}} = q^{\text{sol}} + q^{\text{pol}} = \frac{q^{\text{sol}}}{\epsilon_0}, \quad (82)$$

$$\mathbf{d}^{\text{tot}} = \mathbf{d}^{\text{sol}} + \mathbf{d}^{\text{pol}}, \quad (83)$$

$$Q^{\text{tot}} = Q^{\text{sol}} + Q^{\text{pol}}. \quad (84)$$

Again, when computing the correction to the electrostatic energy of the system, the PCC approach gives the same result obtained with the Makov-Payne scheme:

$$\begin{aligned} \Delta E &= \frac{1}{2} \int \rho^{\text{solute}} \Delta v[\rho^{\text{tot}}](\mathbf{r}) d\mathbf{r} \\ &= \frac{\alpha_0}{2L} \frac{q^{\text{solute}}}{\epsilon_0} - \frac{2\pi}{3L^3} \left[ q^{\text{solute}} \left( \frac{Q^{\text{tot}}}{2} + \frac{Q^{\text{solute}}}{2\epsilon_0} \right) \right. \\ &\quad \left. - (\mathbf{d}^{\text{tot}}) \cdot \mathbf{d}^{\text{solute}} \right] \end{aligned} \quad (85)$$

apart from the fact that the solute charge density is now optimized in the presence of the correct boundary conditions. From the above expressions, it is straightforward to derive the

correction to the interatomic forces, namely,

$$\begin{aligned}\Delta \mathbf{f}_a &= -\frac{d\Delta E}{d\mathbf{R}_a} = z_a \nabla \Delta v[\rho^{\text{tot}}](\mathbf{r})|_{\mathbf{r}=\mathbf{R}_a} \\ &= z_a \frac{4\pi}{3L^3} (-q^{\text{tot}} \mathbf{R}_a + \mathbf{d}^{\text{tot}}),\end{aligned}\quad (86)$$

where we have used the Hellmann-Feynman theorem, following the derivation reported in Sec. III C of Ref. [13], and for the solute charge density we have used

$$\rho^{\text{solute}} = \rho^{\text{elec}} + \sum_a z_a \delta(\mathbf{r} - \mathbf{R}_a), \quad (87)$$

where the nuclei are represented as pointlike charges.

As we are now using the correct potential in the derivation of the polarization charges, no NCB polarization and no periodic polarization appear in the polarization charge. Similarly, provided that the potential is correct up to the region where the dielectric medium becomes uniform, the total polarization charge will sum up to the correct value for an isolated system. As summarized in Appendix, special care needs to be taken in the way nuclear charges are treated when computing the polarization charge and PCC periodic boundary corrections.

A similar approach can be derived for systems of different periodicity, where in particular the exact expression of the corrective potential can be obtained analytically via partial Fourier transforms. A particularly important case is the one of two-dimensional systems, for which a solution involving two-dimensional Fourier transform was derived in Refs. [35,42]. Analogously to what was done with PCC, a simple approximated analytical solution can be devised for the case where the cell size is large enough compared to the size of the system. In this case, only the component for  $\mathbf{g} = 0$  contributes significantly to the corrective potential, that acquires a quadratic form analogous to the PCC results reported for the isolated system in cubic cells. Namely, the expression for the corrective potential of a 2D system is

$$\begin{aligned}\Delta v^{2D}[\rho^{\text{tot}}](\mathbf{r}) &= \Delta v^{2D}[\rho^{\text{tot}}](r_z) \\ &= \frac{\alpha_{1D}}{L_z} q^{\text{tot}} - \frac{2\pi q^{\text{tot}}}{AL_z} r_z^2 + \frac{4\pi}{AL_z} d_z^{\text{tot}} \cdot r_z \\ &\quad - \frac{2\pi}{AL_z} Q_{zz}^{\text{tot}},\end{aligned}\quad (88)$$

where  $\alpha_{1D} = \pi/3$  is the Madelung constant of a one-dimensional periodic array of charges,  $A$  is the area of the slab, while  $L_z$  is the size of the cell axis perpendicular to the plane of the slab. The correction to the energy is readily obtained by integration with the system charge density; namely, for a system in vacuum

$$\Delta E^{2D} = \frac{\alpha_{1D}}{2L_z} (q^{\text{solute}})^2 - \frac{2\pi}{AL_z} [q^{\text{solute}} Q_{zz}^{\text{solute}} - (d_z^{\text{solute}})^2]. \quad (89)$$

Similarly to what was derived for isolated systems, also for slabs the effects of the solvent can be immediately included by defining the corrective potentials in terms of the total dipole moment of the system, thus including the contribution of the

polarization density

$$\begin{aligned}\Delta E^{2D} &= \frac{\alpha_{1D}}{2L_z} \frac{(q^{\text{solute}})^2}{\epsilon_0} \\ &\quad - \frac{2\pi}{AL_z} \left[ q^{\text{solute}} \left( \frac{Q_{zz}^{\text{solute}} + Q_{zz}^{\text{pol}}}{2} + \frac{Q_{zz}^{\text{solute}}}{2\epsilon_0} \right) \right. \\ &\quad \left. - (d_z^{\text{solute}} + d_z^{\text{pol}}) d_z^{\text{solute}} \right].\end{aligned}\quad (90)$$

For the periodic boundary correction contribution to interatomic forces, an expression similar to the one derived for the OD case applies, namely,

$$\begin{aligned}\Delta f_{a,z} &= -\frac{d\Delta E^{2D}}{dR_{a,z}} = z_a \frac{d}{dr_z} \Delta v[\rho^{\text{tot}}](r_z)|_{r_z=R_{a,z}} \\ &= z_a \frac{4\pi}{AL_z} (-q R_{a,z} + d_z).\end{aligned}\quad (91)$$

## F. Martyna-Tuckerman corrections in a dielectric environment

While the approaches derived above aim at correcting the periodic potential by introducing a real-space potential computed *a posteriori*, a different approach has been developed in the literature, which aims to correct directly the periodic potential as computed in reciprocal space. Such an approach, which has its foundation in the screening function formalisms and was pioneered for PBC correction by Martyna and Tuckerman [33,35,36], has received a lot of attention in recent years due to its simple implementation and very reduced computational cost.

