



**HAL**  
open science

# Accelerating QM/MM simulations of electrochemical interfaces through machine learning of electronic charge densities

Andrea Grisafi, Mathieu Salanne

► **To cite this version:**

Andrea Grisafi, Mathieu Salanne. Accelerating QM/MM simulations of electrochemical interfaces through machine learning of electronic charge densities. *The Journal of Chemical Physics*, 2024, 161 (2), pp.024109. 10.1063/5.0218379 . hal-04643797

**HAL Id: hal-04643797**

**<https://hal.sorbonne-universite.fr/hal-04643797v1>**

Submitted on 10 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Accelerating QM/MM simulations of electrochemical interfaces through machine learning of electronic charge densities

Andrea Grisafi<sup>1, a)</sup> and Mathieu Salanne<sup>2, 3, b)</sup>

<sup>1)</sup>*Institut Sciences du Calcul et des Données, ISCD, Sorbonne Université, F-75005 Paris, France*

<sup>2)</sup>*Physicochimie des Électrolytes et Nanosystèmes Interfaciaux, Sorbonne Université, CNRS, F-75005 Paris, France*

<sup>3)</sup>*Institut Universitaire de France (IUF), F-75231 Paris, France*

A crucial aspect in the simulation of electrochemical interfaces consists in treating the distribution of electronic charge of electrode materials that are put in contact with an electrolyte solution. Recently, it has been shown how a machine-learning method that specifically targets the electronic charge density, also known as SALTED, can be used to predict the long-range response of metal electrodes in model electrochemical cells. In this work, we provide a full integration of SALTED with MetalWalls, a program for performing classical simulations of electrochemical systems. We do so by deriving a spherical harmonics extension of the Ewald summation method, which allows us to efficiently compute the electric field originated by the predicted electrode charge distribution. We show how to use this method to drive the molecular dynamics of an aqueous electrolyte solution under the quantum electric field of a gold electrode, which is matched to the accuracy of density-functional theory. Notably, we find that the resulting atomic forces present a small error of the order of 1 meV/Å, demonstrating the great effectiveness of adopting an electron-density path in predicting the electrostatics of the system. Upon running the data-driven dynamics over about 3 ns, we observe qualitative differences in the interfacial distribution of the electrolyte with respect to the results of a classical simulation. By greatly accelerating quantum-mechanics/molecular-mechanics approaches applied to electrochemical systems, our method opens the door to nanoseconds timescales in the accurate atomistic description of the electrical double layer.

## I. INTRODUCTION

Understanding the interface between a metallic electrode and a liquid electrolyte has become an important objective in chemical physics.<sup>1</sup> This interest is driven by the many underlying applications in electrochemistry, electrocatalysis, and energy storage, but it is also a challenging problem from a fundamental point of view. Most of the time, the conventional simulation methods cannot easily be adopted due to the highly heterogeneous nature of the system, the difficulties associated with the application of a voltage between two electrodes, and the large time and length scales required to simulate a representative portion of the electrochemical interface.<sup>2,3</sup>

In the absence of chemical reactions, intermolecular interactions in bulk electrolytes are often well handled using molecular mechanics (MM) approaches.<sup>4</sup> The main difficulty is therefore to account for the electronic structure of the metal. For this purpose, performing quantum mechanics (QM) calculations, most notably at the density functional theory (DFT) level, appears as the most accurate available approach. Although it is possible to use DFT to calculate atomic forces within a molecular dynamics (MD) simulation, the large computational cost limits its application to relatively simple systems.<sup>5-7</sup> Typically, having an electrolyte including a salt at finite concentration remains a hurdle for such simulations. A possible solution is to use a purely classical approach, which leads

to the family of constant-potential MD simulations.<sup>8</sup> In this case, the polarization of the metal is treated using fluctuating atomic Gaussian charges, which depend on both the electrolyte positions and the applied potential. This approach was refined to effectively include some electronic structure effects,<sup>9,10</sup> but the electrode representation remains highly simplified when compared to an explicit QM-based method.

An alternative is to mix the QM/MM approaches by partitioning the system. In fact, by performing the QM calculation on the electrode subsystem only, it is possible to greatly decrease the computational burden. Such calculations were for example used to study supercapacitors,<sup>11</sup> metal-organic frameworks,<sup>12</sup> electrowetting,<sup>13</sup> and electrocatalytic phenomena.<sup>14-16</sup> There are however several limitations. Firstly, performing a QM calculation at every timestep of a MD simulation ( $\approx$  1fs) is not very efficient, so that multi-stepping techniques have to be introduced by updating the QM system every few thousands steps. Secondly, while the energy of the system can be recovered by adequately summing up the energies of the sub-systems, the same is not always true for the forces. In particular, calculating the electrostatic forces acting on the MM subsystem due to the QM subsystem is not straightforward. The problem is generally tackled by introducing equilibrated partial charges on the electrode atoms,<sup>17</sup> which are computed following a quadratic functional form of the electrostatic energy. Beside the assumed quadratic approximation, this type of charge partitioning is not rigorously defined, so that some important physical effects associated with the local anisotropy of the charge distribution may be lost in the process.

