PHYSICAL REVIEW B **102**, 045403 (2020)

Generalized dipole correction for charged surfaces in the repeated-slab approach

Christoph Freysoldt, Arpit Mishra, Michael Ashton, and Jörg Neugebauer
Max-Planck-Institut für Eisenforschung GmbH, Max-Planck-Straße 1, 40237 Düsseldorf, Germany

(Received 8 May 2020; revised 16 June 2020; accepted 17 June 2020; published 1 July 2020)

First-principles calculations of surfaces or two-dimensional materials with a finite surface charge invariably include an implicit or explicit compensating countercharge. We show that an ideal constant-charge counterelectrode in the vacuum region can be introduced by means of a simple correction to the electrostatic potential in close analogy to the well-known dipole correction for charge-neutral asymmetric slabs. Our generalized dipole correction accounts simultaneously for the sheet-charge electrode and the huge voltage built up between the system of interest and the counterelectrode. We demonstrate its usefulness for two prototypical cases, namely, field evaporation in the presence of huge electric fields (20 V/nm) and the modeling of charged defects at an insulator surface. We also introduce algorithmic improvements to charge initialization and preconditioning in the density functional theory algorithm that proved crucial for ensuring rapid convergence in slab systems with high electric fields.

DOI: 10.1103/PhysRevB.102.045403

I. INTRODUCTION

Charges at surfaces, interfaces, and two-dimensional (2D) materials play an important role for creating or tuning macroscopic electric fields in a variety of technical applications. Examples are electronic and optoelectronic devices, chemical sensors, or electrochemistry and corrosion. The origin and distribution of charges at the atomic scale, as well as their effect on chemical bonding and chemical reactions, are intimately linked to the electronic structure. It is therefore highly desirable to simulate charged surfaces with electronic-structure theory, notably density functional theory [1]. However, treating nominally charged systems with feasible system sizes (typically a few hundred atoms) remains challenging because long-range Coulomb interactions can introduce significant artifacts. In particular, the energy changes associated with forming and charging surfaces depend on the location of the countercharge, i.e., on the specific experimental scenario [2–6]. Truncating the Coulomb interactions in the simulation [7–9] alleviates the convergence issues encountered with the implicit homogeneous background approach, but the countercharge setup corresponding to the truncation requires in general further modeling considerations to connect to the situation of interest.

For charged point defects, it has become common practice to extrapolate to the dilute limit of vanishing space charge via corrections [6,10–15] or scaling considerations [16,17]. For surfaces, which are typically modeled in the repeated-slab approach, this requires an estimate of the electrostatic energy arising from the chosen configuration of the countercharge, namely, whether it is homogeneously distributed across the cell [14,15] or confined to the slab [6]. However, the effect of an artificial field on the calculated properties remains unclear. There are still other cases where a finite field is actually required, e.g., for field-induced reconstructions or field evaporation [4,18–20]. It is therefore desirable to introduce a computational counterelectrode facing the charged surface as an alternative approach [cf. Fig. 1(a)]. This has been done previously using a Gaussian sheet charge [18,19]. These investigations employed symmetric slabs because for asymmetric slabs the field on either side of the slab cannot be controlled easily. Alternatively, one can achieve this by imposing special boundary conditions [3]. Here, we present a conceptually different solution to introduce a computational counterelectrode with full control over the field on both sides. The concept is based on an easy to implement correction to the standard periodic-boundary potential with implicit (homogeneous) countercharge. The approach is in close analogy to the well-established dipole correction for charge-neutral slabs [21], which introduces a sheet dipole layer in the vacuum to compensate for the intrinsic dipole of asymmetric slabs. We therefore name our scheme "generalized dipole correction."

The remainder of this paper is organized as follows. In Sec. II, we summarize the idea and working equations of the generalized dipole correction. Some methodological improvements of the DFT algorithms relevant for actual calculations are presented in Sec. III. The application of the method for field-dependent evaporation curves and charged defect calculations are presented in Secs. IV A and IV B, respectively. In the Appendix, we exemplify for the used SPHINX code [22,23] which settings must be made in a typical DFT input file. A complete example input as well as a jupyter notebook within the pyiron framework [24] is provided as Supplemental Material [25].

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

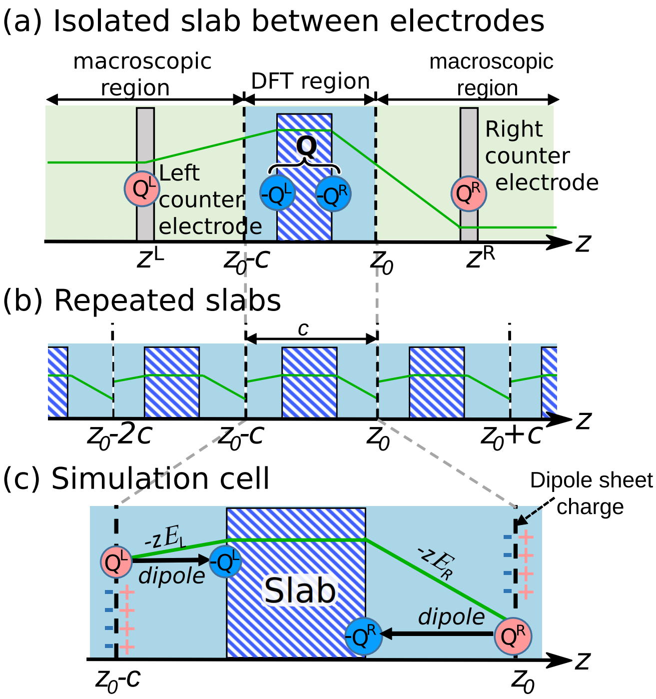

FIG. 1. Sketch of the potential (a) of a charged metallic slab with counterelectrodes, (b) in the repeated-slab approach including the generalized dipole correction, (c) the DFT region. The potential discontinuity at the cut position z_0 corresponds to a sheet dipole layer that compensates the net dipole of the simulation cell (sum of arrows).

II. GENERALIZED DIPOLE CORRECTION

A. Idea

The main motivation for using the repeated-slab approach is to employ the highly efficient computer codes and algorithms available for crystalline solids. These codes are designed to deal with periodic-boundary conditions in *three* dimensions. By adding enough empty space (vacuum) between the slabs and by making the slabs sufficiently thick, the presence of the periodic images and of the backside surface can be safely ignored for the surface of interest. The main goal of our generalized-dipole correction is therefore to decouple the simulation cell electrostatically from its periodic images, and to link the energetics to a macroscopic system with external electrodes as sketched in Fig. 1(a). For simplicity of notation in the following, we will set the coordinate system such that the slab is parallel to the x-y plane. The electric field is parallel to the z axis. The surface unit-cell area will be denoted by A. c is the size of the repeated-slab cell in the third dimension, i.e., along the z axis. As we plot the z axis horizontally from left to right in all diagrams, we will use the words “left” as synonym for “toward decreasing z values,” and “right” for “toward increasing z values” throughout this work. We will further assume that the surface of interest is on the “right” side.

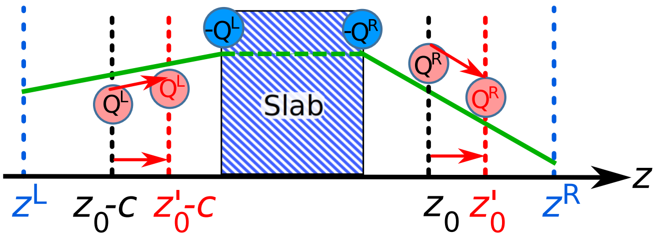

FIG. 2. Potential and charges within the DFT region (cf. Fig. 1) for different cut positions. Changing the position of the DFT boundaries ($z_0 \rightarrow z'_0$) should not change the potential near the slab. Real charges placed at z_0 would incur changes in energy.

The macroscopic system sketched in Fig. 1(a) consists of the charged slab and two ideal counterelectrodes. The total system is charge neutral, i.e., the slab charge is compensated by the two electrodes. The charges per area A on each of the two electrodes are Q^L and Q^R, respectively. They define the electric field between the respective electrode and the slab via

$$\mathcal{E}^L = +\frac{4\pi}{A} Q^L, \tag{1}$$

$$\mathcal{E}^R = -\frac{4\pi}{A} Q^R. \tag{2}$$

As the electronic density rapidly drops toward zero outside the slab, we can safely restrict the quantum-mechanical [density functional theory (DFT)] treatment to a small region around the slab, which will take notice of the external electrodes by the electrostatic boundary conditions. Apart from the electrostatic boundary conditions, the DFT region can be safely approximated with a standard repeated-slab approach, as the kinetic energy, short-range electrostatics, exchange-correlation effects, and the ionic potentials are very short ranged compared to macroscopic dimensions. The first task is therefore to reconcile the desired potential shape for the DFT part with periodic-boundary conditions.

