Equivalence of dipole correction and Coulomb cutoff techniques in supercell calculations

Liping Yu,1 V. Ranjan,1 W. Lu,1,2 J. Bernholc,1,2 and M. Buongiorno Nardelli1,2
1Center for High Performance Simulation and Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA
2CSMD, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

(Received 21 March 2008; revised manuscript received 2 May 2008; published 3 June 2008)

In ab initio calculations of surfaces or nonperiodic systems, one frequently relies on the supercell approximation, where the periodic replicas of the system are separated by enough empty space to avoid spurious interactions between the successive images. However, a vacuum separation is not sufficient to screen the dipolar interaction that appears in asymmetrically charged or polar systems. The dipole correction and Coulomb cutoff methods are often used to eliminate such interactions between the periodic replicas. In this work, these methods are compared under the same conditions in the framework of plane-wave based density-functional theory. The dipole correction method is shown to be equivalent to the rigorous Coulomb cutoff formalism in the calculations of total energy, force, charge density, and self-consistent potential. We demonstrate that the band structures obtained by these methods coincide for the localized bound states and that the corrections have essentially no influence on the occupied energy bands, only substantially affecting the unoccupied bands. By comparing the results of the two methods, the localized bound states of interest can be easily distinguished from the highly delocalized unoccupied states using a relatively small supercell. This comparison offers substantial savings in the computational time when ascertaining convergence with supercell size. The accuracy of the dipole correction method is also confirmed by comparing the results for a model ferroelectric BaTiO3 slab with a Berry-phase calculation of polarization for the bulk system.

I. INTRODUCTION

The plane-wave method with Born–von Kármán periodic boundary conditions (PBCs) is one of the most important state-of-the-art techniques in electronic structure calculations for systems ranging from bulk solids to molecules and surfaces.1 The systems are considered to be repeated periodically in three dimensional space and the supercell approximation is commonly adopted when nonperiodic systems such as surfaces, nanostuctures, or molecules are investigated. When necessary, e.g., for modeling surfaces, a vacuum region is inserted into the supercell to separate the periodic replicas along the aperiodic direction(s). However, the supercell approximation artificially introduces spurious dipole interactions between the periodic images when there is a net dipole moment or charge in the cell.2 The dipole interactions lead to results that display strong dependence on cell size and shape, possibly introducing artifacts into the calculations.

The issue of eliminating the undesired interactions between the periodic images has been extensively discussed in the literature. Several different methods have been proposed:3–17 Makov and Payne6 corrected the electrostatic energy functional (but not the potential) for an aperiodic system by adding a dipole dependent term in a cubic supercell. Kantorovich16 generalized this method to cells of arbitrary shape. Another kind of method obtains the electrostatic potential by subtracting out the countercharge matching multiple moment of the charge distribution and solving for that part using boundary conditions of an isolated system.11,16 A different approach is to introduce a modified Coulomb potential with a spatial cutoff.7,14,15 This Coulomb cutoff formalism can remove all spurious image interactions introduced by the long-range part of the Coulomb potential, including dipole and all higher order multipole interactions. It has been demonstrated to be exact and rigorous for clusters,7,8 slabs,12 and even wires.14 For polar slabs, a simple to implement dipole-correction method, which introduces a correcting dipole in the vacuum region, has been widely used in calculations.18,19 It was originally proposed by Neugebauer and Scheffler,5 and later corrected by Bengtsson.9

Unfortunately, the above methods have never been consistently compared to each other. It is therefore difficult to know which one is the most suitable for the system at hand. In particular, it is not yet clear whether the simple dipole correction5,9 is sufficiently effective in cancelling the spurious interactions due to the PBCs in planar polar slabs. In this paper, we compare the dipole correction method9 with the exact Coulomb cutoff formalism12,14 that we implemented in the plane-wave based density-functional code PWSCF,20 in which the dipole correction method was already coded as an option. We studied the convergence of these two techniques via computation of the total energy, force, charge density, potential, and electron band structure. Our results demonstrate that the dipole correction method is equivalent to the Coulomb cutoff formalism. As such, both methods can completely cancel the spurious interactions between the periodic images in a polar slab system. As already noted in Ref. 14, we find that the corrections to the band structure substantially affect the unoccupied electronic states but have negligible influence on the occupied ones. This finding is important when studying optical or excited state properties of polar slabs. In fact, as will be described below, comparisons between the two methods provide a cost-effective avenue for confirming the convergence of the results with respect to the supercell size.
II. COULOMB CUTOFF FORMALISM AND DIPOLE CORRECTION TECHNIQUE