The main idea behind the approach of Martyna and Tuckerman (MT) [33] and similar approaches [42] is the following: When the electrostatic problem is solved in reciprocal space, the use of the Fourier transform of the differential operator [Eq. (13)] implies that the potential is obtained from the periodic sum of the real-space potentials. In other words, the analytic Fourier transform of the aperiodic Green's function  $G(\mathbf{r} - \mathbf{r}')$  corresponds to the reciprocal-space coefficients of the periodic Green's function  $\mathbf{G}(\mathbf{g})$ , namely,

$$\begin{aligned}\tilde{\mathbf{G}}(\mathbf{g}) &= \int_{\infty} G(\mathbf{r}) e^{-i\mathbf{g}\cdot\mathbf{r}} d\mathbf{r} = \int_{\mathcal{D}} \sum_{\mathbf{R}} G(\mathbf{r} + \mathbf{R}) e^{-i\mathbf{g}\cdot\mathbf{r}} d\mathbf{r} \\ &= \int_{\mathcal{D}} G(\mathbf{r}) e^{-i\mathbf{g}\cdot\mathbf{r}} d\mathbf{r} = \mathbf{G}(\mathbf{g}).\end{aligned}\quad (92)$$

In particular, for the case of an isolated system in vacuum [Eq. (8)] we have

$$\tilde{\mathbf{G}}(\mathbf{g}) = \mathbf{G}(\mathbf{g}) = \frac{1}{g^2}. \quad (93)$$

If one, instead, were to use the Fourier series coefficients of the potential kernel

$$G(\mathbf{g}) = \int_{\mathcal{D}} \frac{e^{-i\mathbf{g}\cdot\mathbf{r}}}{|\mathbf{r}|} d\mathbf{r}, \quad (94)$$

one could build the first-image form of the potential

$$\hat{\mathbf{G}}(\mathbf{r}) = \sum_{\mathbf{g}} G(\mathbf{g}) e^{-i\mathbf{g}\cdot\mathbf{r}}, \quad (95)$$

i.e., a periodically repeated approximation of the isolated Green's function  $G(\mathbf{r})$ . The periodicity which is introduced by using the Fourier series and the first-image form only affects the potential at the boundary of the simulation cell. For this reason, in short-ranged functions which decay well within the boundaries of the unit cell, the first-image form, the true potential, and the periodic sum are identical in the region of interest, i.e.,

$$\hat{G}^{\text{short}}(\mathbf{r}) = \mathfrak{G}^{\text{short}}(\mathbf{r}) = G^{\text{short}}(\mathbf{r}) \quad \text{for } \mathbf{r} \in D. \quad (96)$$

For long-ranged functions, as is the case for the Coulomb potential, it is generally accepted that if a cell twice as large as the system studied is used, the first-image form is a good approximation of the correct potential in all the relevant domain, where the quantum charge density is different from zero.

In order to make the algorithm more stable and readily compatible with PBC, the Fourier series coefficients of the potential can be written in an auxiliary-function formalism, i.e., as  $\mathbf{g}$ -dependent coefficients which correct the analytical Fourier-transform coefficients:

$$G(\mathbf{g}) = \tilde{G}(\mathbf{g}) + G(\mathbf{g}) - \tilde{G}(\mathbf{g}) = \tilde{G}(\mathbf{g}) + \Delta G(\mathbf{g}). \quad (97)$$

For short-ranged potentials, for the reasons discussed above, one has that  $\Delta G^{\text{short}}(\mathbf{g}) = 0$ . On the other hand, for the Coulomb potential one needs to compute the long-range correction in reciprocal space numerically using fast Fourier transforms. Once the values of  $\Delta G(\mathbf{g})$  are known, the first-image form of the electrostatic potential of the system can be easily obtained as

$$\begin{aligned} \hat{v}[\rho](\mathbf{r}) &= \sum_{\mathbf{g}} \frac{4\pi}{V} \rho(\mathbf{g}) \left( \frac{1 - \delta_{\mathbf{g}\mathbf{0}}}{g^2} + \Delta G(\mathbf{g}) \right) e^{i\mathbf{g}\cdot\mathbf{r}} \\ &= v[\rho](\mathbf{r}) + \Delta \hat{v}[\rho](\mathbf{r}), \end{aligned} \quad (98)$$

where the Kronecker  $\delta_{\mathbf{g}\mathbf{0}}$  is 1 for  $\mathbf{g} = \mathbf{0}$  and 0 otherwise, where

$$\Delta G(\mathbf{0}) = \lim_{\mathbf{g} \rightarrow \mathbf{0}} \left( \tilde{G}(\mathbf{g}) - \frac{1}{g^2} \right), \quad (99)$$

and where the corrective potential is now computed in reciprocal space as

$$\Delta \hat{v}[\rho](\mathbf{r}) = \sum_{\mathbf{g}} \frac{4\pi}{V} \rho(\mathbf{g}) \Delta G(\mathbf{g}) e^{i\mathbf{g}\cdot\mathbf{r}}. \quad (100)$$

The coefficients entering the calculation of the potential are only dependent on the geometry of the cell and on the type of potential that is computed (depending, e.g., whether the whole Coulomb potential is computed or just its long-range part). Thus, these coefficients can be computed once and for all at the beginning of a calculation and the overall computational cost of the procedure becomes negligible. This is not the case for real-space approaches, where a new potential needs to be computed during the SCF cycle following the change in the multipole moments of the charge distribution, even though it is usually not necessary to update it at each SCF step [40,41]. On the other hand, real-space approaches are in principle able to provide a good approximation of the exact potential profile over the entire cell and can usually adopt smaller cell sizes

compared to MT approaches, where the imposed periodicity can significantly alter the potential at the cell boundaries.