<sup>a)</sup>Electronic mail: andrea.grisafi@sorbonne-universite.fr

<sup>b)</sup>Electronic mail: mathieu.salanne@sorbonne-universite.fr

In this work, we introduce a data-driven method that allows us to improve both the efficiency and the accuracy of QM/MM simulations of electrochemical interfaces. This method is grounded on recent developments in a machine-learning (ML) approach capable of predicting the electronic charge density of metal electrodes, including its non-local response to far-field perturbations.<sup>18</sup> In particular, we implement an interface between the SALTED ML program<sup>19</sup> and the MetalWalls MD code for electrochemical simulations.<sup>20</sup> As a key aspect of this interface, we provide an efficient calculation of the electric field generated by the predicted electronic charge distribution, which we derive following a spherical harmonics generalization of the Ewald summation method. In turn, this allows us to compute accurate electrostatic forces acting on the electrolyte atoms, positioning our method in stark contrast with approaches that are based on partial charges. By shifting the largest share of the computational costs on the learning stage, simulation times of the order of nanoseconds can be achieved with a relatively small overhead compared to conventional full MM simulations. As an application of our method, we study the interface between a gold electrode and an aqueous solution of NaCl at 5 M concentration under applied electric potentials of 0 and 1 V. The structure of the liquid phase as well as interfacial properties such as the capacitance are calculated and compared to previous approaches.

## II. GENERAL METHOD

We adopt a finite-field simulation setup, where a single electrode is put in contact with an electrolyte solution under three-dimensional periodic boundary conditions. This setup can be conveniently used to simulate a pair of equivalent electrodes that are kept at a constant potential difference.<sup>21</sup> In particular, a uniform and constant electric field,  $\varepsilon_z$ , is applied along the direction normal to the electrode surface, such that  $\Delta V = -\varepsilon_z L_z$ , with  $L_z$  the length of the simulation box, corresponds to a physical electrochemical cell potential. The representation of the electrode/electrolyte interface follows the approach already presented in Ref. 18. Specifically, we adopt a quantum-mechanics/molecular-mechanics (QM/MM) strategy, where an *ab initio* representation of the electrode is used in combination with a classical representation of the electrolyte defined via fixed Gaussian charges. We aim to rely on SALTED predictions of the electrode charge density in order to compute accurate electrostatic forces that can be used to drive the dynamics of the system.

### A. Representation of the electrode charge density

Within SALTED, the electronic charge density of the electrode is represented as a linear spherical harmonics expansion that follows density-fitting (DF) approximations

commonly adopted in electronic-structure methods:<sup>22,23</sup>

$$n_e(\mathbf{r}) = \sum_{inlm} c_i^{nlm} \sum_{\mathbf{u}} R_{nl}(|\mathbf{r} - \mathbf{r}_i - \mathbf{u}|) Y_m^l(\widehat{\mathbf{r} - \mathbf{r}_i - \mathbf{u}}). \quad (1)$$

Here, the index  $i$  identifies the electrode atoms,  $nlm$  the basis function indexes, while  $\mathbf{u}$  represent the three-dimensional translation vectors over the periodic images of the system. We assume that radial auxiliary functions are of Gaussian-type, i.e.,  $R_{nl} \propto r^l e^{-r^2/(2\sigma_n^2)}$ , while  $Y_m^l$  are defined as orthonormalized real spherical harmonics.

Reference calculations of  $n_e$  in the presence of the classical electrolyte field are performed using the CP2K simulation program.<sup>24</sup> Specifically, quantum-mechanical calculations are performed at the Kohn-Sham DFT level using the PBE functional<sup>25</sup> with DZVP-MOLOPT-SR basis sets<sup>26</sup> and GTH pseudo-potentials.<sup>27</sup> To extract the expansion coefficients,  $c_i^{nlm}$ , we adopt a density-fitting approximation that makes use of a truncated Coulomb metric,<sup>28,29</sup> a choice that is especially suitable to represent the electrostatic properties of the system.

The total electrode charge density is obtained by adding to  $n_e$  the effective nuclear-charge density associated with the classical local part of the GTH pseudopotential. For each electrode atom  $i$ , this is simply defined by an isotropic Gaussian charge, i.e.,

$$n_{\text{ion}}(r) = \frac{Z_{\text{eff}}}{\sqrt{8\pi^{3/2}} r_{\text{loc}}^3} \exp\left[-\frac{1}{2} \left(\frac{r}{r_{\text{loc}}}\right)^2\right], \quad (2)$$

where  $Z_{\text{eff}}$  is the effective nuclear charge of the electrode atoms and  $r_{\text{loc}}$  is the local pseudopotential cutoff.