In the supercell approach, the periodic images of the system interact with each other through the long-range Coulomb potential, i.e., the electrostatic potential $V(\mathbf{r})$ that satisfies the Poisson equation with PBCs. It has the form of a convolution,

$$V(\mathbf{r}) = \int_{\text{all space}} n(\mathbf{r}') \delta^3 \mathbf{r}' = \int_{\text{all space}} n(\mathbf{r}') v(|\mathbf{r} - \mathbf{r}'|) d\mathbf{r}' . \quad (1)$$

In reciprocal space for a 3D periodic system with reciprocal lattice vectors $\mathbf{G}$, Eq. (1) can be written using the convolution theorem as

$$V(\mathbf{G}) = n(\mathbf{G}) v(\mathbf{G}) , \quad (2)$$

where $n(\mathbf{G})$ and $v(\mathbf{G})$ are the Fourier transforms of the charge density $n(\mathbf{r})$ and long-range Coulomb potential $v(\mathbf{r})$.

To remove all undesired interactions between the periodic replicas in the Coulomb cutoff formalism, one replaces Eq. (2) with a modified one

$$\tilde{V}(\mathbf{G}) = \tilde{n}(\mathbf{G}) \tilde{v}(\mathbf{G}) , \quad (3)$$

where the Fourier integral

$$\tilde{v}(\mathbf{G}) = \int_{\text{all space}} \tilde{v}(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} d^3 \mathbf{r} \quad (4)$$

corresponds to a modified Coulomb interaction

$$\tilde{v}(|\mathbf{r} - \mathbf{r}'|) = \begin{cases} \frac{1}{|\mathbf{r} - \mathbf{r}'|} & \text{for } |z - z'| < L \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

The modified charge density $\tilde{n}(\mathbf{r})$ is still 3D periodic but the range of the cutoff parameter $L$ must be equal to the supercell dimension in the aperiodic direction and at least twice the slab thickness. Here, the slab thickness is defined as the distance between the two surfaces outside the slab where charge densities are essentially zero, and not the distance between the top and bottom atomic surfaces of the slab. By choosing $L$ and the slab thickness in this way, the densities belonging to the different slab images will not overlap and cannot interact through $\tilde{v}(\mathbf{G})$.

In plane wave based density-functional theory (DFT) calculations, the electrostatic energy is decomposed into three terms: Hartree, ion–electron, and ion–ion interactions. Within the Coulomb cutoff formalism, each term has an analytic form in reciprocal space for a 2D slab system (refer to Refs. 12 and 14 for the exact expressions). The divergent terms at $G=0$ are shown to cancel exactly when the cutoff is applied to both the ionic and Hartree potentials. The regular spherical energy cutoff can still be applied after separating the surface terms that need a higher cutoff $G_c$ and special care (see discussion in Appendix B of Ref. 12).

The dipole correction method is based on the observation that the electrostatic potentials on the two sides of an isolated slab with a nonzero net dipole moment $m$ are different and the use of PBCs introduces an artificial uniform electric field in the supercell that cancels the potential jump across the slab. To eliminate the unwanted electric field, one introduces a dipole layer in the midvacuum region of the supercell. Furthermore, a true external electric field $E_{\text{ext}}$ can also be applied to the slab in the same way by adjusting the strength of the additional sawtoothlike potential.$^{18,21,22}$

$$V^{\text{dip}}(z) = -e \left( \frac{4\pi m}{c} - E_{\text{ext}} \right) z , \quad \frac{c}{2} < z < \frac{c}{2} . \quad (6)$$