To this end, we introduce, similar to what is done in the dipole correction [21], a cut in the vacuum region at z_0, where the field and potential change discontinuously [see Fig. 1(b)]. The potential between the cut left to the slab and the cut right to it [Fig. 1(c)] should then be equivalent to that of the slab sandwiched between two electrodes [Fig. 1(a)]. If, as in Fig. 1, the fields on the left side $\mathcal{E}^L$ and right side $\mathcal{E}^R$ shall differ, a corresponding charge (in Hartree atomic units)

$$Q = \frac{A}{4\pi}(\mathcal{E}^R - \mathcal{E}^L) = -(Q^L + Q^R) \tag{3}$$

must sit on the slab per surface unit-cell area A. This condition is independent of the slab’s screening properties, i.e., whether it is metallic (as sketched in Figs. 1 and 2) or behaves as a dielectric. The main difference between a metal and a dielectric is how the total charge is distributed within the slab: in a metal, the surface charges on either side will exactly compensate the external electrode charges they are facing in order to ensure that the slab is field-free inside. In a dielectric, there is no such guarantee. This internal distribution of the excess charge, however, does not affect the formalism. Indeed, the total polarization of the slab, the only relevant aspect for the method, is directly accounted for by the total dipole moment. Regarding the external distributions of fields, a particular important case is $\mathcal{E}^L = 0$, as this suppresses any changes from the left surface compared to a field-free neutral slab.

Reversely, any charged slab produces such a field difference in open boundary conditions.

B. Potential

However, the repeated-slab approach is typically used for codes which do not implement *open*-boundary conditions, but *periodic*-boundary conditions (notably to employ plane-wave basis sets). Solving the Poisson equation of a nominally charged system with standard periodic-boundary conditions, e.g., by means of the Fourier method and setting the $\mathbf{G} = 0$ component to zero, is equivalent to implicitly including a homogeneous compensating background [11,12]. Correcting from the periodic-boundary solution V^{pbc} to the desired potential shape sketched in Fig. 1

$$V^{\text{tot}} = V^{\text{pbc}} + V^{\text{corr}} \tag{4}$$

requires a correction potential V^{corr} that (1) accounts for the sheet charge of both counterelectrodes at the cut position, (2) removes the implicit compensating background, (3) fixes the potential jump at the cut position such that the left and right fields attain to their desired values, and (4) shifts the potential by a spatially constant offset to follow some alignment convention.

Formally, the corrections to the potential could be introduced by adding suitable countercharges in the vacuum region. However, correcting the potential rather than introducing explicit countercharges and using the Fourier method to produce the desired potential shape has two advantages. First, in order to get a converged Fourier series the countercharges have to be smeared out [4,18]. The smeared-out charge distribution introduces an additional artificial self-energy compared to an ideal, flat electrode. This self-energy must be accounted for to arrive at an artifact-free total energy [4]. In general, the electrode smearing width requires an additional convergence test with respect to the chosen plane-wave cutoff. Second, in order to accommodate the total potential drop across the cell, one would have to separate the left and right counterelectrodes within the vacuum. This requires, however, a larger vacuum region where the electron density must be neglible. To create the required voltage between the two counterelectrodes, one could use the (conventional) dipole correction with a cut in-between the two electrodes. Alternatively, one could adjust dynamically the values of the electrode charges such that an additional counterdipole is created. For the latter choice, the artificial interelectrode capacitor energy requires another correction. This complex procedure is circumvented in the generalized dipole correction that we propose here, by separating corrections to the potential and to the energy. First, we fix the potential without explicitly introducing the electrode charges. For the energy expression, we then add a simple extra term that accounts for electrode charges at an arbitrary position with respect to an *isolated* slab (see Sec. II C).

The correction potential for Eq. (4) can be compactly written as (inside the interval $0 \leqslant z < c$)

$$V^{\text{corr}}(z) = \begin{cases} z < z_0: & V_0 - \frac{2\pi Q}{cA} z^2 \quad -\mathcal{E}^{\text{corr}} z, \\ z \geqslant z_0: & V_0 - \frac{2\pi Q}{cA}(z-c)^2 \quad -\mathcal{E}^{\text{corr}}(z-c) \end{cases} \tag{5}$$

with

$$\mathcal{E}^{\text{corr}} = \mathcal{E}^{\text{L}} + \frac{2\pi Q}{A} - \frac{4\pi\mu}{cA}, \tag{6}$$

where the slab dipole with respect to z_0

$$\mu = \int dx\, dy \int_{z_0}^{z_0+c} (z - z_0)\rho^{\text{es}}(x, y, z) \tag{7}$$

is obtained from the electrostatic charge ρ^{es}, i.e., electrons + nuclear charges. Note that this dipole depends on both the slab position and z_0 whenever the slab is charged. V_0 is the alignment constant adjusted such that the laterally averaged potential $\overline{V}^{\text{tot}}$ right to the slab approaches $-z\mathcal{E}^{\text{R}}$ near the cut position for reasons discussed below.

The quadratic term in the correction potential compensates for the implicit homogeneous background. It is furthermore easy to verify that Eq. (5) correctly describes the desired jump in the electric field [cf. Eq. (3)]. The jump in the potential is equivalent to that of the standard dipole correction [21] for an augmented system comprised of the charged slab and its counterelectrodes placed at the cut. Thus, the generalized dipole correction can be thought of as simultaneously introducing a sheet monopole and sheet dipole at z_0.

The cut position z_0 can be anywhere in the vacuum region where the electron density is sufficiently low. If the density were exactly zero, the slab's dipole and the resulting generalized dipole correction to the potential energy would become independent of z_0. In practice, the electronic charge always decays into the vacuum region from both sides of the slab. Moving the cut position z_0 then gives rise to a small error, roughly proportional to the total charge contained between the original and final positions of the cut, and proportional to the voltage jump at the cut position. This uncertainty also exists in the standard dipole correction for charge-neutral slabs, but the voltage drop is typically much smaller (on the order of 1 eV) than in the generalized dipole correction (up to order of 100 eV). Therefore, in order to minimize the impact of density tail error, we set the cut position z_0 to the minimum of the laterally averaged electrostatic density at each step in the self-consistency iteration. Technically, this optimal choice is not required for applying the correction: the cut position could alternatively be set by an additional input parameter. In either case, the numerical noise from the unavoidable tail error can be further reduced by increasing the vacuum thickness.

Our approach would fail, however, when the desired target situation is inherently unstable against the escape of electrons into the vacuum region. This occurs regularly for positively charged counterelectrodes (e.g., for negatively charged slabs). If the counterelectrode is put at sufficient distance, the voltage drop across the vacuum exceeds the binding of electrons to the slab, and produces unwanted electron states near the positive counterelectrode that become occupied during the self-consistent iterations. Trying to adapt z_0 to the density minimum will in general not produce a stable solution. The experimental analog of this is field emission: a negatively charged surface will resolve the instability ultimately by constantly emitting the electrons toward the positive counterelectrode, hindered only by the tunneling barrier near the surface. In ground-state DFT, such a kinetic barrier does not exist and the probably best pragmatic solution is to place the counterelectrode so close that the critical voltage drop is not reached.

C. Energy

Until this point, we have corrected the *potential* such that it corresponds to that of the slab sandwiched between two electrodes. So far, the electrode positions $z^{\rm L}$ and $z^{\rm R}$ have not been needed. For arriving at a consistent energy expression, we have to make a choice where exactly the electrodes are located. An obvious choice would be to place them at the boundary of the supercell, i.e., at $z_0 - c$ for the left electrode and z_0 for the right one. Unfortunately, this introduces energy jumps when we want to dynamically adapt z_0 during the calculation. Figure 2 highlights the problem: if the position of the electrode depends on z_0, the potential and hence the electrode charge's potential energy varies if the position shifts from z_0 to a different position z_0'. It is therefore advantageous to choose fixed, but otherwise arbitrary, electrode positions $z^{\rm L}$ and $z^{\rm R}$, which corresponds to the situation sketched in Fig. 1(a). For similar reasons, the alignment convention for the potential is chosen such that it does not depend on z_0 directly (e.g., by setting the potential at z_0 to zero) or indirectly (e.g., by setting the average potential to zero). Instead, we aim at keeping the right electrode at constant potential. When we additionally leave the left electrode at constant charge, this choice makes it easy to interpret charge-transfer processes between the active right electrode and the slab (see below in Sec. II D).

