8-7-2023

Uncoupling System and Environment Simulation Cells for Fast-Scaling Modeling of Complex Continuum Embeddings

G. Medrano
University of North Texas

E. Bainglass
Paul Scherrer Institute

O. Andreussi
Boise State University

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Uncoupling system and environment simulation cells for fast-scaling modeling of complex continuum embeddings

Cite as: J. Chem. Phys. 159, 054103 (2023); doi: 10.1063/5.0150298
Submitted: 13 March 2023 • Accepted: 7 July 2023 •
Published Online: 1 August 2023

G. Medrano,[1] E. Bainglass,[2] and O. Andreussi[3,a)]

AFFILIATIONS

[1] Department of Physics, University of North Texas, Denton, Texas 76203, USA
[2] Materials Software and Data Group, Paul Scherrer Institut, 5232 Villigen PSI, Switzerland
[3] Department of Chemistry and Biochemistry, Boise State University, Boise, Idaho 83725, USA

[a)] **Author to whom correspondence should be addressed:** olivieroandreuss@boisestate.edu

ABSTRACT

Continuum solvation models are becoming increasingly relevant in condensed matter simulations, allowing to characterize materials interfaces in the presence of wet electrified environments at a reduced computational cost with respect to all atomistic simulations. However, some challenges with the implementation of these models in plane-wave simulation packages still persists, especially when the goal is to simulate complex and heterogeneous environments. Among these challenges is the computational cost associated with large heterogeneous environments, which in plane-wave simulations has a direct effect on the basis-set size and, as a result, on the cost of the electronic structure calculation. Moreover, the use of periodic simulation cells is not well-suited for modeling systems embedded in semi-infinite media, which is often the case in continuum solvation models. To address these challenges, we present the implementation of a double-cell formalism, in which the simulation cell used for the continuum environment is uncoupled from the one used for the electronic-structure simulation of the quantum-mechanical system. This allows for a larger simulation cell to be used for the environment, without significantly increasing computational time. In this work, we show how the double-cell formalism can be used as an effective periodic boundary conditions correction scheme for nonperiodic and partially periodic systems. The accuracy of the double-cell formalism is tested using representative examples with different dimensionalities, both in vacuum and in a homogeneous continuum dielectric environment. Fast convergence and good speedups are observed for all the simulation setups, provided the quantum-mechanical simulation cell is chosen to completely fit the electronic density of the system.

I. INTRODUCTION

Materials simulations have seen an incredible growth in recent years, thanks to the increasing power of computer hardware and simulation software, together with the development of new computational infrastructures to handle high-throughput simulations.[1–5] Unbiased first-principles simulations based on density functional theory (DFT) represent the workhorse of most computational approaches to materials design,[6] with many research groups pushing to improve the accuracy of computational predictions and extend the scope of first-principles simulations to more challenging systems and properties.

Traditional DFT simulations of materials focus on bulk periodic crystalline structures. For this reason, most condensed matter and materials simulation packages rely on basis sets composed of periodic functions (e.g., plane-waves, PW) uniformly occupying the simulation cell. This approach provides a systematically improvable basis set, whose accuracy can be tuned with a single parameter, usually associated with the kinetic energy of the corresponding plane-wave. PWs also provide easy access to the solution of the core electrostatic equations involved in first-principles simulations, with the Poisson equation in vacuum being readily solved in reciprocal space. Fast Fourier transforms (FFTs)[7] have proven instrumental to the widespread success and fast scaling of PW-based DFT packages, by providing real to reciprocal space mapping, and vice versa, at a cost proportional to $N_{pw}\log(N_{pw})$, with N_{pw} being the number of PWs in the simulation. However, what is advantageous for bulk

crystalline materials can be a source of artifacts when nonperiodic or partially periodic systems need to be characterized. Traditionally, materials interfaces are simulated using the slab approach, in which a few-atoms thick slab of the material is described in a three-dimensional periodically repeated simulation cell. By increasing the thickness of the slab and by increasing the cell size in the direction orthogonal to the material interface, the periodic replicas of the slab are effectively uncoupled from each other and the results are considered to be representative of the surface of a semi-infinite bulk material. A similar strategy can be devised for systems that are only periodic in one dimension, such as nano-wires or nanotubes, or systems that are nonperiodic. While adjusting the cell size can provide an easy solution to periodic boundary conditions (PBCs) artifacts, the number of PWs grows linearly with the cell volume, leading to a polynomial increase in the cost of the corresponding DFT calculation. For this reason, multiple alternative strategies to correct for PBC artifacts have been proposed in the past years, including real-space and reciprocal-space approaches. Most of these strategies are focused on simulations of two-dimensional and zero-dimensional (isolated) systems, while one-dimensional correction schemes are usually less widespread.

Unfortunately, as the typical scaling of DFT simulations varies as N^3, where N is the number of electrons in the system, complex and large systems with more than a few hundreds of electrons are still out of reach for systematic DFT studies. To overcome this limitation, a variety of hierarchical approaches[8,9] and/or divide and conquer[10–12] strategies have been developed in the literature. Without sacrificing the important atomistic characteristics of the relevant part of the process under investigation (the system), these methods take advantage of different, possibly simplified, models to handle the more macroscopic or complex part of the process (the environment), thus allowing a significant reduction of the computational cost.

In particular, implicit solvation models based on continuum embedding media have proven to provide a reasonable qualitative accuracy while significantly reducing the need for statistical sampling of disordered configurations.[13–16] Indeed, removing the atomistic details of solvent molecules not only reduces the number of electrons involved in the DFT simulation but allows to avoid the use of molecular dynamics (MD) to sample the configurations of the molecules of the liquid. Continuum methods in computational chemistry usually involve more or less empirical definitions of the different contributions to solvation free energy, which are then parameterized on experimental databases generated from solubility data. Methods such as the Polarizable Continuum Model (PCM)[15] of Tomasi and co-workers have shown good success in predicting ground state and response properties of molecules in solution. More recently, similar approaches have been translated into condensed matter simulation packages in order to study wet, possibly electrified, interfaces of materials.[8] When compared to the more widespread continuum solvation models in the computational chemistry community, the models implemented in periodic PW-based simulation packages rely on a smoothly varying boundary between the quantum-mechanical system and the continuum environment, instead of a sharp two-dimensional interface. Different definitions of the interface, based off the electronic density of the solute[17] or its atomic positions,[18] have been developed and show similar accuracy, when properly parameterized. While computationally more demanding, the use of smooth interfaces allows continuum models in condensed matter to more seamlessly introduce nonlocal corrections[19,20] to avoid some of the typical artifacts of these solvation methods.

Beyond the use of continuum solvation to model bulk neutral solutions, extensions to handle liquid crystals and diluted ionic solutions were implemented exploiting the integral equation formalism (IEF) of PCM.[21] Still following the same strategy, PCM was extended to the study of a single two-dimensional interface between different dielectric media (liquid–air, liquid–liquid).[22] The authors of the work of Corni *et al.* extended the use of the polarizable dielectric model to handle the electrostatic effects of metal surfaces and nanoparticles on nearby molecular dyes, thus allowing a more quantum-mechanical characterization of surface-enhanced spectroscopies.[23–25]

Similar to the PCM literature, more complex and heterogeneous environments can be introduced in condensed matter simulations. In particular, a significant focus has been devoted to study electrolyte solutions,[22,26–35] so as to unlock the use of these models for simulations of electrified interfaces. Going beyond neutral solutions and electrolyte distributions, Campbell and Dabo[36,37] exploited a continuum approach to model charge reorganization in a semiconductor substrate.

As the definition of the continuum environment in smooth-interface solvation models follows a function defined everywhere in the simulation cell, it is relatively straightforward to introduce more complex and heterogeneous environments, e.g., by introducing multiple continuum media, each with its own boundaries and physical properties (e.g., dielectric constant or surface tension). These media can be used to model substrate effects on the properties of overlaying materials or molecular compounds. Following this idea, Bononi and co-workers[38,39] exploited the possibility to introduce flat two-dimensional dielectric regions in the simulation cell in order to characterize in a computationally inexpensive way the effect of an ice substrate on the absorption properties and photodegradation of molecular dyes.