Reciprocal-space approaches are particularly suited for calculations in the presence of a continuum dielectric medium. In particular, by extending the use of the auxiliary-function coefficients computed for the potential also to the calculation of the gradient of the potential

$$\nabla \hat{v}[\rho](\mathbf{r}) = \sum_{\mathbf{g}} \frac{4\pi}{V} \rho(\mathbf{g}) i\mathbf{g} \left( \frac{1 - \delta_{\mathbf{g}\mathbf{0}}}{g^2} + \Delta G(\mathbf{g}) \right) e^{i\mathbf{g}\cdot\mathbf{r}}, \quad (101)$$

it is straightforward to compute the ideal polarization charge by using Eq. (30). All the standard SCCS equations can then be used straightforwardly, but particular care has to be given to the calculation of the forces. Indeed, the interatomic forces can be computed from the Hellman-Feynman theorem as

$$\mathbf{f}_a = - \frac{dE[\rho^{\text{solute}}]}{d\mathbf{R}_a} = - \int \rho^{\text{solute}} \frac{\partial v[\rho^{\text{ions}}](\mathbf{r})}{\partial \mathbf{R}_a} d\mathbf{r}, \quad (102)$$

while the Martyna-Tuckerman correction to the potential introduces the following contribution:

$$\begin{aligned} \Delta \mathbf{f}_a &= - \int \rho^{\text{solute}} \frac{\partial \Delta \hat{v}[\rho^{\text{ions}}](\mathbf{r})}{\partial \mathbf{R}_a} d\mathbf{r} \\ &= - \sum_{\mathbf{g}} \rho^{\text{solute}}(\mathbf{g}) \frac{\partial}{\partial \mathbf{R}_a} \left[ \frac{4\pi}{V} \left( \sum_b^{N_{\text{ions}}} z_b e^{i\mathbf{g}\cdot\mathbf{R}_b} \right) \Delta G(\mathbf{g}) \right] \\ &= - \frac{4\pi z_a}{V} \sum_{\mathbf{g}} i\mathbf{g} \rho^{\text{solute}}(\mathbf{g}) e^{i\mathbf{g}\cdot\mathbf{R}_a} \Delta G(\mathbf{g}). \end{aligned} \quad (103)$$

Similarly, since the forces in the SCCS framework require an additional term due to the solvent polarization density

$$f_{a,i}^{\text{pol}} = - \frac{\partial E^{\text{pol}}[\rho^{\text{solute}}, \rho^{\text{pol}}]}{\partial R_{a,i}} = - \int \rho^{\text{pol}} \frac{\partial \Delta \hat{v}[\rho^{\text{ions}}](\mathbf{r})}{\partial R_{a,i}} d\mathbf{r}, \quad (104)$$

the proper MT correction needs to be included:

$$\Delta \mathbf{f}_a^{\text{pol}} = - \frac{4\pi z_a}{V} \sum_{\mathbf{g}} i\mathbf{g} \rho^{\text{pol}}(\mathbf{g}) e^{i\mathbf{g}\cdot\mathbf{R}_a} \Delta G(\mathbf{g}). \quad (105)$$

### III. RESULTS

#### A. Numerical details

The methods reported in the previous section have been implemented in a development version of the open-source QUANTUM ESPRESSO distribution [51].

Calculations of 0D systems are performed on a pyridine molecule, using the local-density approximation (LDA) of DFT with a wave-function cutoff of 50 Ry. The Brillouin zone is sampled only at the gamma point. Norm-conserving pseudopotentials from the 0.2.2 version of the library of Dal Corso *et al.* [52] are adopted.

The analysis of 2D systems is performed for a CO molecule adsorbed in the atop geometry on a fcc Pt (111) surface. In order to speed up the computational cost for the calculation, a simplified structure is adopted for the slab, composed by only two layers of metal atoms in a  $\sqrt{3} \times 2$  supercell, with a lattice

constant of 2.828 Å. Marzari-Vanderbilt [53] cold smearing with a smearing width of 0.03 Ry is used, the Brillouin zone is sampled with a shifted  $4 \times 4 \times 1$  reciprocal-space integration grid, ultrasoft pseudopotentials and the LDA are adopted, with wave-function and density cutoffs of 35 and 280 Ry, respectively.

For both 0D and 2D systems, the accuracy of the methods is tested by comparing the Hellmann-Feynman forces against the ones obtained by finite differences of the energy (a displacement step of 0.01 a.u. is adequate for all the systems studied).

### B. Isolated (0D) systems

In Figs. 1 and 2, we report the behavior of the energy of a pyridine cation as a function of the cell size, for a system in periodic boundary conditions and for the different correction schemes presented in the sections above, both without (Fig. 1) and with (Fig. 2) a continuum solvent as described by the SCCS method. As expected, both for the molecule in vacuum and for the solvated one, the periodic energy decays as the inverse power of the cell size, with a much less marked dependence for the solvated cases, due to the dielectric which screens the total charge and dipole moment of the system. Corrected results from the different methods are converged for cell sizes of 23 (27) a.u. for the vacuum (solvent) case, reflecting the larger size of the solvated system. The MP and PCC energies are found to converge as fast as  $L^{-5}$  (as seen in the log-log plot of the residual error, see Fig. 4), while the Martyna-Tuckerman energies become constant and exact for cell sizes larger than 30 a.u. On the other hand, while for small cells the energies computed with the real-space approaches and the Makov-Payne are still less than 1 mRy away from the converged results, the Martyna-Tuckerman approach shows significant errors, even exceeding the uncorrected periodic energy. The same trend is reflected in the calculation of the electrostatic contribution to solvation free energies.

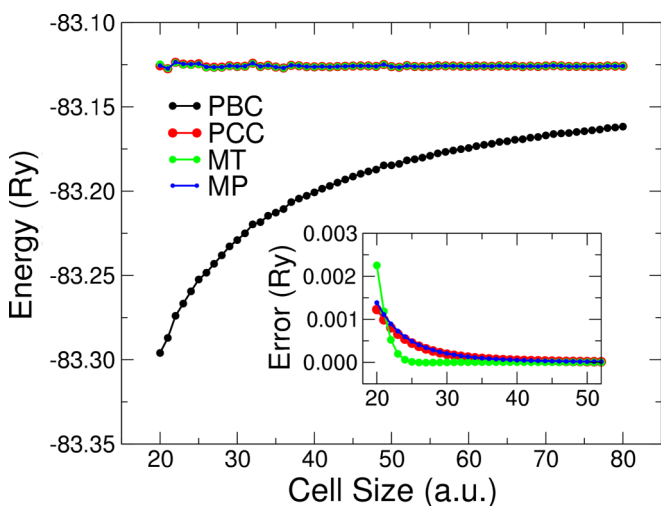


FIG. 1. (Color online) Total energy of a pyridine cation in vacuum as a function of cell size, for PBC calculations and for the three correction schemes analyzed: Makov-Payne (MP, in blue), Martyna-Tuckerman (MT, in green), and point-countercharge (PCC in red).