The relevant electrode potentials can be extrapolated from the the potential at the cut position, or rather infinitesimally above and below it, and the known asymptotic fields $\mathcal{E}^{\rm L}$ and $\mathcal{E}^{\rm R}$. In Fig. 1(a) or Fig. 2, this corresponds to following the potential trend beyond the boundaries at $z_0 - c$ and z_0 until the counterelectrodes at $z^{\rm L}$ and $z^{\rm R}$, respectively. If we denote the laterally averaged potential near z_0 approached from below and above as $\overline{V}^{\rm tot}_<(z_0)$ and $\overline{V}^{\rm tot}_>(z_0)$, respectively, the electrode potentials are given by

$$\overline{V}^{\rm R} = \overline{V}^{\rm tot}_<(z_0) - \mathcal{E}^{\rm R}(z^{\rm R} - z_0), \tag{8}$$

$$\overline{V}^{\rm L} = \overline{V}^{\rm tot}_>(z_0) - \mathcal{E}^{\rm L}(z^{\rm L} - z_0 + c). \tag{9}$$

Using the alignment convention introduced above, and noting that the potential jump at z_0 is [cf. Eq. (5)]

$$\overline{V}^{\rm tot}_>(z_0) - \overline{V}^{\rm tot}_<(z_0) = \mathcal{E}^{\rm L}c - \frac{4\pi}{A}(\mu - Qz_0), \tag{10}$$

this simplifies to

$$\overline{V}^{\rm R} = -\mathcal{E}^{\rm R}z^{\rm R}, \tag{11}$$

$$\overline{V}^{\rm L} = -\frac{4\pi\mu}{A} - \mathcal{E}^{\rm L}z^{\rm L}. \tag{12}$$

The total electrostatic energy is then given by

$$\begin{aligned} E &= E^{\rm DFT} + \tfrac{1}{2}(Q^{\rm L}\overline{V}^{\rm L} + Q^{\rm R}\overline{V}^{\rm R}) \\ &= E^{\rm DFT} - \tfrac{1}{2}\mu\,\mathcal{E}^{\rm L} - \tfrac{1}{2}(Q^{\rm L}\mathcal{E}^{\rm L}z^{\rm L} + Q^{\rm R}\mathcal{E}^{\rm R}z^{\rm R}), \end{aligned} \tag{13}$$

where $E^{\rm DFT}$ is the DFT energy using the corrected electrostatic potential $V^{\rm tot}$. The last term in Eq. (13) highlights the trivial dependence on the chosen (virtual) electrode positions $z^{\rm L}$ and $z^{\rm R}$. In particular, for a given choice of $\mathcal{E}^{\rm L}$ and $\mathcal{E}^{\rm R}$ we can use the first two terms as an internally consistent total energy expression. This is the expression used in our implementation in SPHINX [22,23]. It corresponds to the (unphysical) situation of extrapolating the two virtual electrodes to $z^{\rm R} = z^{\rm L} = 0$. Satisfyingly, it also corresponds to the energy expression of the traditional dipole correction (for $Q = 0$) in the presence of an additional external field $\mathcal{E}^{\rm L}$.

D. Energy derivatives

The above energy expression corresponds to a well-defined charge-neutral DFT setup, namely, the charged slab sandwiched between the two virtual electrodes with an overall compensating charge. Therefore, all common derivatives can be immediately evaluated in the usual way. No explicit correction terms are needed for forces (derivative with respect to nuclear coordinates), eigenvalues (derivatives with respect to occupation numbers), or the Fermi energy (derivative with respect to electron number), provided that the corrected electrostatic potential $V^{\rm tot}$ [Eqs. (4) and (5)] is used instead of the periodic-boundary version.

However, some additional comments are appropriate. First, fixing the electrode positions in space breaks the usual translational invariance along z in periodic boundary conditions. The sum of the nuclear forces is therefore not zero in general, but amounts to

$$\begin{aligned} \sum_i F_z^i &= \frac{1}{2}Q\mathcal{E}^{\rm L} + \frac{1}{2}Q\mathcal{E}^{\rm R} \\ &= \frac{A}{8\pi}[(\mathcal{E}^{\rm R})^2 - (\mathcal{E}^{\rm L})^2]. \end{aligned} \tag{14}$$

This is not a computational artifact, but reflects the balance between the well-known Maxwell stresses on the left and right sides. Ignoring these forces would introduce an unphysical compensating force acting directly on each atom without a physical basis.

For the case of an external counterelectrode facing the surface, i.e., $\mathcal{E}^{\rm L} = 0$, the Maxwell stress in a real experimental setup is compensated by mounting the charged sample mechanically on the back side. In geometry optimization, we simulate this by fixing the positions of the first few layers on the back side (left). The Maxwell stress then puts the slab under tensile strain, which is physical. Indeed, we observe an expansion in the free part of the slab in this case. The stresses for fields relevant to field evaporation experiments (~10–100 V/nm) amount to 0.4–44 GPa. If the countercharge in experiment was located inside the material (e.g., band bending due to Fermi-level pinning at the surface), this would be simulated by setting $\mathcal{E}^{\rm R} = 0$ and $\mathcal{E}^{\rm L} = 4\pi Q/A$, and fixing the left side, which would put the slab under compressive strain.

A second comment concerns the Fermi energy, i.e., the derivative with respect to electron number and hence with respect to the charge Q. If we keep the left side field fixed (assuming that the surface of interest is the right one), any change in the slab charge must be compensated by the right electrode charge. The Fermi energy therefore corresponds to the voltage between the right electrode and the electronic subsystem in the slab. Any changes to the nominal charge would be realized by a current flowing between the electronic subsystem and the right electrode. The voltage then changes like in a standard plate capacitor. Of course, if the left electrode's charge is nonzero, the change in the total voltage drop between the left and right electrode $\Delta U = \Delta(\overline{V}^{\mathrm{L}} - \overline{V}^{\mathrm{R}})$ upon charging contributes $\frac{1}{2}\Delta U Q^{\mathrm{L}}$ to the overall energy.

Of course, we verified the agreement between the analytical derivatives and their numerical counterparts from finite differences for our implementation in SPHINX [22]. We found excellent agreement within the numerical accuracy of the calculation (10^{-4} eV for Fermi energy, 10^{-5} hartree/bohrs for forces). Moreover, we used the convergence behavior of our direct energy minimization algorithm [26] down to 10^{-10} hartree as a sensitive check for the consistency between the total energy and the Hamiltonian (derivative with respect to the density matrix) and also the Kohn-Sham eigenvalues. Finally, the energy is independent of z_0 to within the expected accuracy, i.e., it has the desired property of being invariant to shifts in the position of the compensating charge and dipole layer.

III. IMPROVEMENTS ON SELF-CONSISTENT FIELD ITERATIONS

One of the challenges for actual calculations turned out to be the high sensitivity of the potential to incorrect distributions of the charge. For instance, shifting 0.01 e from the left to the right side of a 20 Å thick slab with a surface area of 10 Å^2 incurs a voltage drop of 3.6 V between the surfaces. This is sufficient to dramatically modify the occupation of surface-related states. The occupation changes imply even larger redistribution of charge across the surfaces than the original fluctuation. In consequence, they render the self-consistent iterative procedure unstable.

The misplacement of charge originates from the charge initialization, as discussed below in Sec. III A, or from fluctuations arising from the self-consistent field iterations. The latter are commonly and efficiently suppressed by density preconditioners such as the Kerker preconditioner $\hat{K}$, written in reciprocal space as [27,28]

$$\hat{K}\Delta\rho(\mathbf{G}) = \frac{|\mathbf{G}|^2}{q_{\mathrm{TF}}^2 + |\mathbf{G}|^2}\Delta\rho(\mathbf{G}). \tag{15}$$

The physical model underlying the Kerker preconditioner is a homogeneous Thomas-Fermi metal with screening length q_{TF}. At the slab surface, this model is very inaccurate and produces density fluctuations that leak into the vacuum region.