Despite the flexibility of continuum embedding approaches in condensed matter environments, PW-based simulations can pose significant challenges to these class of multiscale methods. The use of periodic simulation cells is intrinsically unfit to model systems that are embedded in a semi-infinite medium, i.e., most solvated systems would typically require nonperiodic simulation cells in the directions where the continuum solvent resides. Moreover, the fact that the simulation cell size is directly related to the basis-set size N_{pw} makes simulating large heterogeneous environments computationally more challenging. Extensions of both real-space and reciprocal-space PBC correction schemes have been developed to account for continuum dielectric media in simulations of isolated (0D) and slab (2D) systems.[40] In both continuum electrolyte and semiconductor models, long-range charge reorganizations are handled implicitly, by relying on analytical solutions to the electrostatic problems and/or appropriate boundary conditions. However, simulating the effects of nano-sized environments, such as plasmonic nanoparticles, micelles, liquid nanodroplets, would require the use of simulations cells, and thus basis sets, that are beyond reach of standard computational resources.

The straightforward solution to the aforementioned problem is to uncouple the simulation cell exploited for the continuum

environment from the one used for the DFT simulation of the quantum-mechanical system. The use of a large simulation cell for the environment would not cause significant limitations to the simulation time: The time-consuming part of the continuum environment calculation is related to computing the electrostatic potentials, by solving modified and generalized forms of the Poisson or Poisson–Boltzmann equations. Most of the algorithms implemented rely on iterative strategies in which each iteration involves the solution of a simple Poisson equation, which can be performed via FFTs in a time almost linear in cell size. By keeping the simulation cell of the DFT calculation small, the largest overhead in computational time is avoided. While ideal in theory, in practice this double-cell formalism requires full uncoupling of the two simulation methods and careful design of the mapping between the two simulation cells.

While designed to account for large heterogeneous environments, the double-cell formalism can also be used for nonperiodic and partially periodic systems as an effective PBC correction scheme. By increasing the length of the lattice vectors going along the nonperiodic directions in the environment cell, electrostatics can be fully converged without affecting the basis-set size and cost of the quantum-mechanical calculation. This general strategy works equally well for isolated (0D), slab (2D), but also one-dimensional (1D) systems. In this work, the details of the double-cell methodology are presented and the approach is tested as a PBC correction scheme for some representative examples with different dimensionalities, both in vacuum and in a continuum dielectric environment. Along with this new methodology and in order to test the accuracy and the speedup of the double-cell formalism for PBC corrections, the auxiliary-function countercharge corrections library (LIBAFCC) of Li and Dabo[41] was coupled to the continuum embedding software, giving access to an alternative exact correction scheme for all types of partially periodic systems.

The paper is organized as follows: In Sec. II, we review the main details of the implementation of the double-cell formalism, as well as the coupling of the auxiliary-function countercharge corrections library (LIBAFCC);[41] in Sec. III, the computational details of the DFT calculations as well as how the double-cell calculations are reported. Eventually, in Sec. IV we report a detailed benchmark of the accuracy and performance of the double-cell approach for modeling partially periodic systems, including 0D, 1D, and 2D systems, in vacuum and in a uniform dielectric medium.

II. METHODS

A. Double-cell formalism

The design of a double-cell algorithm is strongly dependent on the numerical details that are associated with the calling program (e.g., the PW DFT simulation package), the continuum embedding library, and with their coupling. In particular, the key aspects that affect the implementation are

1. which scalar fields need to be passed from one cell to the other,
2. how the scalar fields are stored and parallelized in the two cells, and
3. the way the grid points of the two cells are aligned.

In the following, we will refer to the specific implementation involving Quantum Espresso as the DFT driver and Environ as the continuum embedding library. A schematic representation of the steps involved in the algorithm is visualized in Fig. 1, while a more detailed example in a simplified two-dimensional orthorhombic cell is reported in Fig. 2. In order to generate the continuum boundary in Environ, information on the electronic density of the system and/or its atomic positions is required. The same information is also needed for the calculation of the electrostatic potential, i.e., the position and valence of the ions as well as the full electronic density of the system are required inputs for an Environ calculation. In alternative polarizable dielectric models, such as PCM, the electrostatic potential of the system in vacuum is passed to the continuum embedding module, to be used to solve for the polarization charge on the environment. However, passing the full information of the charge of the solute allows us to have a consistent definition of the electrostatic potential, e.g., when smeared ionic charges are used to account for core-electrons. In the case of the double-cell formalism, passing the full charge density of the solute also allows us to seamlessly adapt the calculation to a different simulation cell. In the presented implementation, the key scalar quantity that needs to be passed from the system cell into the environment cell is thus the electronic density of the system.

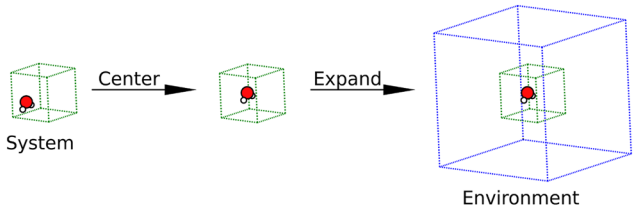

FIG. 1. A schematic view of the double-cell formalism with an isolated water molecule. The green cell is the system cell where all the PWs are introduced while the blue cell is the environment cell used for PBC correction.

The results of the continuum embedding calculations are usually reported in terms of corrections to the corresponding quantities computed in vacuum and periodic boundary conditions by the host DFT program. In particular, corrections to the total energy, interatomic forces, and Kohn–Sham potential are the key outputs of Environ calculations. Of these quantities, only the latter needs to be mapped from the environment cell into the system cell.

Scalar fields in a 3D simulation cell are discretized in PW simulation packages in terms of their values on a 3D structured grid, whose spacing is inversely related to the PW cutoff specified in the simulation input. However, the values on the 3D grid are stored by Quantum Espresso as one-dimensional arrays, by scanning grid points sequentially along the three cell axis. As part of the hierarchy of parallelization schemes adopted by Quantum-Espresso, parallelization on the real-space simulation cell is implemented by dividing it into slices along the third axis or into sticks along the second and third axes, depending on the total number of available processors. This allows us to perform operations that only depend on local values of the scalar field in a fully parallelized way while still requiring gathering and scattering of the whole simulation data for real to reciprocal space transformations (FFTs and inverse FFTs). Owing to the origin of Environ as an internal plugin of Quantum Espresso, the same strategies and the same numerical libraries are