The dipole corrected potential $V(\mathbf{r})$, total energy $E_{\text{tot}}$, and Hellmann–Feynman forces $F_I$ become,

$$V(\mathbf{r}) = V^{\text{dip}}(z) + V^0(\mathbf{r}) , \quad (7)$$

$$E_{\text{tot}} = E_{\text{tot}}^0 + \left( \frac{2\pi m}{c} - E_{\text{ext}} \right) Am , \quad (8)$$

$$F_I = F_I^0 + eZ_I \left( E_{\text{ext}} - \frac{4\pi m}{c} \right) \hat{z} , \quad (9)$$

where $V^0(\mathbf{r})$, $E_{\text{tot}}^0$ and $F_I^0$ are the calculated electrostatic potential, total energy, and force without dipole correction, $Z_I$ is the ionic charge of the $I$-th atom, and $A$ is the area of the surface unit cell.

We should stress that the Coulomb cutoff and dipole correction methods correspond to two different physical procedures rather than mathematical transformations of the same expression. The former is considered to be exact, rigorous, and able to remove all spurious interactions between the periodic images. However, it is more complex to implement and needs modifications for including nonzero electric field. In contrast, the dipole correction can simultaneously cancel the undesired artificial electric field due to the PBCs and include an external electric field applied to a slab but it is difficult to ascertain whether the artificial electric field completely cancels all interactions between the periodic replicas. The two methods may also differ in their convergence properties with respect to the supercell size. In order to investigate these issues, it is necessary and fruitful to compare these methods under the same conditions and for the same physical system.

To accomplish this, we implement the Coulomb cutoff formalism in the PWSCF code, where the dipole correction method was already programmed. We are thus able to perform calculations with no correction, dipole correction, and Coulomb cutoff, in the same supercell and with the same inputs (such as cutoff energy, $k$ points, pseudopotentials, etc.). The slab was centered in the middle of the supercell along the $z$ direction and the supercell size was chosen to be large enough so that the Coulomb cutoff formalism exactly removes all the undesired interactions. Our
calculations use the cutoff energy of 30 Ry, Vanderbilt ultrasoft pseudopotentials\(^\text{24}\) with Perdew–Zunger (LDA) parametrization of the exchange correlation functional\(^\text{25}\) and an in-plane Monkhorst–Pack \(k\)-point grid of \(8 \times 8\).\(^\text{26}\)

III. RESULTS AND DISCUSSION

As a first test of the implementation of the Coulomb cutoff formalism in the PWSCF code, we computed the total energy and the electronic band structure of a nonpolar Na sheet in a supercell of size \(7.6 \times 7.6 \times 19\) a.u. As expected, the total energy from the calculations with and without the Coulomb cutoff are the same, i.e., \(-0.651\,110\) Ry, because the sheet is nonpolar. The electronic band structure, which agrees very well with that in Ref.\(^\text{14}\), shows that the occupied states obtained with the Coulomb cutoff correction applied are essentially the same as the ones obtained without the correction while the unoccupied bands start to differ as their energies increase [see Fig. 1(a)]. In Ref.\(^\text{14}\), the authors also found the same effect for a Si chain and stated it is due to greater delocalization of states at higher energies.

To understand this effect more explicitly, we plot in Fig. \(2(a)\) the planar average of the total potential and squares of the wave functions of the three lowest band states at \(\Gamma\). It shows that the supercell of 19 a.u. thickness is too small for the potential to become flat in the vacuum region. In this small cell, the third lowest \(\Gamma\) state does not decay to zero in a vacuum, which inevitably leads to overlap interactions between the states belonging to neighbor slab images. This state then behaves like a bulk one and the Coulomb cutoff correction is no longer valid. It changes both the bulklike wave function and its band energy. When calculations are performed in a larger supercell (76 a.u.) with more vacuum, the flat vacuum potential around the slab is recovered [see Fig. \(2(b)\)] and the band structure difference between the results with and without the Coulomb cutoff correction disappears [see Fig. \(1(b)\)]. In this large supercell, the third lowest \(\Gamma\) state is well localized within about 25 a.u. from the slab and the Coulomb correction does not affect it.