However, similar to the standard dipole correction for neutral slabs, a temporary escape of electron density into the vacuum region during the self-consistent iterations must be strictly avoided. The reason lies in the calculation of the dipole moment and the associated electric correction field: in the absence of density in the vacuum, the total dipole moment depends on the cutting position only trivially via the shift of the reference position [cf. Eq. (7)]. Otherwise, if the cutting position is shifted by Δz, the dipole additionally changes by $\Delta z \overline{\rho}(z_0) c A$. If the cutting position is determined dynamically in each step from the minimum of the density profile $\overline{\rho}$, fluctuations of the charge density may shift the cutting position from iteration to iteration. The associated quasirandom fluctuations in the correction field prevent progress in the self-consistency unless severe damping is employed. But even when the cutting position is fixed, charge fluctuations may cross the cutting boundary and lead to severe convergence problems.

The solution to this dilemma is a preconditioner that accounts for the local material's behavior. A conceptually straightforward, very powerful model for the related problem of potential mixing has recently been proposed by Lin and Yang [29]. They termed it "elliptic preconditioner" due to the appearance of an elliptic differential equation that must be solved numerically. We adapted the idea to density mixing, and describe it in Sec. III B.

A. Charge initialization

In many DFT codes the initial charge density is set up from an overlap of atomic charge densities. In order to arrive at a charged slab, it is tempting to rescale the electron density

$$\tilde{\rho}(\mathbf{r}) = \lambda\rho(\mathbf{r}) \qquad \text{with } \lambda = 1 - Q/N^{\mathrm{el}}, \tag{16}$$

where N^{el} denotes the number of electrons in the charge-neutral case. This corresponds to distributing the excess charge over the entire slab. In the self-consistent solution, however, the excess charge will sit in the near-surface region. For conditions relevant to field evaporation ($\mathcal{E}^{\mathrm{L}} = 0$, $\mathcal{E}^{\mathrm{R}} = 10^{10}$–$10^{11}$ V/m corresponding to surface charge densities of 0.006–0.06 e/Å^2), this discrepancy may lead to dramatic distortions of the effective potential, raising the surface region far above the bulklike part [see Fig. 3(a)] and spilling the associated electron density all over the supercell.

To prevent this, we introduce the charge not by rescaling, but by adding an additional Gaussian layer density

$$\overline{\rho}(z) = Q\frac{\sqrt{\pi}}{\beta}e^{-(z-z_c)^2/2\beta^2} \tag{17}$$

with a default broadening of $\beta = 1$ bohrs. The center position z_c is placed at the rightmost atomic layer. For positive charges (removing electrons), this reduces the electron density. Depending on the surface roughness, this might lead to artificial negative electron densities in the setup. These may be avoided by increasing the broadening and/or shifting the center of charge z_c deeper into the slab. Alternatively, one can also place spherical Gaussian charges on each surface atom, which yields similar results. SPHINX [22,23], our plane-wave DFT code that we used to implement the approach, supports both Gaussian layers and localized Gaussian charges at arbitrary positions as well as combinations thereof.

The key advantage is that the initial potential is already much closer to the final potential [see Fig. 3(b)], producing a realistic electronic structure right from the outset, even if the surface dipoles are vastly off.

B. Elliptic preconditioner

As elegantly outlined by Lin and Yang [29], the self-consistent procedure can be viewed as fixed-point iterations

$$V^{(n)}(\mathbf{r}) \to V^{(n+1)}(\mathbf{r}). \tag{18}$$

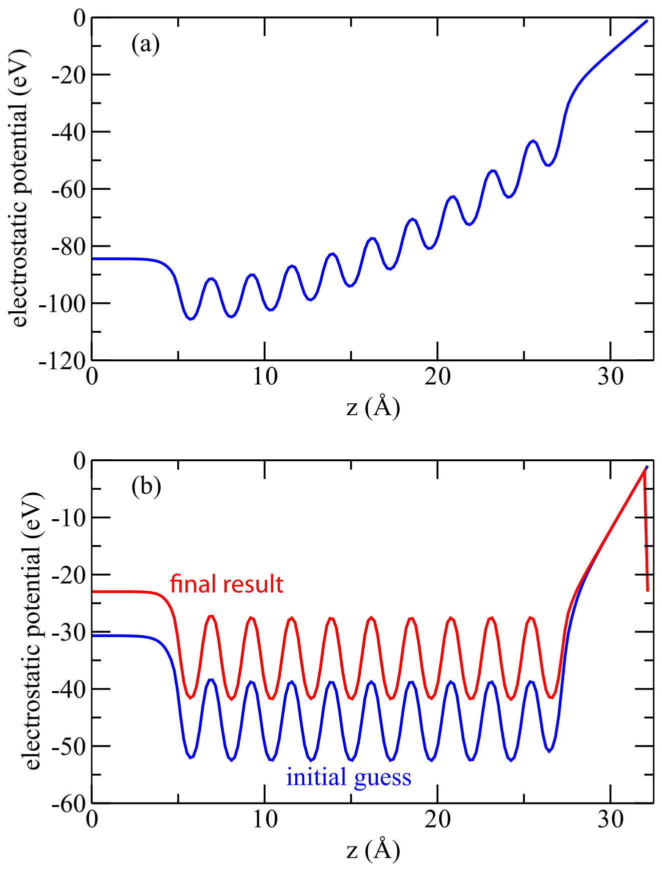

FIG. 3. Potential for electrons, averaged over the x-y plane, of a charged slab from (a) a rescaled electron density and from (b) an initial density with a additional Gaussian layer charge (blue curve) with comparison to the self-consistent potential (red curve). Note the different scale.

The mapping is performed by taking a given DFT potential, solving the Kohn-Sham equation, constructing a new density, and computing a new potential from it. As is well known, this can be rewritten as iterations of the error

$$\delta V^{(n)} = V^{(n)} - V^{\mathrm{scf}} \tag{19}$$

from the self-consistent potential V^{scf}. From this viewpoint, each step corresponds to calculating the electronic response to a slightly perturbed ground-state system. Near the solution, this response can be linearized. The mapping from one iteration to the next is then dictated by the response properties of the self-consistent solution: The density $\delta\rho_{\mathrm{out}}^{(n)}$ induced by $\delta V^{(n)}$ reflects the polarizability (written here as a linear operator $\hat{\chi}$)

$$\delta\rho_{\mathrm{out}}^{(n)} = \hat{\chi}\,\delta V^{(n)}. \tag{20}$$

The resulting potential is obtained by computing the change in the Coulomb (also known as Hartree) and exchange-correlation potential. This can be written in linearized form as the Hartree-exchange-correlation kernel $\hat{v}_{\mathrm{Hxc}}$:

$$\delta V_{\mathrm{out}}^{(n)} = \hat{v}_{\mathrm{Hxc}}\delta\rho_{\mathrm{out}}^{(n)}. \tag{21}$$

Combining Eqs. (20) and (21) leads to

$$\delta V_{\mathrm{out}}^{(n)} = \hat{v}_{\mathrm{Hxc}}\,\hat{\chi}\,\delta V^{(n)}, \tag{22}$$

which can be inverted to give the ideal step toward the self-consistent solution

$$\Delta V = -\delta V^{(n)} = (1 - v_{\mathrm{Hxc}}\,\hat{\chi})^{-1}\big(V_{\mathrm{out}}^{(n)} - V^{(n)}\big). \tag{23}$$

The main idea of the elliptic preconditioner introduced in Ref. [29] is to approximate Eq. (23) by a differential equation

$$\{-\nabla(a(\mathbf{r})\nabla) + 4\pi b(\mathbf{r})\}\Delta V = -\nabla^2\big(V_{\mathrm{out}}^{(n)} - V^{(n)}\big). \tag{24}$$

The motivation is that setting $a(\mathbf{r}) = 1$ and $b(\mathbf{r}) = \frac{1}{4\pi}q_{\mathrm{TF}}^2$ yields the Kerker preconditioner, while $b(\mathbf{r}) = 0$ and $a(\mathbf{r}) = \varepsilon$ yields a simple scaling $1/\varepsilon$ appropriate for a homogeneous dielectric. In other words, $a(\mathbf{r})$ accounts for the local dielectric screening and $b(\mathbf{r})$ for the local metallic screening in the Thomas-Fermi approximation.