### B. SALTED electrode model

In analogy with the prescription adopted in Ref. 18, we will assume that the positions of the electrode atoms are kept fixed, so that the learning can be focused on the sole electrolyte-induced variations of the electrode charge density. Moreover, as the applied voltage  $\Delta V$  is typically of the order of a few volts over a cell that covers several nanometers along  $z$ , we can further assume that the external electric field, defined as  $\varepsilon_z \equiv -\Delta V/L_z$ , introduces a small linear perturbation to  $n_e$ . Under these assumptions, we can conveniently define the SALTED prediction target as the following density difference,

$$\Delta n_e(\mathbf{r}) = n_e(\mathbf{r}; \varepsilon_z) - n_e^0(\mathbf{r}; \varepsilon_z), \quad (3)$$

where  $n_e^0$  is the electron density of the isolated electrode under  $\varepsilon_z$ , used as a constant baseline for the regression procedure. Upon training,  $n_e^0$  is eventually added back to the predicted  $\Delta n_e$  in order to recover the total electron density  $n_e$ . Notably, the possibility of treating the external field as a small perturbation implies that  $\Delta n_e$  can be considered independent from  $\varepsilon_z$ . From a ML point of view, this represents a great advantage, as it allows us to

simulate a continuum spectrum of cell potentials without the need to retrain a new SALTED model whenever a different value of  $\varepsilon_z$  is applied.

Consistently with the density-fitting metric used to generate the reference quantum-mechanical data, SALTED models can be trained by minimizing a loss function that makes use of a (truncated) Coulomb metric.<sup>30</sup> In compact notation, this reads as follows,

$$\begin{aligned} \mathcal{L}(\mathbf{w}) &= \sum_{I=1}^N \langle n_e^{\text{ML}}(\mathbf{w}) - n_e^{\text{DF}} | \frac{1}{\mathbf{r} - \mathbf{r}'} | n_e^{\text{ML}}(\mathbf{w}) - n_e^{\text{DF}} \rangle_I \\ &= \sum_{I=1}^N [\mathbf{c}_I^{\text{ML}}(\mathbf{w}) - \mathbf{c}_I^{\text{DF}}]^T \mathbf{J}_I [\mathbf{c}_I^{\text{ML}}(\mathbf{w}) - \mathbf{c}_I^{\text{DF}}], \end{aligned} \quad (4)$$

where  $N$  is the number of training configurations. In the second line of this equation,  $\mathbf{J}$  is the matrix of 2-centers Coulomb integrals between the auxiliary functions, i.e.,  $\langle i'n'l'm' | r^{-1} | inlm \rangle$ , as computed by CP2K,  $\mathbf{c}^{\text{DF}}$  and  $\mathbf{c}^{\text{ML}}$  are the density-fitting and machine-learning vectors of density coefficients associated with the expansion of  $\Delta n_e$ , respectively, and  $\mathbf{w}$  is the vector of regression weights. In practice, a regularization term is also added to the loss function to avoid overfitting; we refer to Ref. 22 for an in-depth discussion of the SALTED method.

Following Ref. 18, the symmetry-adapted kernels used for the approximation of the  $\mathbf{c}^{\text{ML}}$  density coefficients can be defined to include long-range information. In particular, kernel functions between atomic environments are constructed from long-distance equivariant (LODE) structural descriptors,<sup>31–33</sup> as implemented in the “rascaline” atomistic representations package.<sup>34</sup> This feature of the method is fundamental to capture the non-local variations of the electron density induced by electrolyte charges that are located arbitrarily far from the electrode surface.<sup>18</sup> We remark that, unlike charge equilibration schemes, this is made possible without relying on any self-consistent procedure. In fact, the use of long-range descriptors allows us to learn directly the self-consistent electronic polarization of the system, thus bypassing the need to incorporate self-consistency at the ML level. This would still hold true in the scenario where the electrolyte is also undergoing an electronic polarization.

### III. ELECTRIC FIELD CALCULATION

From the predicted electrode charge distribution, a nontrivial problem is that of performing an efficient calculation of the electric field needed to drive the dynamics of the electrolyte. To achieve this goal, we provide in this section an original analytical derivation of the Ewald summation method, properly extended to account for the spherical harmonics expansion of the electron density.