From Figs. 1 and 2, we conclude that the differences between the results obtained with and without the Coulomb cutoff only occur for those unoccupied states that overlap with their periodic images due to the limited size of the vacuum region. These differences disappear in a very large supercell, as they should, since there are no dipole interactions between nonpolar repeated slab images.

For a more meaningful comparison between the two techniques, we now turn to a highly polar slab: a layer of water molecules in a supercell geometry, which is for direct comparison with a previous investigation.\(^\text{9}\)

Table I shows the total energies, forces, and dipole moments for a water-molecule sheet in a supercell of size \(5.67 \times 5.67 \times 23\) a.u., for the cases of no correction, Coulomb cutoff, and dipole correction, respectively. The dipole moment is calculated from the electron charge density and the point charges of ions. As a reference, we also performed the same calculation in a nonpolar double cell, where we positioned two water slabs with dipoles in the opposite directions so that no net dipole moment exists in the simulation cell. The results show that both the Coulomb cutoff and the dipole correction give essentially the same total energy, forces, and dipole moment, which also agree very well with the results of the double cell calculation where no correction is needed.

Figure 3 shows the convergence of the total energy with respect to the supercell size \(c\) (\(a\) is fixed at 5.67 a.u.) with and without the corrections. At supercell sizes \(c/a \geq 3\), the Coulomb cutoff and the dipole corrections are essentially converged while without a correction, the convergence is much slower. This behavior agrees well with what is reported.

---

**TABLE I. Calculated total energy, force on the oxygen atom, and dipole moment per unit area for a water-molecule sheet in a supercell of size \(5.67 \times 5.67 \times 23\) a.u.**\(^\text{3}\) See text.

<table>
<thead>
<tr>
<th></th>
<th>Total Energy (Ry)</th>
<th>(F_z) (Ry/a.u.)</th>
<th>Dipole moment (e/a.u.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No correction</td>
<td>(-34.335063)</td>
<td>0.014020</td>
<td>0.018996</td>
</tr>
<tr>
<td>Dipole correction</td>
<td>(-34.329478)</td>
<td>0.014475</td>
<td>0.016516</td>
</tr>
<tr>
<td>Coulomb cutoff</td>
<td>(-34.329477)</td>
<td>0.014497</td>
<td>0.016518</td>
</tr>
<tr>
<td>Double cell</td>
<td>(-34.329482)</td>
<td>0.014355</td>
<td></td>
</tr>
</tbody>
</table>

---

FIG. 1. (Color online) Band structures of Na sheet with and without the Coulomb cutoff correction in supercell thicknesses of (a)\(19\) and (b)\(76\) a.u. The Fermi energies in the two supercells have been aligned.

FIG. 2. (Color online) Plane averaged total potential (thick black curves) along the normal direction of the Na-layer slab with no correction in supercells with thicknesses of (a) \(19\) and (b) \(76\) a.u. The three lowest energy bands at the \(\Gamma\) point and their corresponding electron densities without the correction and with the Coulomb cutoff correction are also shown.
for the Coulomb cutoff formalism in Refs. 7 and 14. In supercells with sizes $c/a < 3$, the Coulomb cutoff correction is not converged while the dipole correction is essentially converged at $c/a = 2$. As mentioned in Sec. II, the supercell size for the Coulomb cutoff method should be at least twice the slab thickness. The thickness of the water sheet is about 1.3$\alpha$; therefore, the Coulomb cutoff cannot eliminate all spurious interactions between the replicas for supercells smaller than $c = 2.6\alpha$.

In Fig. 4, we plot the planar average of the electrostatic potential in the $z$ direction for a water sheet in a supercell of size $5.67 \times 5.67 \times 23$ a.u. Except in the midregion of the vacuum, the dipole correction and Coulomb cutoff methods give essentially the same electrostatic potential. Both potentials introduce characteristic dipole layers in the vacuum, which allows for the proper extraction of work functions of the two inequivalent slab surfaces. As evident from the figure, the dipole potential in the Coulomb cutoff method has a shape that is similar to the electrostatic potential of the slab but with an opposite sign. This is why the size of the supercell has to be at least twice the slab thickness. In contrast, the simple steplike potential in the dipole correction method allows for a significantly smaller vacuum region and, thus, a smaller supercell. The dipole method should thus be preferred for large slab systems.