For density mixing, we seek to obtain the optimal density step $\Delta\rho$ from the residue $R^{(n)} = \rho_{\mathrm{out}}^{(n)} - \rho^{(n)}$. For this, we propose an almost identical differential equation to Eq. (24), namely,

$$\left\{1 + \underbrace{\left(-\frac{1}{4\pi}\nabla\big([a(\mathbf{r}) - 1]\nabla\big) + b(\mathbf{r})\right)\hat{v}_C}_{\hat{P}}\right\}\Delta\rho = R^{(n)}. \tag{25}$$

The subtle difference is that the Hartree-exchange-correlation operator $\hat{v}_{\mathrm{Hxc}}$ that connects the error in potential to the one in density [cf. Eq. (22)] has been replaced by a pure Coulomb potential $\hat{v}_c = -4\pi(\nabla^2)^{-1}$ in periodic-boundary conditions with a special alignment convention. Namely, the operator $\hat{P}$ on the left-hand side of Eq. (25) should be interpreted as adding the Hartree-potential-induced density. We force this induced density to carry no net charge by adjusting the alignment of the potential $\hat{v}_C\Delta\rho$: only the metallic contribution $b(\mathbf{r})$ is able to change the net charge. A constant shift in the potential will therefore add charge proportional to $b(\mathbf{r})$. Enforcing charge neutrality is straightforwardly achieved by first evaluating a preliminary induced density ρ_{ind}' at an arbitrary alignment, and then setting

$$\rho_{\mathrm{ind}}(\mathbf{r}) = \rho_{\mathrm{ind}}'(\mathbf{r}) - b(\mathbf{r})\frac{\int d^3\mathbf{r}\,\rho_{\mathrm{ind}}'(\mathbf{r})}{\int d^3\mathbf{r}\,b(\mathbf{r})}. \tag{26}$$

An approximate solution S to the differential equation (25) is obtained iteratively by a conjugate-gradient scheme until

$$|(1 + \hat{P})S - R| < \epsilon|R| \tag{27}$$

for a relative accuracy $\epsilon = 10^{-5}$. This approximate solution then acts as the preconditioned residue in a standard DIIS density-mixing scheme [28].

The final task is to find a suitable recipe for dielectric and metallic polarizability functions $a(\mathbf{r}) - 1$ and $b(\mathbf{r})$. Since we are mostly interested in metallic slabs, we use

$$a(\mathbf{r}) = 1, \tag{28}$$

$$b(\mathbf{r}) = \frac{1}{4\pi}q_{\mathrm{TF}}^2(\tilde{\rho}_e(\mathbf{r})) \tag{29}$$

as a pragmatic approach. Here, $\tilde{\rho}_e$ is a Gauss-broadened version ($\beta = 1$ bohrs) of the initial density, and

$$q_{\mathrm{TF}}^2(\tilde{\rho}) = 4\left(\frac{3\tilde{\rho}}{\pi}\right)^{1/3} \tag{30}$$

is the Thomas-Fermi screening length. For small electron densities $\tilde{\rho} < 10^{-3}$ e/bohrs^3, we set b to exactly zero. The resulting function is again Gauss broadened ($\beta = 1$ bohrs). The motivation for the broadening is to smoothen the rapid variation of the screening behavior near the atomic cores, which would lead to an artificial localization of screening charge due to the most problematic long-wavelength fluctuations near the cores within the model. The chosen broadening is a compromise: to dampen fluctuations near the atomic cores, large values are desirable, but too large values spill the screening response at the surfaces into the vacuum region.

One advantage of this setup is that by construction, the preconditioner cannot produce artificial densities in the vacuum region. Of course, for complex systems containing metallic and nonmetallic regions, more refined recipes must be developed to avoid metallic behavior (which we associate here with the presence of electron density) in insulating material. Nevertheless, the present procedure even works reliably for insulator slabs such as the ones discussed in Sec. IV B. For semiconducting slabs, however, direct minimization, [26] turned out to be the fastest algorithm in our calculations, even if the slabs are charged.

In our calculations, the computational effort to solve the preconditioner equation (25) amounted to 1%–2% of the total time. We therefore did not explore the various opportunities to improve the algorithm. For instance, the broadening may allow for a reduced mesh when solving the elliptic preconditioner equation. However, care must be taken to make sure that the preconditioning does not remove components that cannot be represented in the reduced mesh. This would introduce an implicit filtering on the preconditioned residue and prevent any improvement on the filtered-out components. Therefore, we always replace the approximate, potentially filtered solution S' from the employed solver by

$$S = R + \mathcal{F}^{-1}[S' - (1 + \hat{P})S']. \tag{31}$$

The $\mathcal{F}^{-1}$ operation interpolates from the reduced mesh to the fine mesh, e.g., by a Fourier interpolation. The third term corresponds to the filtered representation of R contained in S' (to within the accuracy of the solution). In other words, we replace the filtered component of R contained in S' by its unfiltered one (first term).

The advantage of the elliptic preconditioner is visible already for neutral slabs. In Fig. 4, we show the convergence of the self-consistent field iterations for a 10-layer Al(111) slab with various thicknesses of the vacuum separation. No dipole correction was used. Using a dipole correction even worsens the convergence for the Kerker preconditioner. It is apparent that the Kerker-preconditioned cases sensitively depend on the vacuum separation, probably due to temporary charge fluctuations that leak into the vacuum region. In contrast, the elliptic preconditioner not only significantly speeds up convergence (similar to the observations made in Ref. [29]), but also entirely removes the vacuum dependence. The impact for thicker slabs and charged slabs is even larger, as such calculations often failed to converge with the Kerker preconditioner. In contrast, the elliptic preconditioner guaranteed rapid convergence in all test cases.

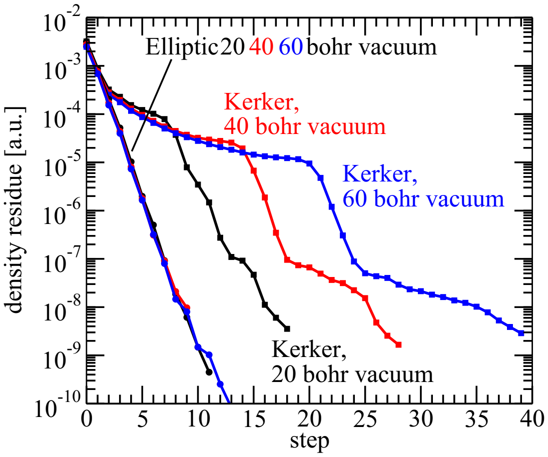

FIG. 4. Convergence behavior of neutral 10-layer Al(111) slabs with three differerent vacuum thicknesses for the Kerker and elliptic preconditioners (see text).

IV. APPLICATIONS

A. Field evaporation

In order to investigate thermally driven field evaporation, it is necessary to explore the potential energy surface in the presence of a surface electric field. In particular, we are interested in positive biases, i.e., positively charged surfaces. The left side of the slab is kept field free ($\mathcal{E}^{\mathrm{L}} = 0$). The field driving the evaporation is thus solely described by the field on the right side $\mathcal{E} = \mathcal{E}^{\mathrm{R}}$, and we omit the R superscript from the right field from now on. The generalized dipole correction is a valuable framework for investigating such potential energy surfaces. As demonstration, we consider the case of an adsorbed Al atom on Al(111). Sanchez *et al.* were the first to study this prototypical case by means of DFT [19]. Lacking the generalized dipole correction, they employed an explicit Gaussian countercharge and a necessarily symmetric setup (with adatoms on both sides of the slab), which restricted the affordable complexity to such a relatively simple case. Moreover, they neglected relaxation within the surface in response to adatom desorption, probably to avoid the additional computational cost. Our approach, on the other hand, is capable of handling complex situations such as desorption from a step or from a kink site and we have recently shown how such investigations provide new insights into the previously poorly understood desorption mechanism [20].

Sufficiently high above the surface, any atom will be able to transfer some electrons to the surface and acquire a positive charge [19,20]. This ion will then move in the electric field toward the counterelectrode. In the potential energy surface, this yields a steady lowering of the energy as the atom moves away from the surface. For the desorption of a single ion, its z coordinate is therefore a suitable reaction coordinate to describe its evaporation [19]. In order to explore the energy landscape orthogonal to the reaction coordinate, the z coordinate of the atom of interest is kept at a fixed value, while all other atoms are relaxed except for the left 2–3 layers to prevent the slab from drifting in the field as described in Sec. II D. This is in contrast to the approach taken by Sanchez

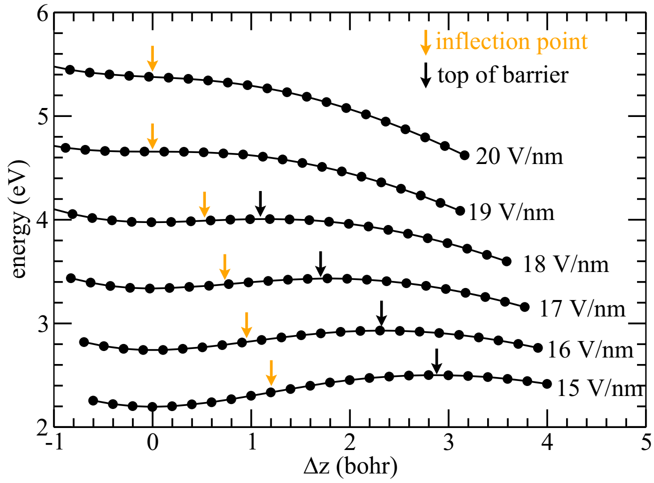

FIG. 5. Binding curves for an adatom on Al(111) for various electric fields. Δz is displacement from the equilibrium position for fields below 19 V/nm. Black arrows mark the top of the desorption barrier. For higher fields, $\Delta z = 0$ corresponds to the inflection point (see text). Orange arrows mark the inflection point.