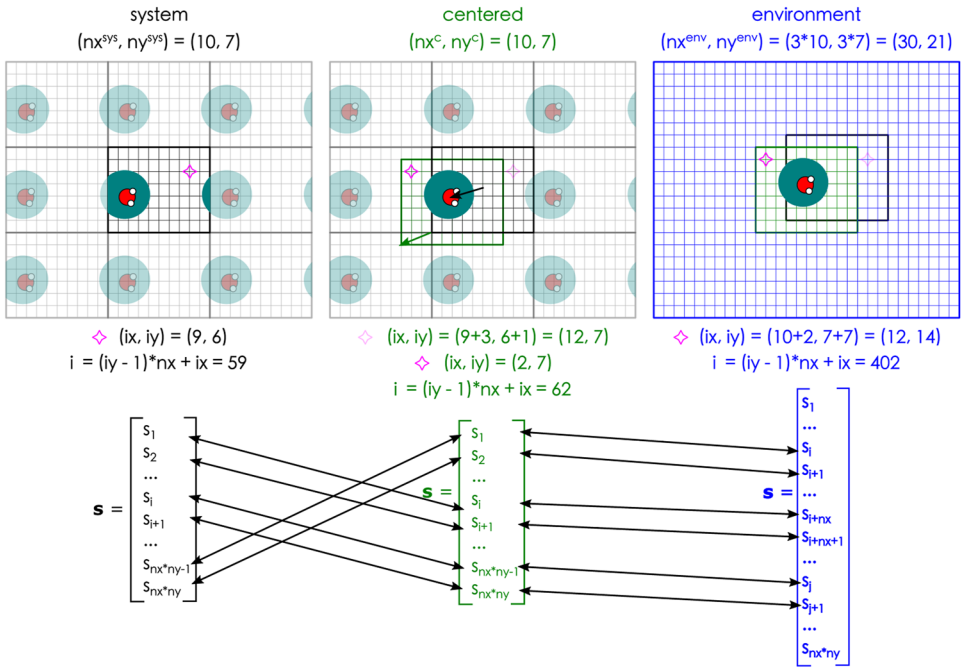

FIG. 2. Visualization of the mapping algorithm in a simplified two-dimensional rectangular grid. Left panels: A single water molecule and its electronic density (dark green circle) are mapped on a periodically repeated system cell (black rectangle), with a grid size of ten times seven points along the horizontal (x) and vertical (y) axes, respectively. A generic grid point (the pink diamond) can be identified by its pair of indices (ix and iy) or by a combined single index (i). A scalar field in the cell is thus represented as a one-dimensional array with index i ranging from 1 to $nx \times ny$. Central panels: A re-centering of the system cell on the center of mass of the system is required to ensure that the electronic density of the system is fully connected when the cell is expanded into the environment cell. The position of the center of mass with respect to the center of the system cell (black arrow) is used to determine the shift (green arrow) of the origin of the re-centered cell (green rectangle). To ensure that grid points of the two cells are at the same physical positions, integer operations (NINT and FLOOR) are adopted in the algorithm, resulting in an origin shift that may be slightly different from the real-valued shift of the center of mass of the system: The green and black arrows are meant to be different. While the number of grid points in the centered cell is the same as the system cell, their indices will be different, possibly including the fact that a grid point from a periodic image of the system cell needs to be selected. In the example, the original diamond is now replaced by one of its periodic images, its new ix and iy indices give rise to a new global index i, with the result that each entry in the original scalar field array is now mapped to a different entry of the re-centered array in a one-to-one fashion. Right panels: The centered cell is expanded in the two directions by one replica per side, for a total environment cell nine times larger than the original system cell. The new grid point indices need to account for the additional grid points in the replicas and the global grid point index is computed using the expanded grid size. While each grid point index in the system cell has a corresponding location in the environment cell, the contrary is not true, as most of the grid points in the environment cell are not meant to be mapped back to the system cell.

exploited by Environ for its internal calculations. For sake of simplicity in the equations, the following discussion will only focus on the serial algorithm, but the developed strategy is seamlessly extended to parallel implementations.

In practice, given a simulation cell with cell axes

$$\vec{h}_1 = \begin{pmatrix} h_{1x} \\ h_{1y} \\ h_{1z} \end{pmatrix}, \quad \vec{h}_2 = \begin{pmatrix} h_{2x} \\ h_{2y} \\ h_{2z} \end{pmatrix}, \quad \vec{h}_3 = \begin{pmatrix} h_{3x} \\ h_{3y} \\ h_{3z} \end{pmatrix}, \tag{1}$$

and a corresponding cell matrix

$$\hat{h} = \left(\vec{h}_1, \vec{h}_2, \vec{h}_3\right) = \begin{pmatrix} h_{1x} & h_{2x} & h_{3x} \\ h_{1y} & h_{2y} & h_{3y} \\ h_{1z} & h_{2z} & h_{3z} \end{pmatrix}, \tag{2}$$

the choice of the density cutoff in the simulation input determines the number of grid points along the three axes, $\mathbf{n}^s = (n_1^s, n_2^s, n_3^s)$. A generic scalar property (e.g., the electronic density or the electrostatic potential) in the simulation cell, $s(\vec{r})$, is thus discretized into an array, $s(\mathbf{i}^s) \equiv s(i_1^s, i_2^s, i_3^s) \equiv s(i^s)$ that contains the values of the property at each grid point's position, $\vec{r}(\mathbf{i}^s)$, with

$$\vec{r}(\mathbf{i}^s) = \hat{h} \cdot ((\mathbf{i}^s)/\mathbf{n}^s) \equiv \begin{pmatrix} h_{1x} & h_{2x} & h_{3x} \\ h_{1y} & h_{2y} & h_{3y} \\ h_{1z} & h_{2z} & h_{3z} \end{pmatrix} \cdot \begin{pmatrix} i_1^s/n_1^s \\ i_2^s/n_2^s \\ i_3^s/n_3^s \end{pmatrix}, \tag{3}$$

where the integer grid point indices i_{1-3}^s range from 0 to $n_{1-3}^s - 1$. A one-dimensional version of the array, with $i^s = 0, \ldots, n_1^s n_2^s n_3^s - 1$, is obtained by combining the three indices, e.g., by using the expression

$$i^s = 1 + i_1^s + i_2^s n_1^s + i_3^s n_1^s n_2^s, \tag{4}$$

which assumes that the scanning of grid points follows the order of the axes, i.e., faster along the first axis, than along the second, and slower on the third. Vice versa, given the global index i^s of a grid point, its three-dimensional indices can be reconstructed as

$$i_3^s = \frac{i^s - 1}{n_1^s n_2^s}, \tag{5}$$

$$i_2^s = \frac{i^s - 1 - i_3^s n_1^s n_2^s}{n_1^s}, \tag{6}$$

$$i_1^s = i^s - 1 - i_3^s n_1^s n_2^s - i_2^s n_1^s, \tag{7}$$

where the fractions are intended to represent integer division operations.

In designing the coupling between system and environment simulation cell, it is crucial to realize that the quantum-mechanical system lives in a fully periodic simulation cell: In the DFT calculation, the system energy and properties are not affected by arbitrary translations of the system's degrees of freedom. In particular, it is possible for a system to have some of its atoms or part of its electronic density on opposite sides of the simulation cell (e.g., as schematized in the left panel of Fig. 2), as the minimum image convention would still give rise to a well-defined fully connected system. However, when expanding the system cell into the environment cell, care must be taken to ensure that a fully connected system is passed to the continuum embedding module and that the environment is properly defined with respect to a meaningful center of the system. Thus, for the mapping of scalar fields between the two simulation cells, a two-step process was implemented as visualized in Figs. 1 and 2: first a centering of the system and second an expansion of the environment cell along the designated axes.

The first step is accomplished by computing the center of mass, $\vec{r}_{cm}$, of the system within its original simulation cell, using the minimum image convention to account for potential splits in neighboring cells. We then introduce a new cell with the same cell parameters and the same grid spacing as the system cell, but with an origin shifted from the Cartesian origin by a vector $\vec{o}$, so as to place the system's center of mass near its central grid point. In practice, we set the indices of a central grid point as $\mathbf{c} = \mathrm{NINT}(\mathbf{n^s}/2)$, with the NINT() operation selecting the nearest integer to the result of the division, in case of grids with odd numbers of grid points along an axis. We then select a grid point near the center of mass of the system as

$$\mathbf{m} = \mathrm{NINT}(\hat{h}^{-1} \cdot \vec{r}_{cm} \cdot \mathbf{n^s}). \tag{8}$$

The integer shift between the indices of the center of mass and the indices of the center of the cell, $\mathbf{d} = \mathbf{c} - \mathbf{m}$, provides the origin vector for the new re-centered cell,

$$\vec{o} = -\hat{h} \cdot (\mathbf{d}/\mathbf{n^s}) = -\begin{pmatrix} h_{1x} & h_{2x} & h_{3x} \\ h_{1y} & h_{2y} & h_{3y} \\ h_{1z} & h_{2z} & h_{3z} \end{pmatrix} \cdot \begin{pmatrix} d_1/n_1^s \\ d_2/n_2^s \\ d_3/n_3^s \end{pmatrix}, \tag{9}$$

as well as the shift used to remap the integer grid point indices, i_{i-3}^c, needed to compute the single-index representation in Eq. (4). However, in applying the shift to the grid point indices, we need to make sure that grid points in periodic images are mapped to the corresponding points in the re-centered cell (e.g., as visualized by the pink star in the central panel of Fig. 2). In the algorithm, this is enforced as follows:

$$i_{1-3}^c = (i_{1-3}^s + d_{1-3}) - \mathrm{FLOOR}((i_{1-3}^s + d_{1-3})/n_{1-3}^s) \times n_{1-3}^s, \tag{10}$$

where the FLOOR() integer operation returns the greatest integer less than or equal to the argument. As defined above, the re-centering of the cell corresponds to a one-to-one mapping of grid points, which is univocally defined by the position of the center of mass and the grid dimensions. Such a mapping does not involve any interpolation and can be computed on the fly without loss of information and without affecting the real-space parallelization scheme. In particular, given an initial grid point in the system cell with global index i^s, its corresponding one-dimensional index in the re-centered cell is given by

$$i^c = 1 + i_1^c + i_2^c n_1^s + i_3^c n_1^s n_2^s, \tag{11}$$

where i_{1-3}^c are computed by applying Eq. (10) to the indices obtained from Eq. (5).