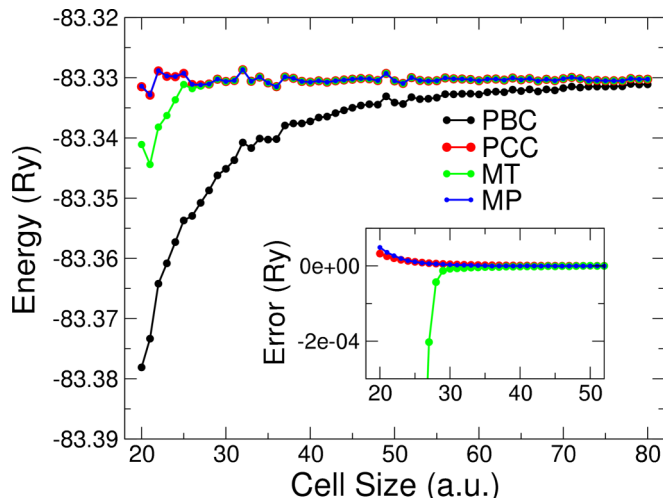


FIG. 2. (Color online) Total energy of a pyridine cation in a dielectric medium as a function of cell size, for PBC calculations and for the three correction schemes analyzed: Makov-Payne (MP, in blue), Martyna-Tuckerman (MT, in green), and point-countercharge (PCC in red). For the dielectric medium, the SCCS parameters optimized to reproduce aqueous solvation of organic compounds, as derived in Ref. [13], have been used, but only the electrostatic contribution has been explicitly considered.

It is important to note that Makov-Payne energies are almost identical to the ones obtained with the self-consistent real-space approach. This validates the hypothesis, assumed by Makov and Payne, that the polarization of the charge density of the system due to periodic images affects only marginally its energy. The same behavior is true for solvated systems, with

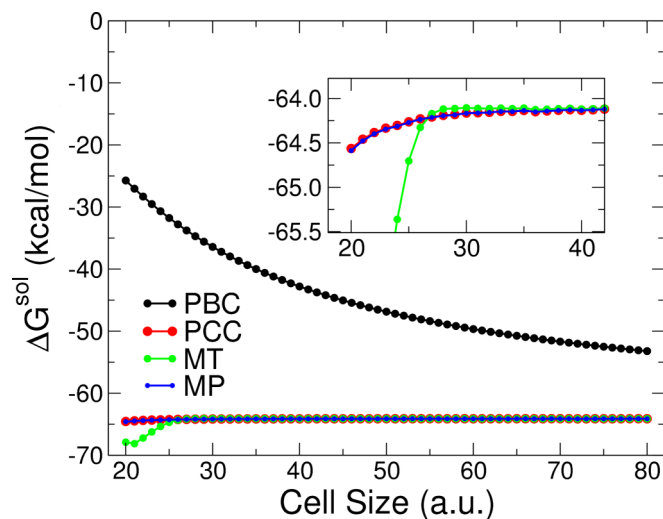


FIG. 3. (Color online) Solvation energies of a pyridine cation in continuum dielectric medium as a function of cell size, for PBC calculations and for the three correction schemes analyzed: Makov-Payne (MP, in blue), Martyna-Tuckerman (MT, in green), and point-countercharge (PCC in red). For the dielectric medium, the SCCS parameters optimized to reproduce aqueous solvation of organic compounds, as derived in Ref. [13], have been used, but only the electrostatic contribution has been explicitly considered.

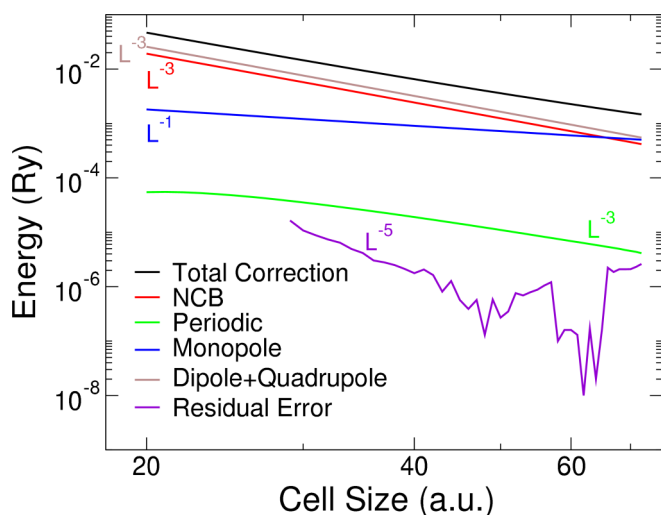


FIG. 4. (Color online) Cell-size dependence of Makov-Payne (monopole and dipole+quadrupole) and polarization-specific (NCB and periodic) contributions to the energy. Reported decay exponents are estimated from fitting the large cell-size part of the figure. The residual error, after the different contributions have been subtracted from the total energy of the system, is shown to decay faster than  $L^{-3}$  for cell sizes up to 40 a.u., while being negligible for larger cell sizes.

the Makov-Payne and the PCC methods almost exactly on top of each other.

Electrostatic solvation free energies, computed as the difference in total energy between the solvated and the vacuum case, are reported in Fig. 3 and substantially reflect what was found above: MP and PCC calculations give well-converged results (errors smaller than 0.5 kcal/mol) for all the system sizes considered, while the MT scheme gives large errors up to cell sizes of 26 a.u.