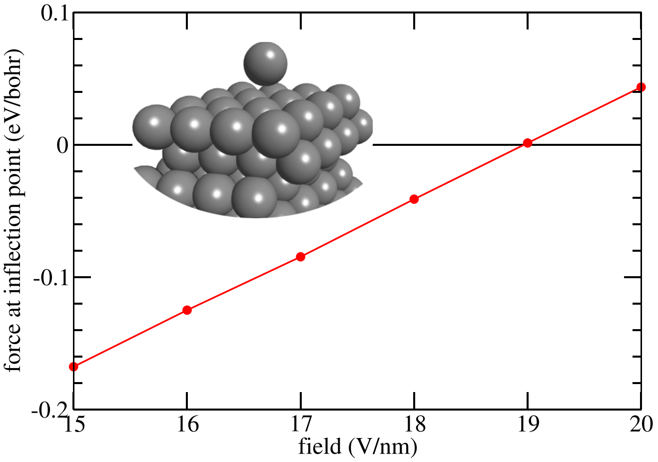

FIG. 6. Force $-dE/d\Delta z$ at the inflection point (cf. Fig. 5) vs applied electric field for an adatom on Al(111). A negative sign indicates a force pointing toward the surface.

et al. in their study on adatoms, where they kept the entire substrate fixed [19].

All DFT calculations for this work were performed in the local-density approximation (LDA) with the SPHINX code [22]. For field evaporation calculations, a plane-wave basis set with an energy cutoff of 20 Ry and norm-conserving pseudopotentials were used. The **k**-point sampling for the 4×4 Al(111) slab with 16 surface atoms and 128 atoms in total (8 atomic layers) was 4×1. Calculations for a 5×5 cell yield very similar results. Atoms were relaxed until the forces were below 10^{-4} hartree/bohrs using our recent geometry optimizer [30], which is ideally suited to capture global strain from the Maxwell stresses.

From a series of calculations for various values of z for a certain field $\mathcal{E}$, we obtain the binding-energy curve $E(z, \mathcal{E})$. Figure 5 shows this for the Al adatom. The binding curve shows a minimum when the atom is in its equilibrium position, a maximum at the barrier for desorption and the aforementioned steady slope far above the surface. The energy difference between the minimum and maximum corresponds to the thermal barrier for desorption $\Delta E^{\rm des}(\mathcal{E})$. Following Sanchez *et al.* [19], one can now straightforwardly obtain barriers as a function of applied field and determine thus the critical "zero-barrier" field $\mathcal{E}^{\rm zbf}$ where the barrier vanishes.

In practice, geometry optimization at fields near the critical one on a more complex surface can be challenging because surface atoms in other sites than the investigated one may be close to evaporation, too. The finite displacements during optimization may then move such an atom beyond its desorption barrier, and remove it from the surface in subsequent steps. It is therefore desirable to extrapolate the zero-barrier field from subcritical fields. As we will see in a moment, the generalized dipole correction provides the perfect framework to achieve this.

For this, we focus on the inflection point of the desorption curve marked with orange arrows in Fig. 5, which is easily determined as the maximum of (retracting) force on the constrained atom along the reaction coordinate. As long as there is a finite barrier, the inflection point falls in-between the local energy minimum (equilibrium position) and the local energy maximum (top of the barrier). At the zero-barrier field, the local minimum and maximum merge into the inflection point, and the retracting force becomes zero. However, in contrast to minimum and maximum, the inflection point is still well defined beyond this critical field. As long as the geometry optimization runs stably, it is easy to determine the force at this point.

Figure 6 shows the force on the adatom at the inflection point for different fields. It is apparent that crossing the zero-barrier field does not induce any special features and the force continues to increase linearly with the field. This implies that the effective charge on the reaction coordinate only weakly depends on the applied field near the critical field strength. Figure 6 also shows the zero-barrier field strength as 19 V/nm (18 V/nm for the 5×5 cell, not shown), in good agreement with Sanchez *et al.* [19]. Apparently, relaxation of the atoms within the close-packed surface plays a minor role for the desorption energetics of a single adatom. For more complex surface structures, however, significant rearrangements occur [20].

Yet, we can go one step further than just observing the linearity of the maximum constraining force with applied field. For this, we recall that varying the field corresponds to varying the surface charge, and that the Fermi energy reflects the corresponding derivative of the DFT energy (13). On the other hand, the constraining force F_ξ (with ξ being the reaction coordinate, in the case of the adatom its z coordinate) is equivalent to the change of the DFT energy with respect to the reaction coordinate. Thus, we can express the total energy's double derivative with respect to field and reaction coordinate (which has the physical dimension of a charge) in two ways as

$$\frac{\partial F_\xi}{\partial \mathcal{E}} = -\frac{A}{4\pi}\frac{\partial^2 E^{\rm tot}}{\partial Q\,\partial \xi} = \frac{A}{4\pi}\frac{\partial E^{\rm Fermi}}{\partial \xi}. \tag{32}$$

The derivative of the Fermi energy with respect to the reaction coordinate can be determined numerically from the calculations used for the binding curve at a given field at negligible cost. Indeed, we find very good agreement for the effective "desorption charge" of 0.81 ($\pm \sim 0.02$ depending on the interpolation scheme for forces and Fermi energies) obtained in this way with the average slope of the force curve in Fig. 6 (0.79, from linear regression). This equivalence offers a way to determine critical fields from a rather coarse mesh of applied fields, thereby reducing the computational effort.

We note in passing that the change in the Fermi energy would be directly related to the change in the slab dipole μ from Eq. (7) if the surface dipole on the left side was independent of the slab charge and the changes on the right surface. However, at finite slab thickness we do observe small, rather unsystematic variations of the left-side work function due to the incomplete decay of Friedel oscillation from the right surface. Therefore, extracting the field dependence of the constraining force from the Fermi energy is more accurate.

B. Surface defect calculations

As a second application, we used the counterelectrode approach realized by the generalized dipole correction for the calculation of the formation energy of a charged defect at a surface. The problem with almost all charged defect calculation is that the charge compensation in the simulation differs from the experimental situation, and in most cases is even "unphysical" in the sense that such smeared-out charge distributions overlapping with the actual atoms, as they are used in simulations, do not exist in reality. In particular, it was criticized that the interaction between the homogeneous background [14,15] or a rescaled nuclear charge [6] at the (sub)atomic level introduces errors that are not well understood and therefore difficult to quantify. The generalized dipole correction gives the chance to move the compensating charge away from the electronic system into the vacuum, where the electrostatics is well understood. The purpose of this part is therefore twofold: demonstrate that (1) charged defects at surfaces can be calculated without artificial charge distributions that overlap with the quantum system, and that (2) the results agree well with the traditional homogeneous background approach.

Specifically, we illustrate the approach for the oxygen vacancy at the MgO(100) surface in the +2 charge state as an example. To model the defect, we use an eight-layer MgO(100) slab with an oxygen vacancy at the right surface. The vacuum region is 15 Å wide. The surface unit cells are multiples of the conventional four-atom square unit cell. For these calculations, the projector-augmented wave (PAW) formalism was used [31]. The PAW setups for Mg and O were taken from the VASP database [32]. The plane-wave cutoff was 28 Ry, and $2 \times 2 \times 1$ **k**-point sampling with offset $[\frac{1}{2}, \frac{1}{2}, \frac{1}{4}]$ was used for all supercells. We tested both the self-consistent field algorithm with the elliptic preconditioner of Sec. III B as well as the direct minimization algorithm [26]. Both work, but in contrast to the metallic Al slab, we found the direct minimization to be more efficient for MgO, a common observation for systems with a gap or a low density of states at the Fermi level within DFT.