Following the representation of Eq. (1), we start by introducing a fictitious electron density,  $n_e^{\text{ewald}}$ , obtained by combining the physical density coefficients,  $c_i^{nlm}$ , with

a new set of smooth radial functions,  $R_{nl}^{\text{ewald}}$ . Crucially, these functions are defined from a Gaussian width  $\sigma_{\text{ewald}}$  that is larger than those entering the definition of  $R_{nl}$ . Following the conventional Ewald summation method, we then separate from the electron density a charge-neutral (screened) contribution, defined as

$$n_e^{\text{screen}}(\mathbf{r}) \equiv n_e(\mathbf{r}) - n_e^{\text{ewald}}(\mathbf{r}). \quad (5)$$

By construction, the electric field generated by  $n_e^{\text{screen}}$  is short-ranged, and can be efficiently computed in real space. Conversely, the remaining (unscreened)  $n_e^{\text{ewald}}$  density will contribute to a long-range electric field that can be efficiently computed in reciprocal space. In what follows, we outline the derivation of both electronic and nuclear contributions to the electric field, as newly implemented in the MetalWalls simulation program.

#### A. Short-range electronic contribution

Starting from  $n_e^{\text{screen}}$ , the short-range part of the Hartree electronic potential can be analytically computed in real space by relying on the Laplace expansion of the Coulomb operator. After some calculations, detailed in the Supplementary Material, we obtain

$$\begin{aligned} \phi_{\text{H}}^{\text{SR}}(\mathbf{r}) &= \int_{\mathbf{R}^3} d\mathbf{r}' \frac{n_e^{\text{screen}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\ &= 4\pi \sum_{i=1}^{N_{\text{at}}} \sum_{nlm} \frac{c_i^{nlm}}{2l+1} \sum_{\mathbf{u}} Y_{lm}(\mathbf{r} - \widehat{\mathbf{r}_i - \mathbf{u}}) \\ &\times \left[ \frac{I_{nl}^<(|\mathbf{r} - \mathbf{r}_i - \mathbf{u}|)}{|\mathbf{r} - \mathbf{r}_i - \mathbf{u}|^{l+1}} + |\mathbf{r} - \mathbf{r}_i - \mathbf{u}|^l I_{nl}^>(|\mathbf{r} - \mathbf{r}_i - \mathbf{u}|) \right]. \end{aligned} \quad (6)$$

Here,  $I_{nl}$  indicate radial integrals performed within (<) and outside (>) a sphere of radius  $r \equiv |\mathbf{r} - \mathbf{r}_i - \mathbf{u}|$ :

$$\begin{aligned} I_{ln}^<(r) &= \int_0^r ds s^{2+l} [R_{nl}(s) - R_{nl}^{\text{ewald}}(s)], \\ I_{ln}^>(r) &= \int_r^\infty ds s^{1-l} [R_{nl}(s) - R_{nl}^{\text{ewald}}(s)]. \end{aligned} \quad (7)$$

While the individual terms of  $I_{ln}^>$  are quickly vanishing for  $r \rightarrow \infty$ , the requirement that  $I_{ln}^<$  vanishes sufficiently fast for increasing values of  $r$  implies that the normalization of the Ewald radial functions must be chosen so that to satisfy the following equality:

$$\int_0^\infty ds s^{2+l} R_{nl}^{\text{ewald}}(s) = \int_0^\infty ds s^{2+l} R_{nl}(s). \quad (8)$$

This choice guarantees that both types of radial integrals in Eq. (7) tend to zero with superexponential rapidity, thus dominating over the algebraic terms of the order of  $1/r^{l+1}$  and  $r^l$  in Eq. (6). Crucially, this screening effect allows us to compute the short-range part of the Hartree potential by only collecting contributions of electrode



atoms that fall within a finite cutoff distance. In particular, we select a cutoff of  $r_{\text{cut}}^l = 4\sqrt{2+l}\sigma_{\text{Ewald}}$ , which is defined to account for the spatial range covered by the Ewald Gaussian-type functions.

To compute the electric field as  $\mathbf{E}_{\text{H}}^{\text{SR}} = -\nabla\phi_{\text{H}}^{\text{SR}}$ , it is convenient to adopt the substitution  $\mathbf{r} - \mathbf{r}_i - \mathbf{u} \equiv (r, \theta, \phi)$  in Eq. (6), so that to make use of the gradient in spherical coordinates, i.e.,

$$\tilde{\nabla} \equiv \left( \frac{\partial}{\partial r}, \frac{1}{r} \frac{\partial}{\partial \theta}, \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \right).$$