The calculated band structures of the H$_2$O sheet are shown in Fig. 5 for the cases of (i) Coulomb cutoff, (ii) dipole correction, and (iii) no corrections. Similar to the case of the nonpolar Na sheet (Fig. 1), we observed differences only in the unoccupied states. In Fig. 5(a) for a small supercell, the Coulomb cutoff and dipole corrections agree for all occupied bands and the lowest unoccupied band within the whole Brillouin zone, and for the second lowest unoccupied band at $k$ points between $X$ and $M$, while disagreeing for all other states. For convenience, we call the former coincident and the latter noncoincident states. By examining their wave functions, we find that all coincident states are well-localized bound states that decay to zero in the vacuum of the supercell [see the $i = 1$ curve in Fig. 6(a) for an example] while all noncoincident states are delocalized and spread over the whole supercell [see the $i = 3$ curve in Fig. 6(a) for an example]. As a result, the two methods here do not work accurately for the overlapping noncoincident states and behave differently.

In a larger supercell ($c = 113$ a.u.), the noncoin-
cident states in Fig. 5(a) become coincident states and both correction techniques coincide again.

The comparison between Figs. 5(a) and 5(b) shows that the noncoincident states in the small supercell \((c=23\ \text{a.u.})\) are far from convergence and cannot be taken as correct band states. It follows that to obtain a converged band structure, especially for the unoccupied states, one needs a huge supercell even with the Coulomb cutoff or dipole correction. Moreover, there also exist unbound noncoincident continuum states (not shown in the figure), which have higher energies and, in principle, need an infinitely large supercell to represent them. Strictly speaking, it is therefore impractical to calculate the high-energy band structure for an isolated polar slab using the supercell technique. However, in this work, we are interested in low-energy unoccupied noncoincident states, which are well localized near the slab. The corrections made to the coincident states by both techniques are substantial in the small supercell but become much smaller as the supercell size increases. For instance, the difference between the direct band gaps at \(E\) calculated with \((2.23\ \text{eV})\) and without \((2.73\ \text{eV})\) the correction is about 0.5 eV in the small supercell [Fig. 5(a)] but it is about 0.14 eV in the large supercell [Fig. 5(b)]. This difference occurs because the empty coincident states are still spread out in the vacuum and are therefore substantially affected by the electric field in the vacuum region, which is due to the periodic boundary conditions. The electric field obviously decreases as the length of the supercell is increased. After removing this field by either the Coulomb cutoff or the dipole correction, it is amazing to find that the corrected coincident bands in this small supercell \([c=23\ \text{a.u.}, \text{see Fig. 5(a)}]\) are almost identical to the ones calculated in a much larger supercell \([c=113\ \text{a.u.}, \text{see Fig. 5(b)}]\) and thus can be taken as converged results.

The above results show that the convergent results for well-localized bound states can be obtained with either the Coulomb cutoff or the dipole correction in a relatively small supercell. To distinguish these states from highly delocalized or unbound continuum empty states, one can simply perform calculations with both techniques and compare the respective results. In this way, one can confirm the convergence for the states of interest without having to carry out calculations for several supercell sizes, thereby saving large amounts of computer time.

To assess the computational cost of the two corrections, we show in Table II the average CPU time needed to run a single self-consistent iteration for a test system. The corrections increase only slightly the computational time compared to the “no correction” case and they both cost about the same.

In general, our tests have shown that the dipole correction and the Coulomb cutoff techniques produce almost identical results for the total energy, force, charge density, electrostatic potential, and the band structure. Therefore, we can conclude that the addition of a self-consistent dipole to screen the interaction between the periodic replicas of a polar supercell leads to very accurate results. Actually, the additional external dipole layer in the vacuum can be seen as an effective capacitor. The charge on the plates of the capacitor produces an electric field that compensates the artificial electric field due to the PBCs imposed on the polar slab. The external field introduces a potential drop in the vacuum region where the wave functions are essentially zero and, thus, are not affected. Consequently, the whole supercell has a zero net dipole moment and all periodic interactions are effectively screened.