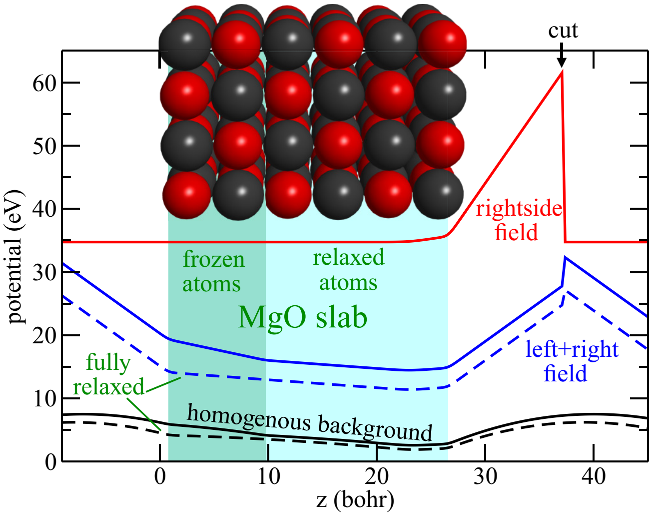

FIG. 7. Sketch of electron potential shape along z direction for three different boundary conditions (see text): right-side field, left + right field, and standard periodic boundary conditions with implicit homogeneous background. The potentials are shifted from their usual alignment for visibility. The shown z range encompasses $\sim$1.15c for clarity. Dashed lines correspond to fully relaxed cases (ionic screening within the entire slab, i.e., without a frozen region), where this is possible (see text). The positive defect is located near the right surface.

The formation energy of a defect depends on the chemical potentials in the target scenario [13]. For oxygen, we choose the oxygen molecule in the triplet state as reference. The electron chemical potential, i.e., the Fermi level in the target scenario, was set to the bulk valence band maximum (within DFT-LDA) in order to allow for a comparison with the formation energy in the bulk. The value, 5.5 eV below the vacuum level, was obtained from a separate bulk calculation, which was then aligned to the defect-free slab's energy scale via the electrostatic potential in the center of the slab. We did not consider band bending, i.e., the value corresponds to the dilute limit at negligible defect concentrations.

We compare three different boundary conditions. Their potentials along the z direction (averaged over the x-y plane) are illustrated in Fig. 7. The three cases are the conventional homogeneous background ("hom. background," black line in Fig. 7) [14], the generalized dipole correction with zero field at the back side ("right-side field," red line), and the symmetric-field case with fields of equal magnitude, but opposite direction on the right and left sides ("left+right field," blue line). The left+right field case is realized by setting the left field $\mathcal{E}^{\mathrm{L}}$ to $+2\pi Q/A$. We also attempted calculations where the entire field is on the left side, but observed dielectric breakthrough in all cases. As explained above, the net force on a charged slab does not vanish in general, and we therefore fixed the positions of the two back-side layers on the left. As shown in Fig. 7, the suppression of ionic relaxation modifies the screening in the lower part of the slab. For the homogeneous background and the left+right field cases, this restriction can be lifted as the net force vanishes in

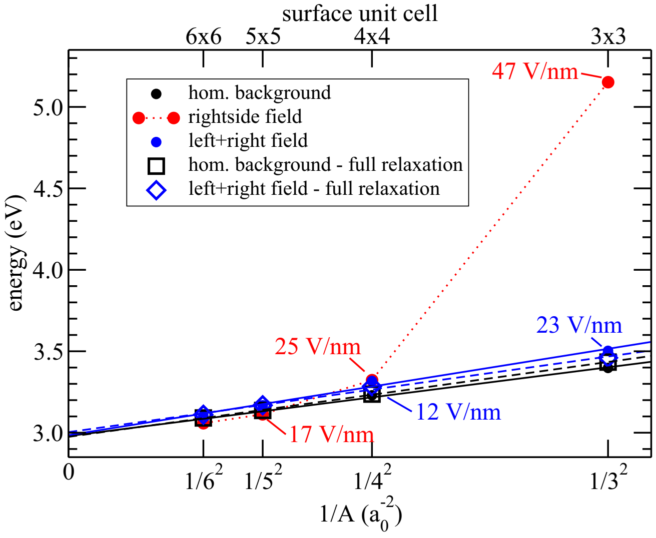

FIG. 8. Comparison of calculated formation energies of the oxygen 2+ vacancy at MgO(001) in the dilute limit obtained from a homogeneous background approach and the present counterelectrode approach for various supercell sizes. All cases have been corrected for artificial interactions in the repeated-slab approach [14]. The chemical potential for the electrons (Fermi energy) is set to the bulk valence band maximum. The solid and dashed lines are linear regressions to the partially relaxed and fully relaxed cases, respectively. A $1/A$ convergence trend is observed. The red dotted line serves as a guide to the eyes, connecting the calculated values for the right-side-field case.

these setups. We therefore included these cases with complete ionic relaxation, too (dashed lines). Outside the slab, in the left-side vacuum region, this leads to a constant potential offset compared to the partially frozen cases, but no qualitative change in the overall profile.

Charged defect calculations need finite-size corrections [6,10–15]. The counterelectrode approach is no exception. We therefore implemented the generalized dipole correction also in the `sxdefectalign2d` code [14]. Indeed, since the total energy strongly depends on the placement of the counterelectrode(s), formation energies without corrections are entirely meaningless. In our case, finite-size corrections reached several 10 eV in some cases, mostly accounting for the effective capacitor energy. For the relaxed part of the MgO slab, we determine a dielectric constant of $\varepsilon_r = 9.8$ from a line fit of the effective potential [14]. For the unrelaxed part (electronic screening only) we find $\varepsilon_r = 3.36$.

In Fig. 8, we compare the resulting formation energies extrapolated to the dilute limit obtained from the various approaches for different supercell sizes. We observe that, for a given, sufficiently large supercell size, the different approaches agree to within ~0.1 eV. The most dramatic exception is the right-field case for the smallest surface cell (3 × 3). However, the applied field in this case is very large (47 V/nm). While we do not observe any indications of bond breaking (note that the Al adatom evaporates spontaneously at half that field!), the field noticeably distorts the surface structure of MgO. In consequence, linearizing the screening response, as assumed in the correction, does not work well. As a result, the effective position of the charge in the correction procedure is too high (≈1 Å above the other cases). Yet, already at 27 V/nm, the corrections work very well.

It is also noticeable that the electrostatic corrections do not fully remove the size dependence with respect to the surface unit cell. This trend is essentially independent on the electrostatic boundary conditions used. The error seems to scale linearly with the surface concentration of the defect, $1/A$. We were not able to identify the underlying interaction mechanism; it might be an insufficient modeling of the lateral electrostatics, residual strains, or a combination thereof. The crucial point here is that the counterelectrode approach and the conventional homogeneous background approach give very similar effects. One motivation for including the fully relaxed calculations was to check if the modified screening at the relaxed/frozen boundary is responsible for the observed effects, but from Fig. 8 we conclude that it is not an important factor.

Assuming that the apparent $1/A$ behavior persists to the limit of infinite supercell sizes, we can extrapolate the formation energy of the vacancy at the surface to 3.0 eV. This can be compared to the bulk value of 3.2 eV obtained with an equivalent computational setup (DFT-LDA, same PAW potentials, same plane-wave cutoff, 2 × 2 × 2 **k**-point sampling in a 4 × 4 × 4 supercell of the conventional eight-atom unit cell, full relaxation, bulk defect corrections [12]). In other words, the +2 vacancy at the surface is slightly (by 0.2 eV) more stable than in the bulk.

V. CONCLUSIONS

In this work, we introduce a "generalized dipole correction" as an efficient means to realize a computational counterelectrode of charged slab calculations in periodic-boundary conditions. The correction can be applied to any plane-wave DFT code at a negligible computational cost. The correction can be thought of as introducing a compensating sheet monopole and sheet dipole layer in the vacuum between the slabs. It thereby efficiently decouples the periodic images and produces boundary conditions equivalent to an isolated slab sandwiched between the electrodes (cf. Fig. 1). At the position of these electrode layers, the electric potential and the electric field are discontinuous. The energy expression is chosen to mimic the isolated case, thereby removing any dependence on the artificial vacuum size or placement of the cut position.