The second step of the double-cell algorithm involves expansion of the system cell in different directions in order to generate a larger environment cell. While arbitrary expansions could in principle be considered, they would usually result in an environment grid spacing and grid point positions that would not match the initial ones on the system cell. This numerical mismatch could be easily addressed by interpolation algorithms, which would add an additional layer in the mapping between the two cells. However, in order to keep the process as simple and robust as possible, the current implementation of the double-cell formalism restricts the cell expansion to integer multiples of the initial cell. An expansion vector $\mathbf{n} = (n_1, n_2, n_3)$ is defined in the input of the environment module, with n_i representing the number of fictitious replicas added on both sides of the re-centered system cell along its ith axis. The combination of these two steps creates the desired environment cell, with lattice vectors $2\mathbf{n} + 1$ times as large as the system cell, with the center of mass of the system located near its central grid point. In practice, the grid point indices in the environment cell, i_{1-3}^e, need to account for the additional replicas added before the cell,

$$i_{1-3}^e = i_{1-3}^c + n_{1-3} n_{1-3}^s. \tag{12}$$

The corresponding global grid point index can thus be computed as

$$i^e = 1 + i_1^e + i_2^e n_1^e + i_3^e n_1^e n_2^e, \tag{13}$$

with $n_{1-3}^e = 2(n_{1-3} + 1) n_{1-3}^s$.

It is important to stress that in the above algorithm, each grid point of the system cell has a corresponding grid point of the environment cell at the same position: The system-to-environment mapping of a smooth charge density only involves a re-indexing of the entries of the array in the system cell and a zeroing of the elements of the environment cell that lie outside of the system cell. Vice versa, the environment-to-system mapping of the electrostatic potential involves re-indexing of the entries of the array in the environment cell that only correspond to grid points of the system cell, while the information stored in the other grid points is discarded. This corresponds to a mapped electrostatic potential that,

in general, will not possess the periodicity of the system cell and will likely present discontinuities at the cell borders corresponding to the axes that were expanded in the environment cell. This is consistent with other real-space PBC correction schemes, such as the parabolic/dipole correction for isolated and slab systems.[40] The presence of such a discontinuity will not affect the energy and convergence of the simulated system, provided that the atomic and electronic charges of the system are not present at the borders of the cell. While the re-centering step of the algorithm enforces that the borders are as far as possible from the system's center of mass, for small system cells it is possible that part of the charges of the system still overlaps or gets close to the borders of the cell. Convergence of the double-cell algorithm with system cell size can provide the necessary validation of the approach, as shown in Sec. IV.

B. Auxiliary-function correction

The long-range nature of Coulomb interactions leads to a non-convergent or ill-convergent result for charged or dipolar systems, when summing over all the periodic replicas of an infinite 3D periodic system.[42] The slow non-integrable behavior of the Coulomb potential at long distances has a mirror counterpart in its singularity at vanishing distances, which affects the calculation of electron interactions and exact exchange terms in periodic DFT simulations.[43,44] These singularities occur when the reciprocal-space vectors approach the Brillouin zone boundaries of partially periodic systems or when the interaction potentials exhibit rapid variations. Most simulation software set these divergences to zero, introducing negligible errors that approach zero as the size of the computational supercell increases.[42] Meanwhile others employ special techniques to mitigate these divergences, such as implementations of real-space countercharge corrections schemes.[45–50] Even though the countercharge correction schemes were created as a method for charged systems, they can still be used with neutral systems for a more comprehensive treatment of long-range electrostatic interactions. Example calculations, both neutral and charged, can be found in Ref. 50. As an alternative, real-space electrostatic corrections have been introduced that can remove these singularities although they can become costly due to these corrections requiring different interaction potentials per self-consistent iteration.

Because of this additional computational cost, reciprocal-space countercharge corrections have been studied in the form of auxiliary-function techniques that are the basis of the LIBAFCC[41] library. These corrections involve adding the auxiliary-function to the point-charge electrostatic kernel allowing for efficient removal of the singularities associated with the plane-wave representation. The LIBAFCC library implements exact point-charge reciprocal-space auxiliary functions for all types of partially periodic systems with the 0D correction giving the same formulas as the Martyna–Tuckerman[51] (MT) correction scheme.

Following an approach similar to the MT implementation, we coupled the library of Li and Dabo to a development version of Environ. The point-charge correction, as computed by the library, is mapped in the real-space grid, centered on the origin. Fast Fourier transforms allow to convert the correction to reciprocal space and to store it, so that it can be used to complement the standard vacuum FFT Poisson solver at no extra cost. More detailed information can be found in the LIBAFCC[41] reference.

III. COMPUTATIONAL DETAILS

All calculations were performed using Quantum Espresso (QE) v7.1[52,53] compiled with the Environ library for continuum embedding effects.[17,52,54] While the double-cell implementation has been officially released in Environ 3.0, for the results reported in the following we used a local development version of the library that includes coupling with the AFCC library for periodic boundary corrections.

All reported benchmarks, in vacuum and in dielectric environments, involve a single self-consistent field (SCF) optimization of the electronic density, starting from a random initial guess. In an effort to keep benchmarks as consistent as possible, we did not include geometry optimization calculations in our results. However, convergence of the interatomic forces with the newly implemented algorithms was thoroughly tested, showing an accuracy consistent with fully periodic simulations and with other PBC correction schemes.

The 0D and 2D simulations are based off the Environ's examples distributed with the library. In particular, the isolated system considered in this study is an acetamine cation, while a thin Pt (111) slab with an adsorbed carbon monoxide molecule was used for testing 2D systems. A boron nitride (BN) nanoribbon was chosen as benchmark for the 1D systems. The selected systems are known to display more pronounced PBC artifacts. However, it is important to stress that the reported analysis is meant to focus on the consistency of the results with alternative PBC correction schemes and on the overall computational performance. For this reason, the physical approximations and the simulation details that are not relevant for the convergence of the electrostatic energy have been selected according to the most widespread choices in the literature. All the calculations were performed using the Perdue–Burke–Ernzerhof generalized gradient approximation (GGA) density functional.[55] For the 0D and 2D systems, the standard ultrasoft pseudopotentials distributed on the QE website were adopted. For the 1D system, the pseudopotentials from the Standard solid-state pseudopotentials (SSSP) efficiency library were selected.[56,57] The simulations were performed by only sampling the gamma point sampling of the Brillouin zone.

A dielectric embedding environment was modeled using the self-consistent interface function presented in the work of Andreussi *et al.*[17] A homogeneous static dielectric permittivity of 100 was selected for the bulk of the environment. The preconditioned conjugate gradient (PCG) approach presented in the work of Fisicaro *et al.*[58] was used for the convergence of the generalized Poisson equation, with a convergence threshold on the computed electrostatic potential of $5.0e - 13$ Ry.[17,52]

For the acetamine cation (0D system), the wave function cutoff and density cutoff were chosen to be 30 and 300 Ry, respectively, with an estimated convergence threshold on the total energy of $5.0e - 6$ Ry. A simple cubic simulation cell was adopted, in order to allow the use of a parabolic correction to PBCs.[40,50] For this isolated system, reference simulations were also performed using the Martyna–Tuckerman reciprocal-space correction,[51] both in vacuum and in the dielectric medium.

For the 2D system, an orthorhombic cell was used with the cell parameter in the z direction being varied to evaluate the interaction with periodic replicas. The wave function cutoff and density cutoff were set to 35 and 300 Ry, respectively, with an

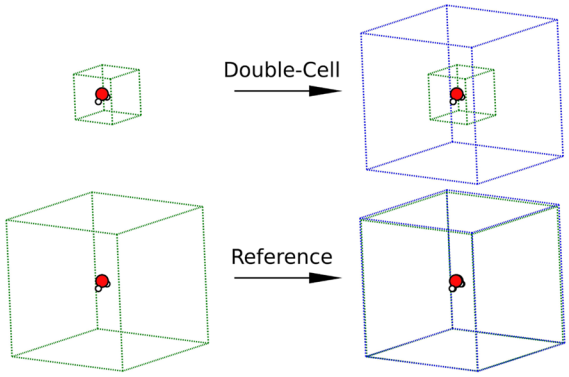

FIG. 3. Graphic representation of the two sets of calculations adopted to investigate convergence of the developed formalism. The double-cell calculations involve a small system cell (green box) embedded in an expanded environment cell (blue box). Reference calculations adopt the same expanded cell for both the system (DFT) and the environment (continuum electrostatic) calculations.

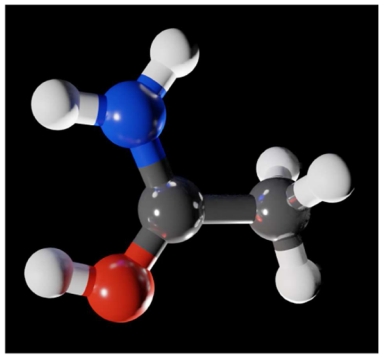

FIG. 4. The acetamine cation used for testing the double-cell formalism in isolated systems.

estimated convergence threshold on the total energy of $1.0e-6$ Ry. Marzari–Vanderbilt[59] smearing of the band occupations was adopted with a spread value of 0.03 Ry. A parabolic correction scheme, as implemented in the Environ library, was adopted for simulations in vacuum and in the dielectric medium.