Upon this change of variables, the contribution to the short-range part of the electric field brought by each electrode atom can be analytically computed. In compact notation, we have

$$\mathbf{E}_{\text{H}}^{\text{SR}}(\mathbf{r}) = 4\pi \sum_{i=1}^{N_{\text{at}}} \sum_{nlm} \frac{c_i^{nlm}}{2l+1} \sum_{\mathbf{u}} \mathbf{U}(\mathbf{r} - \widehat{\mathbf{r}}_i - \mathbf{u}) \cdot \tilde{\nabla} \left[ \left( \frac{1}{r^{l+1}} I_{ln}^<(r) + r^l I_{ln}^>(r) \right) Y_m^l(\theta, \phi) \right], \quad (9)$$

where

$$\mathbf{U} = \begin{pmatrix} \sin \theta \cos \phi & \cos \theta \cos \phi & -\sin \phi \\ \sin \theta \sin \phi & \cos \theta \sin \phi & \cos \phi \\ \cos \theta & -\sin \theta & 0 \end{pmatrix}, \quad (10)$$

represents the unitary transformation matrix between the Cartesian and spherical unit vectors  $(\hat{x}, \hat{y}, \hat{z}) \rightarrow (\hat{r}, \hat{\theta}, \hat{\phi})$ , such that  $\nabla = \mathbf{U} \cdot \tilde{\nabla}$  recovers the Cartesian gradient. Explicit analytical expressions of radial integrals and spherical harmonics derivatives are reported in the Supplementary Material.

## B. Long-range electronic contribution

From a smooth choice of  $n_e^{\text{ewald}}$ , the long-range part of the Hartree potential can be conveniently written as a plane-waves expansion compatible with the three-dimensional periodicity of the system:

$$\phi_{\text{H}}^{\text{LR}}(\mathbf{r}) = \sum_{\mathbf{k} \neq 0} \frac{4\pi}{k^2} \tilde{n}_e^{\text{ewald}}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{r}}. \quad (11)$$

Note that we excluded the  $\mathbf{k} = 0$  term because of the assumed charge neutrality of the overall system, i.e., electrons and nuclei. From the expansion of the plane wave in spherical harmonics, the Fourier components of the electron density are found to get the form of

$$\begin{aligned} \tilde{n}_e^{\text{ewald}}(\mathbf{k}) &= \frac{1}{\Omega} \int_{\Omega} d\mathbf{r} e^{-i\mathbf{k} \cdot \mathbf{r}} n_e^{\text{ewald}}(\mathbf{r}) \\ &= \frac{4\pi}{\Omega} \sum_{i=1}^{N_{\text{at}}} e^{-i\mathbf{k} \cdot \mathbf{r}_i} \sum_{nlm} c_i^{nlm} (-i)^l Y_m^l(\hat{\mathbf{k}}) \hat{I}_{nl}(k), \end{aligned} \quad (12)$$

where  $\Omega$  is the cell volume. Here,  $\hat{I}_{nl}$  are defined as the following radial integrals,

$$\hat{I}_{nl}(k) = \int_0^\infty dr r^2 j_l(kr) R_{nl}^{\text{ewald}}(r), \quad (13)$$

with  $j_l$  a spherical Bessel function of the first kind. Upon plugging Eq. (12) into Eq. (11), the final analytical expression for the long-range Hartree potential is

$$\begin{aligned} \phi_{\text{H}}^{\text{LR}}(\mathbf{r}) &= \frac{32\pi^2}{\Omega} \sum_{\mathbf{k} \neq 0}^{\text{half}} \frac{1}{k^2} \sum_{i=1}^{N_{\text{at}}} \sum_{nlm} c_i^{nlm} \hat{I}_{nl}(k) Y_m^l(\hat{\mathbf{k}}) \\ &\times \begin{cases} (-1)^{\frac{l}{2}} \cos[\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}_i)], & l\%2 = 0 \\ (-1)^{\frac{l+3}{2}} \sin[\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}_i)], & l\%2 \neq 0 \end{cases}. \end{aligned} \quad (14)$$

Note that we exploited the real nature of the electron density to sum  $\mathbf{k}$ -vectors only over a semisphere in reciprocal space of radius  $k_{\text{cut}}$ . In practice,  $k_{\text{cut}}$  is indirectly defined by the chosen value of  $\sigma_{\text{Ewald}}$ , following the numerical convergence criterion implemented in MetalWalls.<sup>35</sup>

Upon Cartesian differentiation of Eq. (14), the long-range part of the electronic electric field is finally given by

$$\begin{aligned} \mathbf{E}_{\text{H}}^{\text{LR}}(\mathbf{r}) &= \frac{32\pi^2}{\Omega} \sum_{\mathbf{k} \neq 0}^{\text{half}} \frac{\mathbf{k}}{k^2} \sum_{i=1}^{N_{\text{at}}} \sum_{nlm} c_i^{nlm} \hat{I}_{nl}(k) Y_m^l(\hat{\mathbf{k}}) \\ &\times \begin{cases} (-1)^{\frac{l}{2}} \sin[\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}_i)], & l\%2 = 0 \\ (-1)^{\frac{l+1}{2}} \cos[\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}_i)], & l\%2 \neq 0 \end{cases}. \end{aligned} \quad (15)$$

An explicit expression of the radial integrals, together with additional implementation details, are reported in the Supplementary Material.