When looking at the different contributions to the Makov-Payne correction for a solvated system (Fig. 4), it appears that all contribute significantly. In particular, the effects of the

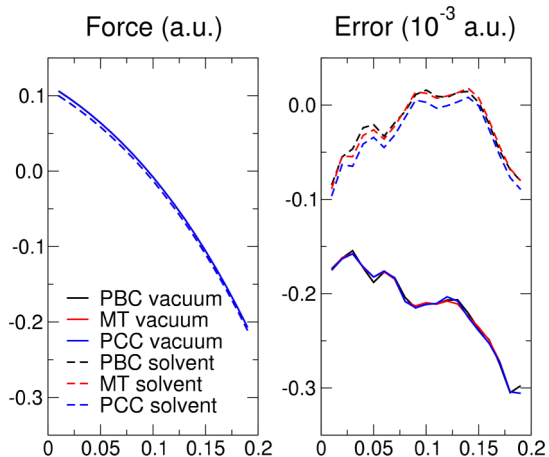


FIG. 5. (Color online) Hellmann-Feynman force (left panel) and error on forces (right panel) computed via finite differences for the nitrogen atom of the pyridine cation, in vacuum and in a continuum dielectric medium, with PBC or with MT and PCC correction schemes.

periodic images on polarizing the dielectric close to the solute are important, especially considering that, although small, such polarization is not charge neutral in the case of charged solutes.

In order to validate the formulas derived and the implementation of the different methods, the errors on the analytic forces have been reported in Fig. 5 for the different approaches considered as well as for the fully periodic case, with and without a continuum solvent. All the different methods show a very similar behavior, with errors almost three orders of magnitude smaller than the absolute value of the computed force. It is important to stress that, even though the reported errors are not negligible, all the methods are in agreement to what is found for the periodic calculations without the solvent,

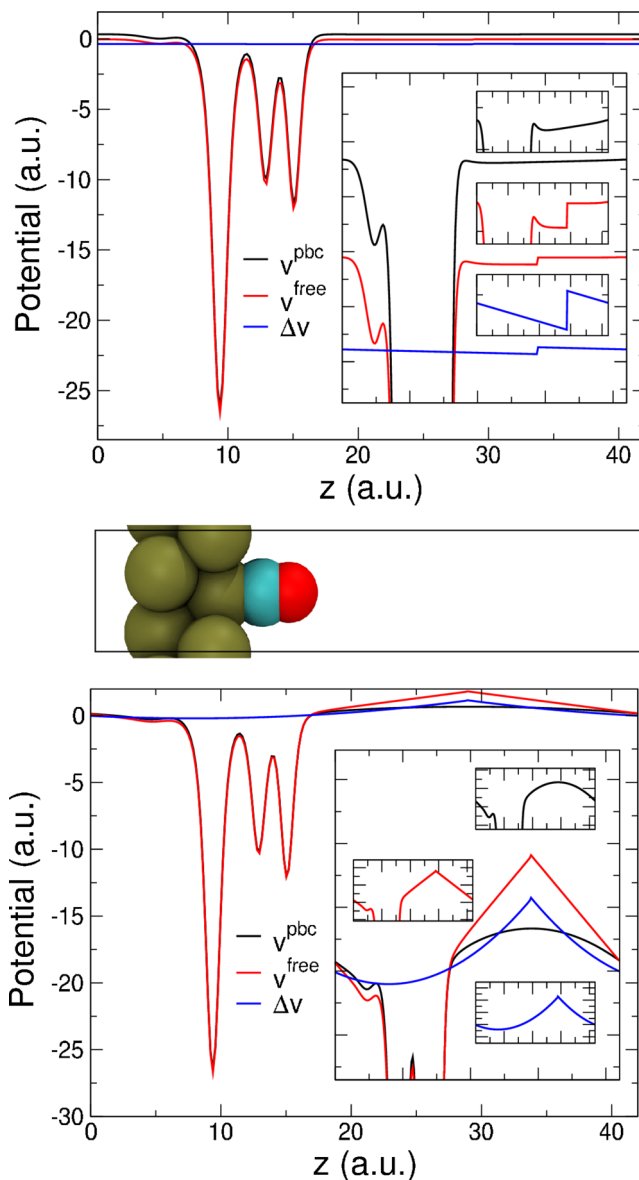


FIG. 6. (Color online) Electrostatic potential along the axis passing through the C-O bond for a neutral (top) and charged (bottom) slab in vacuum, as represented in the middle panel (Pt atoms in yellow, carbon atom in light blue, oxygen atom in red). The PBC potential ( $v^{\text{PBC}}$ , in black), PCC correction ( $\Delta v$ , in blue), and corrected potential ( $v^{\text{free}}$ , in red) are reported and compared in the main panels and in the insets.

which represents the internal benchmark of this work. Lower values for these errors can be obtained by fine tuning the setup of the calculations.

### C. Slab (2D) geometries

When considering neutral two-dimensional systems, the presence of a component of the dipole moment along the axis normal to the plane of the slab will create a step in the electrostatic potential. Due to periodic boundary conditions requiring the potential to be continuous at the boundary of the simulation cell, a deformation of the potential in the entire cell will take