Finally, for the 1D system, an ideal boron nitride (BN) 2D surface, as obtained from the aflowlib[4] repository, was used to create a 1D ribbon. The wave function and density cutoffs were set to 60 and 480 Ry, respectively, with an estimated convergence threshold on the total energy of $1.0e-7$ Ry. A Gaussian smearing with a spread value of 0.02 Ry was adopted, together with a local Thomas–Fermi mixing and a mixing parameter of 0.7. Although the parabolic correction scheme implemented in Environ is only compatible with 0D and 2D systems, a 2D correction with a large simulation cell in the direction along the plane of the ribbon was also exploited as reference for the 1D system.

In order to demonstrate the accuracy and performance of the double-cell formalism, two DFT calculations were done for each

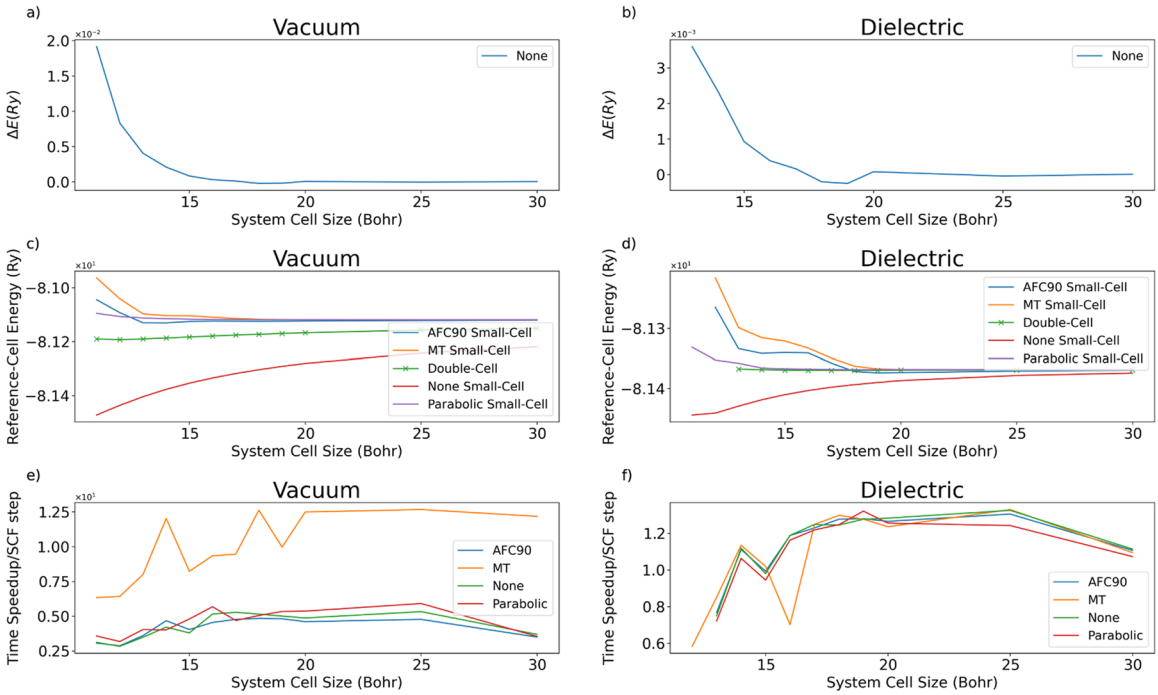

FIG. 5. Plots (a) and (b) show the energy difference, in Ry, between an SCF calculation using the double-cell formalism and the corresponding reference calculation (as visualized in Fig. 2). For plots (c) and (d), the convergence of different PBC schemes as a function of DFT system cell size is reported. Plots (e) and (f) show the time speedup between these simulations. All reported tests were performed for systems in vacuum (left column) and in a continuum dielectric medium (right column).

unit cell size, as schematized in Fig. 3. The first calculation is performed using the double-cell approach, i.e., having the environment cell defined as an integer expansion of the system cell in at least one direction. A reference calculation is then performed by imposing the system and environment cells to be of the same size and to match the environment cell of the first calculation. As an example, for a simulation of an isolated (0D) system with a cubic cell of size 10 bohrs, a double-cell calculation with one replica in every direction, $\mathbf{n} = (1, 1, 1)$, would involve a cubic environment cell with a 30 bohrs side length. The reference calculation would be performed with identical system and environment cells, both set to a size of 30 bohrs. In Sec. IV, the energy difference between the double-cell calculation and the reference are reported together with the speedup in calculation time.

IV. RESULTS

A. Isolated systems (0D)

As discussed above, one of the uses of the double-cell formalism is to reduce PBC artifacts in isolated systems, thus overcoming one of the main sources of artifacts of plane-wave based codes. The system selected for testing the accuracy and performance of the double-cell strategy is an acetamine cation, visualized in Fig. 4. To simplify the handling of PBC artifacts and to allow the use of the parabolic point-counter-charge correction, a cubic cell is used to study this system.

In the top panels of Fig. 5, the changes in total energy with respect to the reference are reported, for both simulations in vacuum and in a dielectric environment. The considered molecule has overall size of about 8.0 bohrs. However, from the convergence of the double-cell calculations with respect to the reference, it is clear that system cells smaller than 15 bohrs present artifacts with respect to the same calculation in the reference cell. As the electrostatic calculations in the double cell and the reference are the same, these small differences must be due to the fact that the double-cell mapping is affecting the DFT convergence. In fact, inspection of the density of the molecule reveals that the difference is due to non-negligible spilling of the electronic density of the molecule beyond the cell boundaries. While in a periodic cell this still corresponds to a connected smooth density, the mapping step of the double-cell formalism breaks the spilling density apart. This result suggests a general rule for the identification of the minimal system cell in a double-cell calculation, which needs to be about 10.0 bohrs larger than the system size.

When looking at the behavior of the double-cell energy as a function of system size, we notice that the results are within 0.1 Ry from the fully converged energy for all simulation cells larger than the identified minimum size. As the simulated system is a molecule with a nonzero total net charge, its electrostatic energy in PBC is expected to vary as $1/L$ with respect to the cubic simulation cell size, L (red curves in Fig. 5). With the parabolic correction scheme, such a dependence is corrected into a $1/L^5$, with energies virtually converged for cell sizes as large as 15 bohrs. Reciprocal-space corrections fully correct such a cell-size dependence, provided cell sizes twice as large as the system size are chosen. In the double-cell algorithm, the $1/L$ behavior is instead still present, although significantly smaller than the small-cell counterpart, as it corresponds to significantly larger periodic cell sizes. The same behavior is less evident for the simulations in a dielectric environment, as a significant fraction of the monopole–monopole interaction between periodic cells is screened by the embedding medium. In the bottom panels of Fig. 5, we report the speedup per SCF cycle of the double-cell formalism when compared with the reference calculations. This speedup accounts for the relative weight of the FFT-based electrostatic calculation and the cost of inverting the DFT Hamiltonian in a single SCF step. The reported results show an average fivefold decrease in computational time when the system cell is allowed to be smaller than the environment cell. In the dielectric embedding, the electrostatic problem needs to be solved iteratively (either using a fixed-point method[17] or using a preconditioned conjugate gradient approach[58]) where each iteration requires to solve a Poisson equation in vacuum. Given the multiple number of times the Poisson equation needs to be solved at each SCF step, the overhead in mapping and remapping scalar fields may become more evident in dielectric environments. In addition, given the fact that the dielectric medium is defined on the electronic density, for small simulation sizes, the convergence of the dielectric medium effects may suffer for the truncation of the electronic density at the cell boundaries.

While the results in this section aim to provide a comparison between alternative PBC schemes, it is important to stress that the double-cell formalism can also be used in combination with any other PBC corrections scheme reported above: The electrostatic calculation in the expanded environment cell can be corrected with a real-space term (parabolic correction) or with any reciprocal-space scheme (Martyna–Tuckerman or AFC) already implemented in Environ. Those results are not reported in this section because the simulations are already fully converged without the need for additional corrections.