## C. Pseudopotential nuclear contribution

From Eq. (2), computing the nuclear contribution to the electric field is well established,<sup>36</sup> and it reduces to the problem of treating a set of classical Gaussian charges. Following the conventional Ewald method, this amounts to compute a screened short-range potential as

$$\phi_{\text{ion}}^{\text{SR}}(r) = \frac{Z_{\text{eff}}}{r} \left[ \text{erf} \left( \frac{r}{\sqrt{2}r_{\text{loc}}} \right) - \text{erf} \left( \frac{r}{\sqrt{2}\sigma_{\text{ewald}}} \right) \right], \quad (16)$$

and a long-range contribution in reciprocal space, whose Fourier components are given by

$$\tilde{\phi}_{\text{ion}}^{\text{LR}}(k) = \frac{4\pi}{\Omega} \frac{Z_{\text{eff}}}{k^2} \exp \left( -\frac{1}{2} k^2 \sigma_{\text{ewald}}^2 \right). \quad (17)$$

For consistency, we select a value of  $\sigma_{\text{ewald}}$  equivalent to that adopted for the electronic field calculation.

The short-range part of the electric field can be computed by relying on the same transformation matrix between Cartesian and spherical unit vectors already introduced in Eq. (10). Considering that Eq. (16) has no

angular components, we can write

$$\mathbf{E}_{\text{ion}}^{\text{SR}}(\mathbf{r}) = \sum_{i=1}^{N_{\text{at}}} \sum_{\mathbf{u}} \mathbf{U}(\mathbf{r} - \widehat{\mathbf{r}_i - \mathbf{u}}) \cdot \left[ -\frac{\partial}{\partial r} \phi_{\text{ion}}^{\text{SR}}(r), 0, 0 \right]. \quad (18)$$

The long-range part of the nuclear electric field is finally computed as follows,

$$\mathbf{E}_{\text{ion}}^{\text{LR}}(\mathbf{r}) = 2 \sum_{\mathbf{k} \neq 0}^{\text{half}} \mathbf{k} \tilde{\phi}_{\text{ion}}^{\text{LR}}(\mathbf{k}) \sum_{i=1}^{N_{\text{at}}} \sin[\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}_i)]. \quad (19)$$

#### IV. SALTED/METALWALLS INTERFACE

Upon implementing the electric field just derived into the MetalWalls molecular dynamics engine, the SALTED/MetalWalls simulation can be easily run thanks to a Python-based interface between the two programs. In what follows, we report a description of the simulation workflow, including pseudocode of the relevant steps.

1. Import SALTED and MetalWalls modules:
 

```
import salted, metalwalls
```
2. Load pretrained SALTED information:
 

```
salted_info = salted.init_pred.build()
```
3. Load isolated electrode information:
 

```
structure = ase.io.read("init.xyz")
coords_electrode =
structure.get_positions()[640:]
coefs_0 = np.loadtxt("coefs_isolated.txt")
```
4. Define MetalWalls variables.
 

```
my_system =
metalwalls.mw_system.MW_system_t()
my_algorithms =
metalwalls.mw_algorithms.MW_algorithms_t()
my_parallel =
metalwalls.mw_parallel.MW_parallel_t()
```
5. Initialize electrode charge density:
 

```
coords_electrolyte = my_system.xyz_ions
coords = np.vstack((coords_electrolyte,
coords_electrode))
structure.set_positions(coords)
my_system.coefs = salted.salted_prediction
.build(salted_info, structure)
my_system.coefs += coefs_0
```
6. Initialize MetalWalls simulation:
 

```
metalwalls.mw_tools.run_step(my_system,
my_parallel, my_algorithms, 0, 0,
do_output, step_output_frequency)
```

7. Start the dynamics:
 

```
for step in range(my_system.num_steps):
```
8. Update electrolyte atomic positions:
 

```
metalwalls.mw_tools.step_setup(my_system,
my_algorithms, my_parallel, step)
coords_electrolyte = my_system.xyz_ions
coords = np.vstack((coords_electrolyte,
coords_electrode))
structure.set_positions(coords)
```
9. Update electrode charge density:
 

```
my_system.coefs = salted.salted_prediction
.build(salted_info, structure)
my_system.coefs += coefs_0
```
10. Compute atomic forces:
 

```
metalwalls.mw_tools.step_compute_forces
(my_system, my_parallel, step)
```
11. Continue from point 8.

Importantly, the electrostatic forces acting on the electrolyte atoms  $j$  are computed at step 10 as follows,

$$\mathbf{f}_j = q_j \mathbf{E}(\mathbf{r}_j), \quad (20)$$

where  $q_j$  are the electrolyte partial atomic charges, while

$$\mathbf{E} = \mathbf{E}_{\text{H}}^{\text{SR}} + \mathbf{E}_{\text{H}}^{\text{LR}} + \mathbf{E}_{\text{ion}}^{\text{SR}} + \mathbf{E}_{\text{ion}}^{\text{LR}}, \quad (21)$$

is the total electric field generated by the predicted electrode charge distribution, as decomposed from the derivation carried out in Sec. III.