In order to confirm even more strongly the validity of the two techniques, we decided to follow an approach similar to that of Fu et al.,\textsuperscript{27} who carried out a Hartree–Fock (HF) calculation for genuinely isolated slabs with a localized basis set. They calculated the bulk spontaneous polarization \(P_s\) of ferroelectric BaTiO\(_3\) from a bulk truncated slab with ideal ferroelectric distortions. Their HF result of 0.245 C/m\(^2\) agrees well with the bulk value of 0.240 C/m\(^2\) obtained independently with a Berry-phase calculation,\textsuperscript{28} demonstrating that indeed one can recover the theoretical bulk value of polarization using genuinely isolated slabs.

However, a similar isolated slab calculation in the density-functional theory is hampered by the fact that the DFT-LDA gap in BaTiO\(_3\) is much too small. Due to the small band gap and the high value of polarization in the bulk-truncated BaTiO\(_3\) slab, Zener (interband) tunneling occurs.\textsuperscript{29,30} This results in metallicity for polar slabs thicker than three atomic layers and, thus, in charge transfer from one side to the other. This transfer reduces the depolarization field and eventually leads to zero polarization as the thickness of the slab increases. Hence, the bulk spontaneous polarization derived from a planar metallic slab is zero. This behavior is in contrast with the HF result where the calculated band gap is much larger.\textsuperscript{31} To circumvent this problem, we created fictitious bulk and slab structures of BaTiO\(_3\) in which the ferroelectric distortions are artificially reduced to 1/10 of the physical ones. In this fictitious system, the corresponding polarization and depolarization fields are also reduced and, thus, the slabs used in our DFT-LDA calculation for deriving bulk polarization are still insulators. We have performed polarization calculations with the reduced distortions using DFT-LDA with dipole correction. From the slab geometry, we obtain \(P_s=0.0230\ \text{C/m}^2\), which is in excellent agreement with the value of 0.0238 C/m\(^2\) derived from the Berry-phase calculations. Therefore, we can conclude that the dipole correction (or equivalently, the Coulomb cutoff formalism) yield results that are equivalent to the ones acquired for isolated slabs in a localized basis scheme.\textsuperscript{27}

### IV. CONCLUSIONS

This work shows that the dipole correction and the Coulomb cutoff formalism techniques are equally effective in...
eliminating the spurious interactions between the periodic images of polar supercells. The results of either technique are equivalent to the ones that would be obtained for isolated slabs, as previously demonstrated in HF calculations with a localized orbital basis. In general, we conclude that both techniques have different relative strengths, which can be exploited in different contexts: The Coulomb cutoff formalism can be applied not only to slab geometries but also for arbitrary shape nanostructures. However, for slab geometries, the dipole correction is preferred because a smaller supercell is sufficient for achieving convergence. Moreover, electric fields can be trivially included in the dipole correction method.

Finally, it is found that both corrections have no influence on the occupied band states and affect only unoccupied bands. Moreover, by comparing the results of the dipole and Coulomb cutoff corrections, the highly delocalized unoccupied states can be easily distinguished from localized bound states in a relatively small supercell calculation. This identification is important when determining excited states and optical properties from the slab calculations. In general, careful convergence studies should be performed when calculating the properties involving unoccupied electronic states. Judicious comparisons between the results of the two corrections make this task much easier and less costly.

ACKNOWLEDGMENTS

We gratefully acknowledge the fruitful discussions with P. Giannozzi. This work was supported in part by the Office of Naval Research, the Department of Energy, and the National Science Foundation. The calculations were performed at DOD ARSC center and at the NCCS facility at ORNL.

2Note that the higher order multipole corrections are negligible in a polar slab system. The quadrupole–quadrupole interaction decays as $L^{-5}$, where $L$ is the linear dimension of a supercell. It is clearly negligible for supercells as large as the ones used in typical calculations (Ref. 6). Moreover, it has been shown that adding the quadrupole corrections does not noticeably improve the accuracy of the results (Ref. 13).
20P. Giannozzi et al., http://www.pwscf.org