B. Slab (2D) system

To study 2D systems, systems that should be periodic in two dimensions, we used a bilayer of platinum atoms in the (111) orientation and in the presence a CO molecule in the atop position, as visualized in Fig. 6. For this system, the slab is included in an orthorhombic cell, oriented in the x-y plane and with the nonperiodic axis along third direction. For the double-cell calculations, the system cell is expanded only in the z direction, while the system and environment cells are the same in the in-plane directions. Convergence tests are reported in Fig. 7 as a function of the size of the system cell along the z direction.

The considered system has an extension along the vertical axis of about 11.0 bohrs. When looking at the difference in

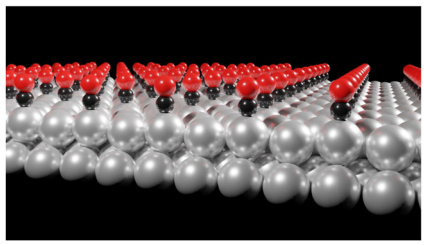

FIG. 6. Visualization of the system used for testing the double-cell algorithm in two dimensions: a Pt (111) surface with two layers of Pt atoms in the presence of a CO molecule adsorbed in the atop position.

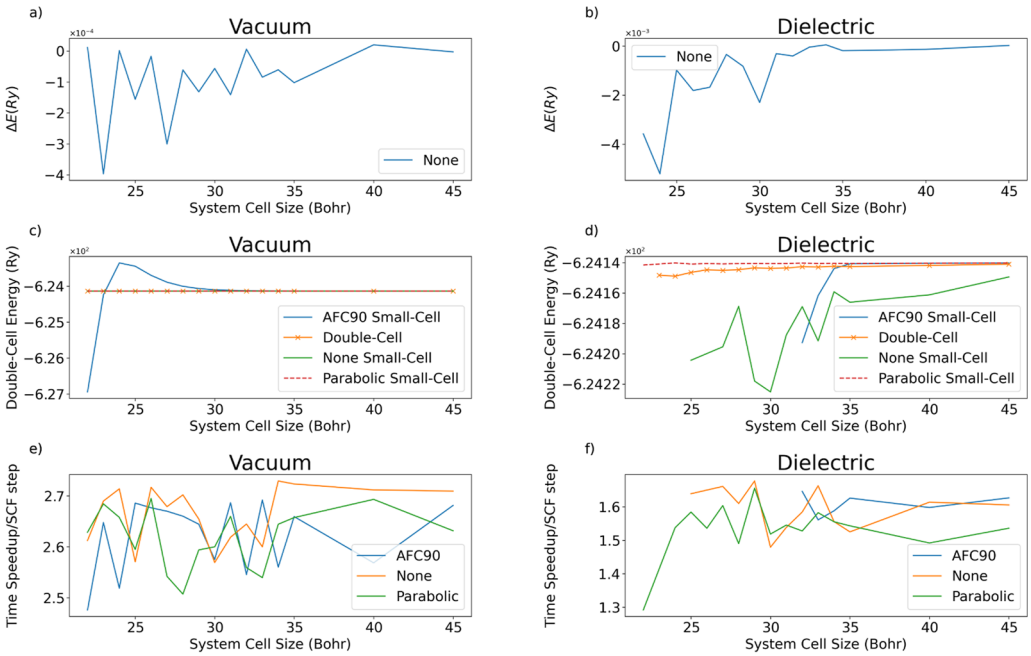

FIG. 7. Plots (a) and (b) show the energy difference, in Ry, between an SCF calculation using the double-cell formalism and the corresponding reference calculation (as visualized in Fig. 2). For plots (c) and (d), the convergence of different PBC schemes as a function of DFT system cell size is reported. Plots (e) and (f) show the time speedup between these simulations. All reported tests were performed for systems in vacuum (left column) and in a continuum dielectric medium (right column).

energy between the double-cell and the reference calculations, it appears that minimal artifacts are present at the smallest cell sizes (22–23 bohrs). This suggests that no significant spills in the electronic density are present for these DFT cells, consistent with our previous observations for the isolated system. As observed in the 0D case, double-cell energies are fully converged with respect to the DFT system cell already for the smallest cells considered. The results are in line with the convergence observed for a parabolic (also known as point-counter-charge or dipole correction) scheme, while AFC90 shows a slower convergence, with exact results for cells larger than 35 bohrs.

Eventually, in the bottom panels of Fig. 7, we report the computational speedup of using a double-cell approach, compared to using the same expanded cell for both electrostatic and DFT simulations. The relative weight of the electrostatic calculation and inverting the DFT Hamiltonian for this small system provide a speedup of about 2.5–2.8 in vacuum, which drops to about 1.5 in a dielectric medium. While the scaling of the FFT-based electrostatic simulation is expected to be better than the DFT one as a function of the cell size, speedups appear to be independent of cell size for the considered system. This is probably due to the relatively small number of electrons in the system and to the other computational overheads associated with larger FFT grids.

C. Wire (1D) system

For the 1D systems, we studied the boron nitride (BN) nanoribbon, as shown in Fig. 8. An orthorhombic cell is adopted, with the ribbon oriented along the x-axis, making the y and z axes the two nonperiodic directions. Moreover, in this case, only these two directions will be considered in the expansion of the system cell into the environment cell, while the system and environment cells have the same size along the x direction. For the sake of simplifying the

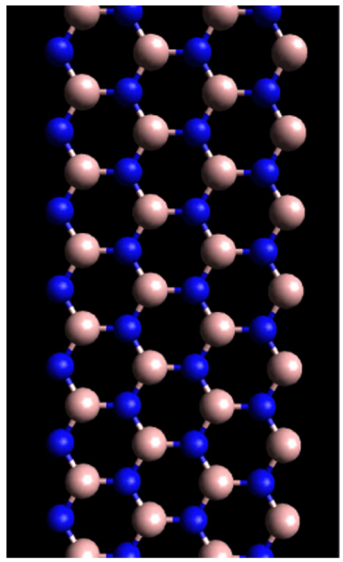

FIG. 8. Visualization of the BN nanoribbon (boron atoms in pink, nitrogen atoms in blue) investigated to test the double-cell formalism for one-dimensional systems.