#### V. RESULTS

To test the simulation workflow previously described, we revisit the example of ionic capacitor introduced in Ref. 18. In particular, we consider a Au(100) electrode made of 4 unit cell repetitions along the  $xy$  plane and 7 atomic layers along  $z$ , which is put in contact with an aqueous solution of NaCl 5 M under 3D periodic boundary conditions (Fig. 1). The electrolyte is made of 200 water molecules and 20 NaCl ions pairs, which, together with the gold electrode, occupy a box size of  $L_{x/y} = 11.54 \text{ \AA}$  and  $L_z = 64.34 \text{ \AA}$ . Classical electrolyte interactions are defined using a TIP4P/2005 water model<sup>37</sup> with  $\text{Na}^+$  and  $\text{Cl}^-$  Lennard-Jones parameters taken from Ref. 38; fixed classical atomic charges are given by  $q_{\text{M}} = -1.1128$ ,  $q_{\text{H}} = +0.5564$  and  $q_{\text{Na/Cl}} = \pm 1.0$ .

Following the discussion reported in Sec. II, we apply a uniform electric field along the direction orthogonal to the Au(100) surface, allowing us to mimic a pair of equivalent gold electrodes that are kept at a constant voltage drop. In particular, we consider two different cell potentials, namely  $\Delta V = 0 \text{ V}$  and  $\Delta V = 1 \text{ V}$ , which correspond to

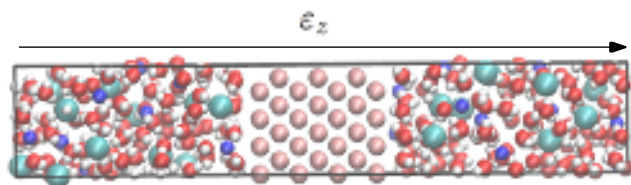


FIG. 1. Representation of an ionic capacitor made of a Au(100) gold electrode put in contact with a water/NaCl electrolyte solution under 3D periodic boundary conditions. According to Ref. 21, a finite electric field  $\varepsilon_z$  is applied along  $z$  to mimic a voltage drop of  $\Delta V = -\varepsilon_z L_z$  between two parallel electrodes.

applied fields of  $\varepsilon_z = 0.0 \text{ V/\AA}$  and  $\varepsilon_z = 0.016 \text{ V/\AA}$ , respectively. Reference QM/MM calculations are performed for 2000 uncorrelated frames selected from a classical Metal-Walls trajectory obtained running a finite-field simulation at a cell voltage of 1 V. The basis set used for the gold electron density representation is defined from an uncontracted version of the automatically generated auxiliary functions implemented in CP2K.<sup>39</sup> In particular, we adopt an auxiliary basis set that includes up to  $l_{\text{max}} = 4$  spherical harmonics, accounting for a total of 182 Gaussian-type functions per gold atom.

#### A. Validation of SALTED electrode model

From the QM/MM dataset so generated, SALTED models of the electrode charge density are trained following the discussion of Sec. II-B. In practice, a subset of  $M = 400$  gold atomic environments is selected to recast the learning problem into a low dimensional space.<sup>22</sup> Symmetry-adapted kernel functions are computed from LODE descriptors that include both atom-density and potential-like structural features.<sup>18,32</sup> In this work, these are defined using Gaussian widths of  $\sigma = 0.5 \text{ \AA}$  and  $\sigma = 2 \text{ \AA}$ , respectively, together with a radial cutoff of  $r_{\text{cut}} = 10 \text{ \AA}$ . To validate the model, we select 1600 random configurations for training and retain the remaining 400 for testing. The prediction error is measured consistently with the definition of the SALTED loss function reported in Eq. (4). In particular, we compute the percentage root mean square error as

$$\% \text{RMSE} = \sqrt{\frac{\mathcal{L}_{\text{test}}(\mathbf{w})}{\sigma_{\text{test}}^2}} \times 100, \quad (22)$$

where  $\mathcal{L}_{\text{test}}(\mathbf{w})$  follows from Eq. (4), while

$$\sigma_{\text{test}}^2 = \sum_{I=1}^{N_{\text{test}}} [\mathbf{c}_I^{\text{DF}}]^T \mathbf{J}_I \mathbf{c}_I^{\text{DF}}, \quad (23)$$

represents the variance of the learning target computed over the test set. Learning curves are reported in the Supplementary Material.