We then introduced two key improvements to the standard DFT formalism. These are critical to overcome the slow electronic convergence of traditional DFT approaches when considering slabs in large electric fields. The first regards the initialization of the charge density. It is required to achieve a realistic field distribution right at the beginning. The second is an implementation of the elliptic preconditioner for self-consistent field iterations, adapted here to the case of density mixing. We provide an explicit, pragmatic approach to determining the screening model for repeated slabs. Despite the screening model's conceptual simplicity, it works extremely well for metallic slabs and reasonably for semiconducting ones.

Finally, we demonstrated that the generalized dipole correction can be used to investigate field evaporation at huge electric fields. It is thus ideally suited to understand the fundamental processes in atom probe tomography. Likewise, the counterelectrode approach realized by the generalized dipole correction provides an efficient way to perform charged-defect calculations at surfaces. We hope (and expect) that the generalized dipole correction is going to be implemented in major plane-wave DFT codes. Thereby, it will open the way for simulating electrified surfaces in a more direct and computationally more efficient way than was possible before.

ACKNOWLEDGMENT

This work was financially supported by the BigMax network of the Max-Planck society.

APPENDIX: INPUT FILE SETTINGS

The methodology presented here requires rather few additions to a typical input file for any DFT code (and hence can be implemented in a user-friendly manner). In order to illustrate this point, we discuss in the following the characteristic items in the input file for the SPHINX code [22]. In addition to the source code [23] for installations on high-performance computers, we provide a compiled binary version of the SPHINXcode as a conda package [33] for testing. We believe that other codes can come up with a similarly lean user interface to simplify application of the methodology. Figure 9 presents an excerpt of the SPHINX input file for a charged-defect calculation (3 × 3 cell). Key lines relevant for the methodology discussed in this work are highlighted and will be discussed in the following. For clarity, we omitted all parts of the input file relating to defining the atomic positions, the PAW potentials, and the basis set.

Lines 20–22, highlighted in light green, reflect the key settings for computing charged slabs with the generalized dipole correction. These lines define, in the given order, the nominal charge (+2 = two electrons less), the use of the (generalized) dipole correction, and $\mathcal{E}^{\mathrm{L}}$ (parameter `zField`, in units of V/bohrs). In this particular case, $\mathcal{E}^{\mathrm{L}}$ was chosen such that the fields are equal and opposite on both sides of the slab. `zField` is an optional parameter, the default is zero field on the left side.

The charge initialization (see Sec. III A) is set up in lines 27–34, highlighted in blue. In this particular case, the total charge of +2 is distributed over the four Mg neighbors next to the vacancy located in the right layer at (0,0,23.5) (in units of bohrs), using a Gaussian charge on each of these sites. This choice was made because it places the initial charge in the rightmost layer (where we expect the dominant localization) in the lateral vicinity of the vacancy.

The elliptic preconditioner for the self-consistent iterations as introduced in Sec. III B is requested in line 47 (highlighted in red). In the present implementation, the preconditioner has no input-controlled parameters. Finally, since the calculation type is geometry optimization (lines 39–41 request our on-the-fly parametrized redundant internal-coordinate optimizer [30]), line 42 (highlighted in purple) switches off the force drift filter which normally removes the net force by adding a small, constant compensating counterforce on each atom. The need for keeping the net force was discussed in Sec. II D.

`input.sx`

```
format paw;

include <parameters.sx>;

a0 = 5.549;
c0 = 47.082822;

structure {
   cell = [[ 3*a0,     0,    0],
           [ 0,     3*a0,    0],
           [ 0,        0,   c0]];
   ...
}

...

PAWHamiltonian {
   ekt = 0.1;    // eV
   xc  = LDA_PW;
   nEmptyStates = 10;
   nExcessElectrons = -2;
   dipoleCorrection;
   zField = -2.4680591/2; // V/bohr
}

initialGuess {
   rho { atomicOrbitals;
      charged {charge=0.5;
               coords=[ a0/2, a0/2, 23.5];}
      charged {charge=0.5;
               coords=[-a0/2, a0/2, 23.5];}
      charged {charge=0.5;
               coords=[ a0/2,-a0/2, 23.5];}
      charged {charge=0.5;
               coords=[-a0/2,-a0/2, 23.5];}
   }
}

main {
   ricQN {
      dF=1e-4;
      ric { withAngles; maxDist=7; }
      driftFilter=false; // keep net force
      bornOppenheimer {
         scfDiag {
            blockCCG {}
            dEnergy=1e-7;
            preconditioner {type=ELLIPTIC;}
            rhoMixing=0.8;
         }
      }
   }
}
```

FIG. 9. Key items of the input file for a charged-defect calculation with SPHINX (see text). For clarity, some parts of the input file are omitted. The full input file is available as Supplemental Material [25].

[1] R. J. Maurer, C. Freysoldt, A. M. Reilly, J. G. Brandenburg, O. T. Hofmann, T. Björkman, S. Lebègue, and A. Tkatchenko, Annual Rev. Mater. Res. **49**, 1 (2019).
[2] G. Schwarz, Untersuchungen zu Defekten auf und nahe der (110)-Oberfläche von GaAs und weiteren III-V-Halbleitern, Ph.D. thesis, Technical University Berlin, 2001.
[3] M. Otani and O. Sugino, Phys. Rev. B **73**, 115407 (2006).
[4] A. Y. Lozovoi and A. Alavi, Phys. Rev. B **68**, 245416 (2003).
[5] Y. Xu, O. T. Hofmann, R. Schlesinger, S. Winkler, J. Frisch, J. Niederhausen, A. Vollmer, S. Blumstengel, F. Henneberger, N. Koch, P. Rinke, and M. Scheffler, Phys. Rev. Lett. **111**, 226802 (2013).
[6] N. A. Richter, S. Sicolo, S. V. Levchenko, J. Sauer, and M. Scheffler, Phys. Rev. Lett. **111**, 045502 (2013).
[7] S. Ismail-Beigi, Phys. Rev. B **73**, 233103 (2006).
[8] C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross, and A. Rubio, Phys. Rev. B **73**, 205119 (2006).
[9] P. A. Schultz, Phys. Rev. Lett. **96**, 246401 (2006).
[10] G. Makov and M. C. Payne, Phys. Rev. B **51**, 4014 (1995).
[11] S. Lany and A. Zunger, Phys. Rev. B **78**, 235104 (2008).
[12] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, Phys. Rev. Lett. **102**, 016402 (2009).
[13] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, Rev. Mod. Phys. **86**, 253 (2014).
[14] C. Freysoldt and J. Neugebauer, Phys. Rev. B **97**, 205425 (2018).
[15] H.-P. Komsa and A. Pasquarello, Phys. Rev. Lett. **110**, 095505 (2013).
[16] C. W. M. Castleton, A. Höglund, and S. Mirbt, Phys. Rev. B **73**, 035215 (2006).
[17] D. Wang, D. Han, X.-B. Li, S.-Y. Xie, N.-K. Chen, W. Q. Tian, D. West, H.-B. Sun, and S. B. Zhang, Phys. Rev. Lett. **114**, 196801 (2015).
[18] C. L. Fu and K. M. Ho, Phys. Rev. Lett. **63**, 1617 (1989).
[19] C. Sánchez, A. Y. Lozovoi, and A. Alavi, Mol. Phys. **102**, 1045 (2004).
[20] M. Ashton, A. Mishra, J. Neugebauer, and C. Freysoldt, Phys. Rev. Lett. **124**, 176801 (2020).
[21] J. Neugebauer and M. Scheffler, Phys. Rev. B **46**, 16067 (1992).
[22] S. Boeck, C. Freysoldt, A. Dick, L. Ismer, and J. Neugebauer, Comput. Phys. Commun. **182**, 543 (2011).
[23] https://sxrepo.mpie.de.
[24] J. Janssen, S. Surendralal, Y. Lysogorskiy, M. Todorova, T. Hickel, R. Drautz, and J. Neugebauer, Comput. Mater. Sci. **163**, 24 (2019).
[25] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.102.045403 for the full jupyter notebook input file.
[26] C. Freysoldt, S. Boeck, and J. Neugebauer, Phys. Rev. B **79**, 241103(R) (2009).
[27] G. P. Kerker, Phys. Rev. B **23**, 3082 (1981).
[28] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
[29] L. Lin and C. Yang, SIAM J. Sci. Comput. **35**, S277 (2013).
[30] C. Freysoldt, Comput. Mater. Sci. **133**, 71 (2017).
[31] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[32] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[33] https://anaconda.org/conda-forge/sphinxdft.