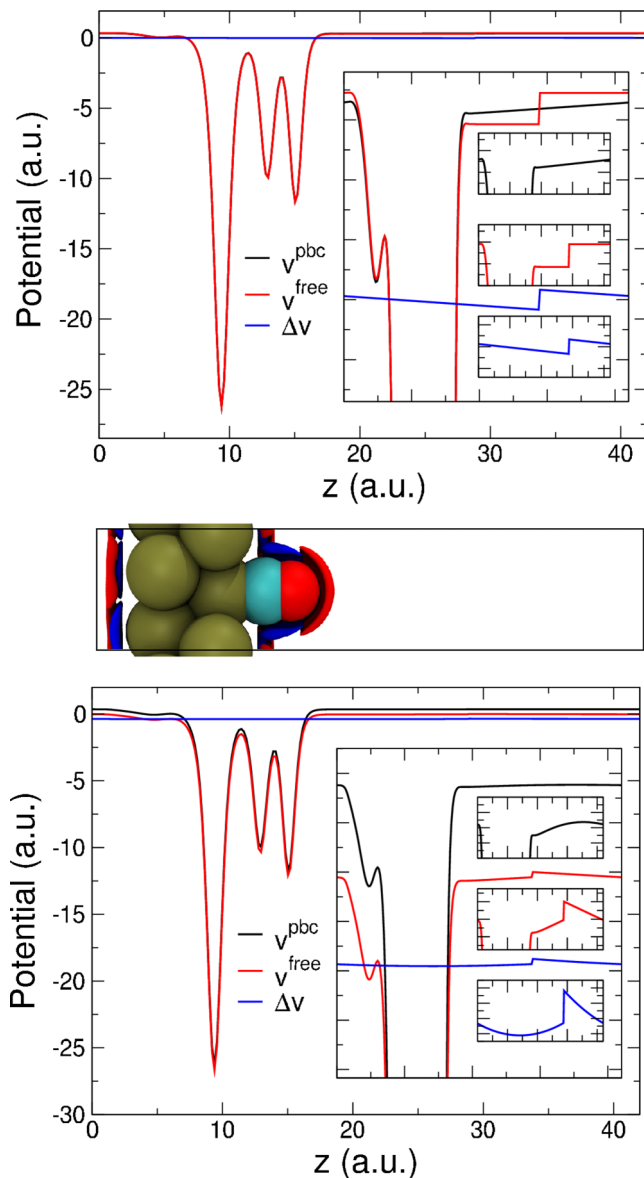


FIG. 7. (Color online) Electrostatic potential along the axis passing through the C-O bond for a neutral (top) and charged (bottom) slab in a continuum dielectric, as represented in the middle panel (Pt atoms in yellow, carbon atom in light blue, oxygen atom in red, positive and negative polarization charges are visualized as red and blue solid isosurfaces). The PBC potential ( $v^{\text{PBC}}$ , in black), PCC correction ( $\Delta v$ , in blue), and corrected potential ( $v^{\text{free}}$ , in red) are reported and compared in the main panels and in the insets.

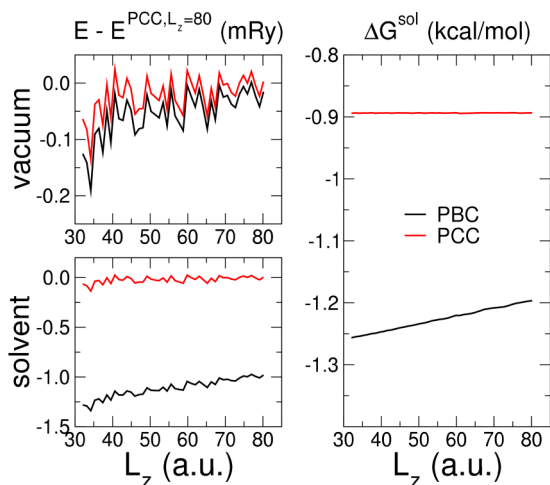


FIG. 8. (Color online) Total energy in vacuum (top left) and in solution (bottom left), and solvation energy (right) of a neutral slab as a function of cell thickness. For the dielectric medium, the SCCS parameters optimized to reproduce aqueous solvation of organic compounds, as derived in Ref. [13], have been used, but only the electrostatic contribution has been explicitly considered.

place (see top panels of Figs. 6 and 7). When looking at the energy of the neutral system as a function of cell size, Fig. 8, it looks that periodic boundary conditions have very different effects in vacuum and in solution. While the correction amounts to only a fraction of mRy for the system in vacuum, the effects of periodicity are more than 10 times larger in solvated systems. The same trends are clearly reflected in the behavior of the solvation energy of the system, which shows PBC errors of  $\sim 0.3$  kcal/mol even for the largest cells considered. In fact, the overall error seems to decay very slowly with cell thickness. This behavior is due to the spurious finite electric field in the interfacial region where the dielectric properties of the environment change: given the expression of the polarization charge [Eq. (30)], the presence of the artificial linear electrostatic potential induces a substantial amount of polarization in the continuum environment. Thus, even though the dielectric medium should compensate the intrinsic dipole of the slab, the interaction with the finite field due to periodic boundary conditions overstabilizes larger polarization charges and increases the PBC artifacts on the energy of solvated 2D systems. The simple 2D PCC correction proposed in this paper is enough to remove this artifact and to provide energies in solutions which are well behaved with respect to the size of the simulation cell, as clearly shown in Fig. 8. The corrected form of the potential (red lines in Figs. 6 and 7) is almost constant in the regions of space above and below the slab, thus providing the right contribution to the polarization of the surrounding medium.

In the case of a charged two-dimensional system, a converged reference value for the energy of the system is not accessible since the field, and thus the electrostatic energy density, of a planar distribution of charge is constant in space. Thus, the electrostatic potential and the total energy will linearly increase with the size of the cell axis perpendicular to the plane of the slab.

Such a behavior is correctly recovered in the trends of the potential in Figs. 6 and 7 for the system in vacuum and in

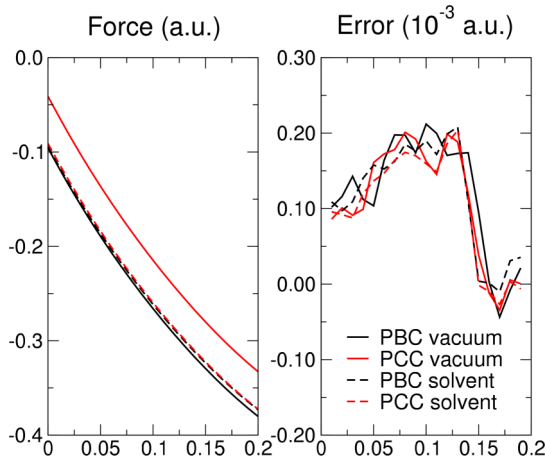


FIG. 9. (Color online) Hellmann-Feynman force (left panel) and error on forces (right panel) computed via finite differences for the carbon atom in the charged slab calculation, in vacuum and in a continuum dielectric medium, with and without PBC.

solution, respectively. The presence of the dielectric medium significantly screens the electrostatic potential, which shows much smaller variations with respect to the case in vacuum.