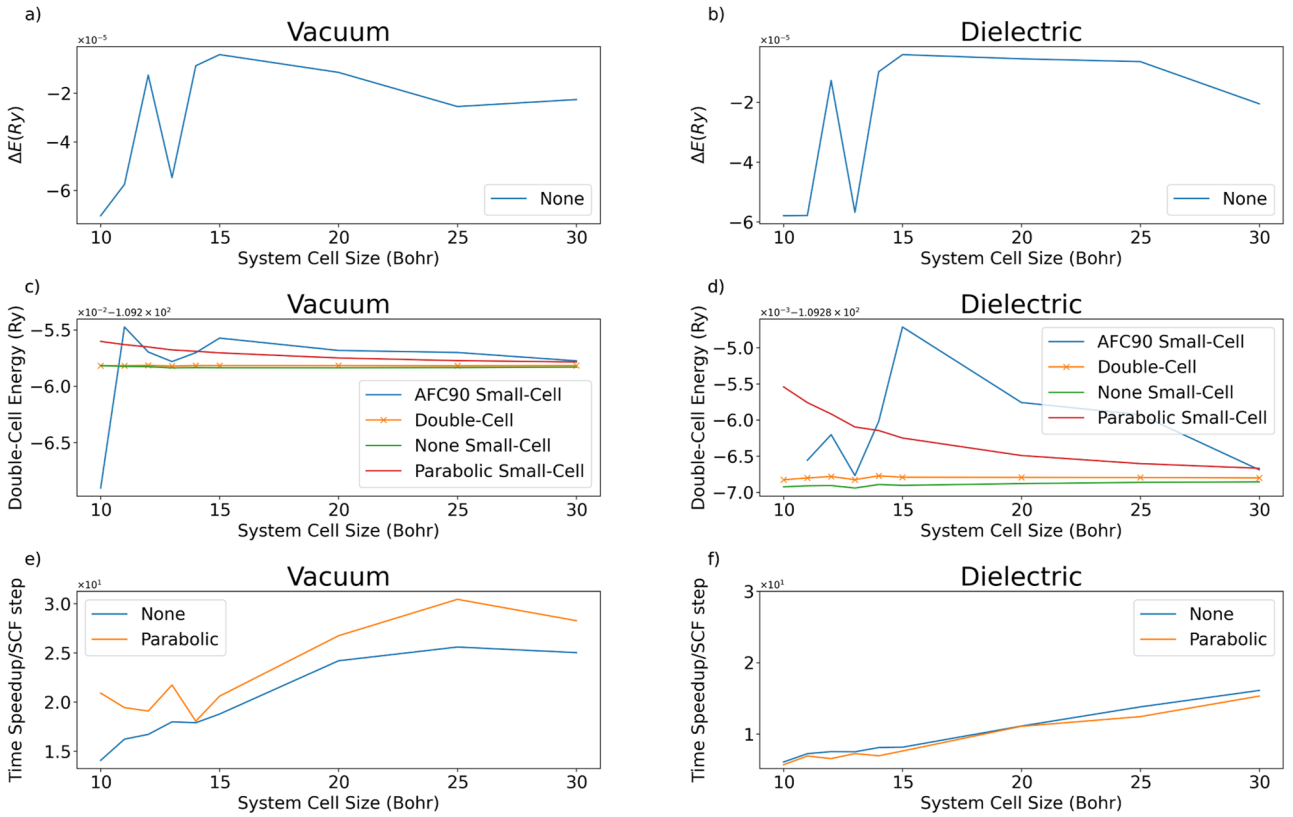

FIG. 9. Plots (a) and (b) show the energy difference, in Ry, between an SCF calculation using the double-cell formalism and the corresponding reference calculation (as visualized in Fig. 2). For plots (c) and (d), the convergence of different PBC schemes as a function of DFT system cell size is reported. Plots (e) and (f) show the time speedup between these simulations. All reported tests were performed for systems in vacuum (left column) and in a continuum dielectric medium (right column).

discussion, the sizes of the cell axes in the nonperiodic directions are kept identical, allowing to focus our analysis on a single parameter. Convergence tests are reported in Fig. 9 as a function of such a parameter.

The BN nanoribbon has a width along the y direction of 13 bohrs, with a one-atom thickness of less than 1 bohr in the z direction. Convergence of the double-cell calculations with respect to the reference setup shows minor artifact at the smallest cells considered (10 bohrs). This is consistent with our understanding of these small-cell artifacts reported for the 0D case and with the size of the studied system. While a small drift in convergence is observed at larger cell sizes, the associated errors are within the convergence threshold of the performed SCF calculations.

Similar to the results reported for 2D systems, double-cell energies appear to be well converged at all the considered system cell sizes. AFC90 results show a slightly slower convergence, with some fluctuations for cell sizes smaller than 20 bohrs. While the parabolic (point-counter-charge) correction is not implemented for 1D systems, in Fig. 9 we report the results of using a 2D scheme applied considering the y axis as the only nonperiodic direction of the system. As the system is very homogeneous along the z direction, we expect this scheme to capture most of the interaction energy between periodic replicas of the nanoribbon. Indeed, our simulations show that this parabolic correction is close to the double-cell results, although some small deviations are still present for cell sizes as large as 20 bohrs.

For this application, more substantial computational speedups are observed in using the double-cell formalism. Moreover, a linear increase in speedup is observed as a function of the cell size, consistent with the expected difference in computational scaling between the electrostatic and DFT components of the calculations. As for the 0D and 2D cases, the speedup is significantly reduced for the simulations in a dielectric medium, as the iterations involved in the solution of the more complex generalized Poisson problem increase the weight of the electrostatic calculation.

V. CONCLUSION

We report the development and implementation of a double-cell numerical algorithm to decouple electrostatic calculations and environment effects from the underlying DFT simulation. The proposed formalism was tested as a tool to remove PBC artifacts for partially periodic and nonperiodic systems. By comparing double-cell calculations with a benchmark reference, we identified a source of potential artifacts for very small system cell sizes. Indeed, if the DFT cell is so small that the electronic density of the system spills

across the cell boundaries, the mapping of the DFT cell into an expanded environment cell introduces artificial cuts in the electronic density, thus forbidding the use of the double cell for minimal DFT cell sizes. However, the presented results should show full convergence with cell size for most simulation setups and system dimensionalities. This fast convergence can be further improved by the additional use of PBC corrections on the environment cell. While a general speedup in simulation is observed by allowing the DFT cell to be smaller than the environment cell, the most impressive result was observed for the 1D system, with speedups of up to 30 times with respect to corresponding single-cell simulations. The effectiveness and accuracy of the presented approach give us confidence in the future use of this formalism for simulations with heterogeneous and nano-scaled continuum environments.

SUPPLEMENTARY MATERIAL

The supplementary material includes information on the effect of the spread value when using LIBAFCC for different system dimensionalities as well as how the double-cell formalism affects the calculated forces.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

G. Medrano: Data curation (equal); Software (equal). **E. Bainglass**: Software (equal); Supervision (equal). **O. Andreussi**: Supervision (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]S. P. Huber, S. Zoupanos, U. Martin, L. Talirz, L. Kahle, R. Häuselmann, D. Gresch, T. Müller, A. V. Yakutovich, C. W. Andersen *et al.*, "AiiDA 1.0, a scalable computational infrastructure for automated reproducible workflows and data provenance," Sci. Data **7**(1), 300 (2020).

[2]U. Martin, S. P. Huber, J. Yu, N. Marzari, and P. Giovanni, "Workflows in AiiDA: Engineering a high-throughput, event-based engine for robust and modular computational workflows," Comput. Mater. Sci. **187**, 110086 (2021).

[3]S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, G. Ceder *et al.*, "Python materials genomics (pymatgen): A robust, open-source Python library for materials analysis," Comput. Mater. Sci. **68**, 314–319 (2013).

[4]S. Curtarolo, W. Setyawan, G. L. W. Hart, M. Jahnatek, R. V. Chepulskii, R. H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy *et al.*, "AFLOW: An automatic framework for high-throughput materials discovery," Comput. Mater. Sci. **58**, 218–226 (2012).

[5]C. Draxl and M. Scheffler, "The NOMAD laboratory: From data sharing to artificial intelligence," J. Phys.: Mater. **2**(3), 036001 (2019).

[6]N. Marzari, A. Ferretti, and C. Wolverton, "Electronic-structure methods for materials design," Nat. Mater. **20**(6), 736–749 (2021).

[7]D. Padua, "Fftw," in *Encyclopedia of Parallel Computing* (Springer, 2011), p. 671.

[8]O. Andreussi and G. Fisicaro, "Continuum embeddings in condensed-matter simulations," Int. J. Quantum Chem. **119**(1), e25725 (2018).

[9]H. Lin and D. G. Truhlar, "QM/MM: What have we learned, where are we, and where do we go from here?," ChemInform **38**(22), 185–199 (2007).

[10]W. Mi, X. Shao, A. Genova, D. Ceresoli, and M. Pavanello, "eQE 2.0: Subsystem DFT beyond GGA functionals," Comput. Phys. Commun. **269**, 108122 (2021).

[11]C. R. Jacob and J. Neugebauer, "Subsystem density-functional theory," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **4**(4), 325–362 (2014).

[12]X. Shao, W. Mi, and M. Pavanello, "GGA-level subsystem DFT achieves sub-kcal/mol accuracy intermolecular interactions by mimicking nonlocal functionals," J. Chem. Theory Comput. **17**(6), 3455–3461 (2021).

[13]M. Orozco and F. Javier Luque, "Theoretical methods for the description of the solvent effect in biomolecular systems," Chem. Rev. **101**(1), 203–204 (2001).

[14]C. J. Cramer and D. G. Truhlar, "Implicit solvation models: Equilibria, structure, spectra, and dynamics," Chem. Rev. **99**(8), 2161–2200 (1999).

[15]J. Tomasi and M. Persico, "Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent," Chem. Rev. **94**(7), 2027–2094 (1994).

[16]J. Tomasi, B. Mennucci, and R. Cammi, "Quantum mechanical continuum solvation models," Chem. Rev. **105**(8), 2999–3094 (2005).

[17]O. Andreussi, I. Dabo, and N. Marzari, "Revised self-consistent continuum solvation in electronic-structure calculations," J. Chem. Phys. **136**(6), 064102 (2012).

[18]G. Fisicaro, L. Genovese, O. Andreussi, S. Mandal, N. N. Nair, N. Marzari, and S. Goedecker, "Soft-sphere continuum solvation in electronic-structure calculations," J. Chem. Theory Comput. **13**(8), 3829–3845 (2017).

[19]O. Andreussi, N. G. Hörmann, F. Nattino, G. Fisicaro, S. Goedecker, and N. Marzari, "Solvent-aware interfaces in continuum solvation," J. Chem. Theory Comput. **15**(3), 1996–2009 (2019).