As in the case of the isolated 0D system, the analytical forces computed with the PBC correction, both with and without the continuum solvent, show errors of similar magnitude to the ones obtained from calculations without the correction (see Fig. 9). Self-consistent convergence of the polarization charge for the calculation of the charged slab in periodic boundary conditions appears to be hindered by the presence of the NCB density. In particular, a very large number of iterations are required in order to achieve the same accuracy on the polarization charge as in other calculations (i.e., with mean-squared variation of the polarization density of the order of  $10^{-12}$  a.u.). Even though the polarization density due to the neutralizing charge background has been shown to add a spurious cell-size-dependent term to the energy, it is important to note that in order to properly describe forces in a charged slab, the NCB polarization needs to be taken into account explicitly.

#### IV. CONCLUSIONS

To conclude, an extension to three current methods (Makov-Payne [31], Martyna-Tuckerman [33], and PCC [40]) to correct for periodic boundary conditions in systems of reduced dimensionality is presented, that allows us to treat quantum systems immersed in a continuum dielectric. Two different geometries have been explicitly addressed, namely, isolated (0D) and slab (2D) configurations. The modified Makov-Payne correction is summarized by Eq. (71), where the energy of an isolated system solvated in a continuum dielectric is expressed in terms of its artificially periodic counterpart plus a post-processing, computationally inexpensive, correction term. The main results for the PCC scheme are, instead, summarized by Eqs. (81) and (88) for the 0D and 2D cases, respectively, where the real-space corrections to the electrostatic potential of the solvated system are reported. While the Martyna-Tuckerman scheme is shown to be intrinsically more easy to adapt to the electrostatic equations defined by the SCCS approach, we

underline the derivation of the correction's contributions to interatomic forces in the presence of a continuum dielectric, as summarized in Eq. (105).

The analytical modifications introduced due to the presence of the solvent have been shown to be accurate in all the different methods. The derivation of the modifications proposed here can be easily extended to more complex methods, such as the density countercharge, or to linear (1D) systems. Results are shown to converge reasonably fast with system size. A comparison of the different approaches indicates that the Martyna-Tuckerman is the technique that converges faster with cell size, but can produce erratic results for cell sizes that are too small. The behavior of Makov-Payne and PCC correction schemes is very similar and is much smoother with system size, so that smaller cells with respect to MT can be used at the price of a small loss in accuracy. The very good agreement of the forces, computed analytically and via finite differences of the energy, is a further test on all derivations and their implementation.

#### ACKNOWLEDGMENTS

The authors acknowledge C. Dupont, I. Dabo, and S. de Gironcoli for useful discussions and the Swiss Platform for Advanced Scientific Computing (PASC) for funding.

#### APPENDIX: IONIC CHARGE DISTRIBUTION IN SCCS AND REAL-SPACE PERIODIC BOUNDARY CORRECTIONS

As discussed in Ref. [13], the way the ionic charge density is described in the SCCS model has no effect on the computed polarization density or on the polarization energy, provided that the ionic charge is well within the range where the dielectric constant is exactly one. Since Gaussian charges of fixed spread are used to model nuclei through the SCCS calculation, solvation energies were shown to be independent of their spread for a large range of values (Fig. 18 of Ref. [13]). Nonetheless, when computing the correction for periodic boundary conditions in real space, either in the self-consistent (PCC) or in the non-self-consistent (Makov-Payne) case, it is important that the ionic cores are treated on the same footing as the calculation of the ionic electrostatic energy. If Ewald summations are used to model the ionic contribution to the total energy, ions need to be considered as pointlike in the calculation of the system's multipole moments  $d^{\text{solute}}$  and  $Q^{\text{solute}}$  that enter in Eq. (44). On the other hand, in order for the polarization energy to be independent of the shape of the ionic density, a consistent description needs to be used in Eqs. (34) and (71).

Different choices of ionic shapes will in general give rise to different multipole moments. Nonetheless, if we chose to work in a coordinate frame originating from the center of ionic charges, the nuclei's contribution to the system's dipole vanishes regardless of the fact that they be described as pointlike or Gaussian charges. Thus, with such a choice of origin, only the quadrupole moment of the system depends on the shape of the ionic charge distribution. In order to correctly remove all the terms depending on the quadrupole, Gaussian-shaped ions have to be used when correcting the solvation energy, while pointlike nuclei have to be used for the

remaining part. Thus, an extra term of the form

$$\Delta E^{0D,\text{Gaussian}} = \frac{\pi}{3L^3} q^{\text{pol}} (Q^{\text{solute,Gaussian}} - Q^{\text{solute,pointlike}}) \quad (\text{A1})$$

for 0D systems, or

$$\Delta E^{2D,\text{Gaussian}} = \frac{\pi}{AL_z} q^{\text{pol}} (Q_{zz}^{\text{solute,Gaussian}} - Q_{zz}^{\text{solute,pointlike}}) \quad (\text{A2})$$

for 2D systems has to be added to the correction of the energy. An extra correction on the forces is also needed in the case of 0D systems (while it vanishes due to symmetry in the 2D case), namely,

$$\Delta f_a^{0D,\text{Gaussian}} = -\frac{2\pi}{3L^3} q^{\text{pol}} \left( \frac{z_a \sigma_a}{\sqrt{\pi}} \right), \quad (\text{A3})$$

where  $z_a$  and  $\sigma_a$  are the atomic charge and spread of atom  $a$ .