[20]M. Truscott and O. Andreussi, "Field-aware interfaces in continuum solvation," J. Phys. Chem. B **123**(16), 3513–3524 (2019).

[21]E. Cancès, B. Mennucci, and J. Tomasi, "A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics," J. Chem. Phys. **107**(8), 3032–3041 (1997).

[22]L. Frediani, R. Cammi, S. Corni, and J. Tomasi, "A polarizable continuum model for molecules at diffuse interfaces," J. Chem. Phys. **120**(8), 3893–3907 (2004).

[23]S. Corni and J. Tomasi, "Enhanced response properties of a chromophore physisorbed on a metal particle," J. Chem. Phys. **114**(8), 3739–3751 (2001).

[24]O. Andreussi, S. Corni, B. Mennucci, and J. Tomasi, "Radiative and nonradiative decay rates of a molecule close to a metal particle of complex shape," J. Chem. Phys. **121**(20), 10190–10202 (2004).

[25]O. Andreussi, A. Biancardi, S. Corni, and B. Mennucci, "Plasmon-controlled light-harvesting: Design rules for biohybrid devices via multiscale modeling," Nano Lett. **13**(9), 4475–4484 (2013).

[26]R. Sundararaman, W. A. Goddard, and T. A. Arias, "Grand canonical electronic density-functional theory: Algorithms and applications to electrochemistry," J. Chem. Phys. **146**(11), 114104 (2017).

[27]R. Sundararaman and K. Schwarz, "Evaluating continuum solvation models for the electrode-electrolyte interface: Challenges and strategies for improvement," J. Chem. Phys. **146**(8), 084111 (2017).

[28]R. Jinnouchi and A. B. Anderson, "Electronic structure calculations of liquid-solid interfaces: Combination of density functional theory and modified Poisson-Boltzmann theory," Phys. Rev. B **77**(24), 245417 (2008).

[29]F. Nattino, M. Truscott, N. Marzari, and O. Andreussi, "Continuum models of the electrochemical diffuse layer in electronic-structure calculations," J. Chem. Phys. **150**(4), 041722 (2019).

[30]R. Sundararaman, K. Letchworth-Weaver, K. A. Schwarz, D. Gunceler, Y. Ozhabes, and T. A. Arias, "JDFTx: Software for joint density-functional theory," SoftwareX **6**, 278–284 (2017).

[31] J. A. Gauthier, S. Ringe, C. F. Dickens, A. J. Garza, A. T. Bell, M. Head-Gordon, J. K. Nørskov, and K. Chan, "Challenges in modeling electrochemical reaction energetics with polarizable continuum models," ACS Catal. **9**(2), 920–931 (2018).
[32] J. A. Gauthier, C. F. Dickens, S. Ringe, and K. Chan, "Practical considerations for continuum models applied to surface electrochemistry," ChemPhysChem **20**(22), 3074–3080 (2019).
[33] S. Ringe, H. Oberhofer, C. Hille, S. Matera, and K. Reuter, "Function-space-based solution scheme for the size-modified Poisson–Boltzmann equation in full-potential DFT," J. Chem. Theory Comput. **12**(8), 4052–4066 (2016).
[34] S. Ringe, H. Oberhofer, and K. Reuter, "Transferable ionic parameters for first-principles Poisson-Boltzmann solvation calculations: Neutral solutes in aqueous monovalent salt solutions," J. Chem. Phys. **146**(13), 134103 (2017).
[35] S. Ringe, N. G. Hörmann, H. Oberhofer, and K. Reuter, "Implicit solvation methods for catalysis at electrified interfaces," Chem. Rev. **122**(12), 10777–10820 (2021).
[36] Q. Campbell and I. Dabo, "Quantum-continuum calculation of the surface states and electrical response of silicon in solution," Phys. Rev. B **95**(20), 205308 (2017).
[37] Q. Campbell and I. Dabo, "Erratum: Quantum-continuum calculation of the surface states and electrical response of silicon in solution [Phys. Rev. B 95, 205308 (2017)]," Phys. Rev. B **96**(3), 039901 (2017).
[38] T. Hullar, F. C. Bononi, Z. Chen, D. Magadia, O. Palmer, T. Tran, D. Rocca, O. Andreussi, D. Donadio, C. Anastasio *et al.*, "Photodecay of guaiacol is faster in ice, and even more rapid on ice, than in aqueous solution," Environ. Sci.: Processes Impacts **22**(8), 1666–1677 (2020).
[39] F. C. Bononi, Z. Chen, D. Rocca, O. Andreussi, T. Hullar, C. Anastasio, and D. Donadio, "Bathochromic shift in the UV–visible absorption spectra of phenols at ice surfaces: Insights from first-principles calculations," J. Phys. Chem. A **124**(44), 9288–9298 (2020).
[40] O. Andreussi and N. Marzari, "Electrostatics of solvated systems in periodic boundary conditions," Phys. Rev. B **90**(24), 245101 (2014).
[41] Y. Li and I. Dabo, "Electronic levels and electrical response of periodic molecular structures from plane-wave orbital-dependent calculations," Phys. Rev. B **84**, 155127 (2011).
[42] G. Makov and M. C. Payne, "Periodic boundary conditions in *ab initio* calculations," Phys. Rev. B **51**(7), 4014–4022 (1995).
[43] P. Broqvist, A. Alkauskas, and A. Pasquarello, "Hybrid-functional calculations with plane-wave basis sets: Effect of singularity correction on total energies, energy eigenvalues, and defect energy levels," Phys. Rev. B **80**(8), 085114 (2009).
[44] F. Gygi and A. Baldereschi, "Self-consistent Hartree-Fock and screened-exchange calculations in solids: Application to silicon," Phys. Rev. B **34**(6), 4405–4408 (1986).
[45] P. E. Blöchl, "Electrostatic decoupling of periodic images of plane-wave-expanded densities and derived atomic point charges," J. Chem. Phys. **103**(17), 7422–7428 (1995).
[46] I. Hamada, M. Otani, O. Sugino, and Y. Morikawa, "Green's function method for elimination of the spurious multipole interaction in the surface/interface slab model," Phys. Rev. B **80**(16), 165411 (2009).
[47] N. D. Hine, K. Frensch, W. M. Foulkes, and M. W. Finnis, "Supercell size scaling of density functional theory formation energies of charged defects," Phys. Rev. B **79**(2), 024112 (2009).
[48] P. A. Schultz, "Charged local defects in extended systems," Phys. Rev. Lett. **84**(9), 1942–1945 (2000).
[49] P. A. Schultz, "Theory of defect levels and the 'band gap problem' in silicon," Phys. Rev. Lett. **96**(24), 246401 (2006).
[50] I. Dabo, B. Kozinsky, N. E. Singh-Miller, and N. Marzari, "Electrostatics in periodic boundary conditions and real-space corrections," Phys. Rev. B **77**(11), 115139 (2008).
[51] G. J. Martyna and M. E. Tuckerman, "A reciprocal space based method for treating long range interactions in *ab initio* and force-field-based calculations in clusters," J. Chem. Phys. **110**(6), 2810–2821 (1999).
[52] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni *et al.*, "Advanced capabilities for materials modelling with quantum espresso," J. Phys.: Condens. Matter **29**(46), 465901 (2017).
[53] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, L. GuidoChiarotti, M. Cococcioni, I. Dabo *et al.*, "QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials," J. Phys.: Condens. Matter **21**(39), 395502 (2009).
[54] E. Bainglass, M. Gabriel, M. Truscott, Q. Campbell, I. Dabo, I. Timrov, N. Marzari, and O. Andreussi, "Environ 3.0: A continuum solvation package" (unpublished).
[55] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," Phys. Rev. Lett. **77**(18), 3865–3868 (1996).
[56] G. Prandini, A. Marrazzo, I. E. Castelli, N. Mounet, and N. Marzari, "Precision and efficiency in solid-state pseudopotential calculations," npj Comput. Mater. **4**(1), 72 (2018).
[57] K. Lejaeghere *et al.*, "Reproducibility in density functional theory calculations of solids," Nature **351**(6280), aad3000 (2016).
[58] G. Fisicaro, L. Genovese, O. Andreussi, N. Marzari, and S. Goedecker, "A generalized Poisson and Poisson-Boltzmann solver for electrostatic environments," J. Chem. Phys. **144**(1), 014103 (2016).
[59] N. Marzari and D. Vanderbilt, "Maximally localized generalized Wannier functions for composite energy bands," Phys. Rev. B **56**(20), 12847–12865 (1997).