- 
- [1] J. Tomasi and M. Persico, *Chem. Rev.* **94**, 2027 (1994).  
 [2] J. Tomasi, B. Mennucci, and R. Cammi, *Chem. Rev.* **105**, 2999 (2005).  
 [3] C. J. Cramer and D. G. Truhlar, *Chem. Rev.* **99**, 2161 (1999).  
 [4] M. Orozco, *Chem. Rev.* **100**, 4187 (2000).  
 [5] J. L. Fattebert and F. Gygi, *J. Comput. Chem.* **23**, 662 (2002).  
 [6] J. L. Fattebert and F. Gygi, *Int. J. Quantum Chem.* **93**, 139 (2003).  
 [7] D. A. Scherlis, J. L. Fattebert, F. Gygi, M. Cococcioni, and N. Marzari, *J. Chem. Phys.* **124**, 074103 (2006).  
 [8] V. M. Sanchez, M. Sued, and D. A. Scherlis, *J. Chem. Phys.* **131**, 174108 (2009).  
 [9] J. Dziedzic, H. H. Helal, C. K. Skylaris, A. A. Mostofi, and M. C. Payne, *Europhys. Lett.* **95**, 43001 (2011).  
 [10] S. A. Petrosyan, A. A. Rigos, and T. A. Arias, *J. Phys. Chem. B* **109**, 15436 (2005).  
 [11] K. Letchworth-Weaver and T. A. Arias, *Phys. Rev. B* **86**, 075140 (2012).  
 [12] D. Gunceler, K. Letchworth-Weaver, R. Sundararaman, K. A. Schwarz, and T. A. Arias, *Modell. Simul. Mater. Sci. Eng.* **21**, 074005 (2013).  
 [13] O. Andreussi, I. Dabo, and N. Marzari, *J. Chem. Phys.* **136**, 064102 (2012).  
 [14] C. Dupont, O. Andreussi, and N. Marzari, *J. Chem. Phys.* **139**, 214110 (2013).  
 [15] J. Shim, E.-K. Lee, Y. J. Lee, and R. M. Nieminen, *Phys. Rev. B* **71**, 035206 (2005).  
 [16] P. A. Schultz, *Phys. Rev. Lett.* **96**, 246401 (2006).  
 [17] C. W. M. Castleton, A. Höglund, and S. Mirbt, *Phys. Rev. B* **73**, 035215 (2006).  
 [18] S. Lany and A. Zunger, *Phys. Rev. B* **78**, 235104 (2008).  
 [19] N. D. M. Hine, K. Frensch, W. M. C. Foulkes, and M. W. Finnis, *Phys. Rev. B* **79**, 024112 (2009).  
 [20] S. Lany and A. Zunger, *Modell. Simul. Mater. Sci. Eng.* **17**, 084002 (2009).  
 [21] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, *Phys. Rev. Lett.* **102**, 016402 (2009).  
 [22] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, *Phys. Status Solidi B* **248**, 1067 (2011).  
 [23] F. Corsetti and A. A. Mostofi, *Phys. Rev. B* **84**, 035209 (2011).  
 [24] H.-P. Komsa, T. Rantala, and A. Pasquarello, *Phys. B (Amsterdam)* **407**, 3063 (2012).  
 [25] H.-P. Komsa, T. T. Rantala, and A. Pasquarello, *Phys. Rev. B* **86**, 045112 (2012).  
 [26] S. T. Murphy and N. D. M. Hine, *Phys. Rev. B* **87**, 094111 (2013).  
 [27] P. H. Hünenberger and J. A. McCammon, *J. Chem. Phys.* **110**, 1856 (1999).  
 [28] T. N. Heinz and P. H. Hünenberger, *J. Chem. Phys.* **123**, 034107 (2005).  
 [29] P. J. Feibelman, *J. Chem. Phys.* **81** (1984).  
 [30] P. J. Feibelman, *Phys. Rev. B* **33**, 719 (1986).  
 [31] G. Makov and M. C. Payne, *Phys. Rev. B* **51**, 4014 (1995).  
 [32] M. R. Jarvis, I. D. White, R. W. Godby, and M. C. Payne, *Phys. Rev. B* **56**, 14972 (1997).  
 [33] G. J. Martyna and M. E. Tuckerman, *J. Chem. Phys.* **110**, 2810 (1999).  
 [34] P. A. Schultz, *Phys. Rev. B* **60**, 1551 (1999).  
 [35] P. Minary, M. E. Tuckerman, K. A. Pihakari, and G. J. Martyna, *J. Chem. Phys.* **116**, 5351 (2002).  
 [36] P. Minary, J. A. Morrone, D. A. Yarne, M. E. Tuckerman, and G. J. Martyna, *J. Chem. Phys.* **121**, 11949 (2004).  
 [37] S. Ismail-Beigi, *Phys. Rev. B* **73**, 233103 (2006).  
 [38] L. Genovese, T. Deutsch, A. Neelov, S. Goedecker, and G. Beylkin, *J. Chem. Phys.* **125**, 074105 (2006).  
 [39] L. Genovese, T. Deutsch, and S. Goedecker, *J. Chem. Phys.* **127**, 054704 (2007).  
 [40] I. Dabo, B. Kozinsky, N. E. Singh-Miller, and N. Marzari, *Phys. Rev. B* **77**, 115139 (2008).  
 [41] I. Dabo, B. Kozinsky, N. E. Singh-Miller, and N. Marzari, *Phys. Rev. B* **84**, 159910(E) (2011).  
 [42] Y. Li and I. Dabo, *Phys. Rev. B* **84**, 155127 (2011).  
 [43] N. D. M. Hine, J. Dziedzic, P. D. Haynes, and C.-K. Skylaris, *J. Chem. Phys.* **135**, 204103 (2011).  
 [44] A. Castro, E. Rasanen, and C. A. Rozzi, *Phys. Rev. B* **80**, 033102 (2009).  
 [45] M. A. Kastholz and P. H. Hünenberger, *J. Chem. Phys.* **124**, 124108 (2006).  
 [46] S. W. de Leeuw and J. W. Perram, *Mol. Phys.* **37**, 1313 (1979).  
 [47] S. W. de Leeuw, J. W. Perram, and E. R. Smith, *Proc. R. Soc. London, Ser. A* **373**, 57 (1980).  
 [48] S. W. de Leeuw, J. W. Perram, and E. R. Smith, *Proc. R. Soc. London, Ser. A* **373**, 27 (1980).  
 [49] S. W. de Leeuw, J. W. Perram, and E. R. Smith, *Proc. R. Soc. London, Ser. A* **388**, 177 (1983).  
 [50] L. Onsager, *J. Am. Chem. Soc.* **58**, 1486 (1936).  
 [51] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, *J. Phys.: Condens. Matter* **21**, 395502 (2009).  
 [52] A. D. Corso, *Pslibrary version 0.1*, <http://qe-forge.org/projects/pslibrary>.  
 [53] N. Marzari, D. Vanderbilt, A. De Vita, and M. C. Payne, *Phys. Rev. Lett.* **82**, 3296 (1999).