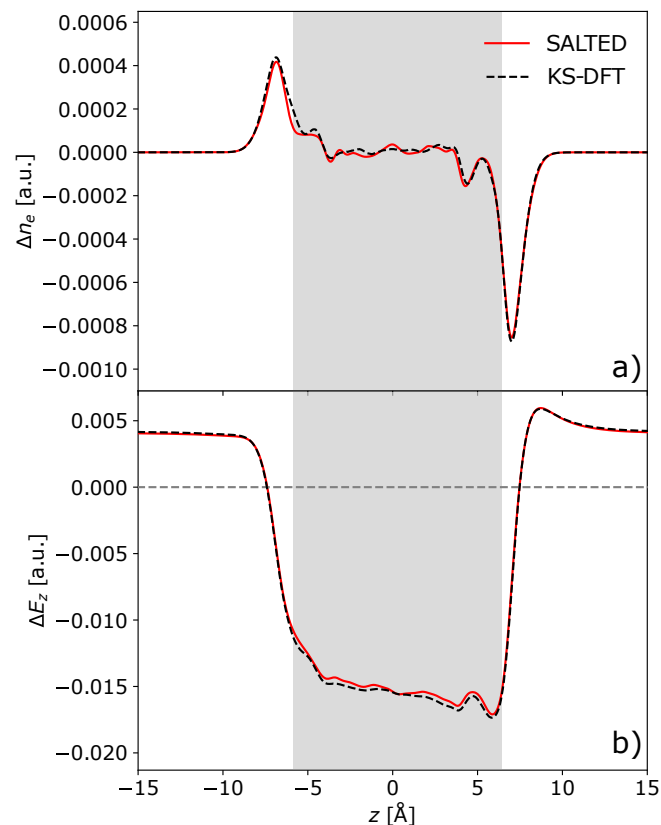


FIG. 2. Electrode charge density response (a) and derived electric field (b) induced by the classical charges of a test electrolyte configuration. Red lines: SALTED prediction results. Black dashed lines: Kohn-Sham DFT reference. Shaded gray area indicates the region occupied by the gold electrode.

With as few as 200 training structures, the SALTED error saturates at  $\sim 3\%$  RMSE, demonstrating the capability of the model of capturing the non-local fluctuations of the electronic charge induced by different spatial arrangements of the electrolyte atoms. This result is confirmed by the excellent agreement between the Kohn-Sham and predicted electron density variation  $\Delta n_e$ , as depicted in Fig. 2-a for a representative test configuration. Despite the slightly less accurate prediction of the charge density profile with respect to what reported in Ref. 18, we find that the use of a Coulomb metric in the SALTED loss function allows us to obtain a highly accurate electric field. This is illustrated in Fig. 2-b, where the longitudinal variation of the predicted electric field with respect to the field of the isolated electrode,  $\Delta E_z$ , is reported against the quantum-mechanical reference. Notably, we observe that the agreement with DFT is especially good outside the electrode region, being the most relevant for driving the dynamics of the electrolyte.

As an ultimate proof of the model accuracy, we compute the predicted Cartesian components of the electrostatic atomic forces associated with the 400 electrolyte configurations used for testing, as compared against the quantum-mechanical reference. Both reference and pre-

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0218379

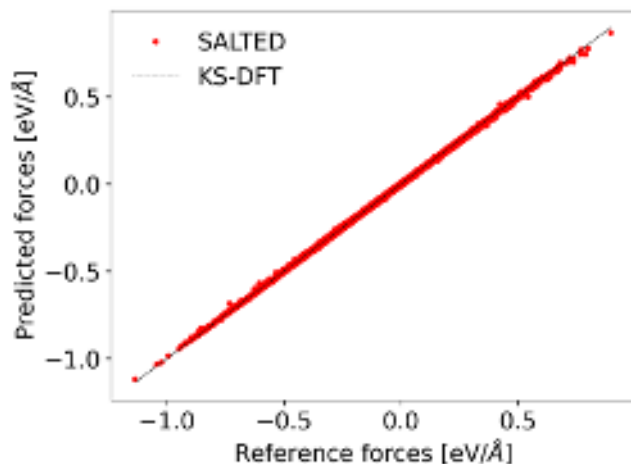


FIG. 3. Correlation plot between the predicted and reference electrostatic forces generated by the electric field of a gold electrode, as computed over 256'000 electrolyte atoms belonging to 400 test configurations. Red dots: SALTED prediction. Black dashed line: KS-DFT reference. Predicted forces are associated with a RMSE = 1.0 meV/Å over a standard deviation of  $\sigma_f = 141.9$  meV/Å.

dicted forces are computed as in Eq. (20), starting from the DFT and SALTED electron-density coefficients, respectively. Consistently with the high level of statistical correlation shown in Fig. 3, we obtain a remarkably small root mean square error of 1.0 meV/Å, which amounts to about 0.7% of the standard deviation of the DFT forces in the test set. By and large, these results highlight the intrinsic effectiveness of a machine learning model that has the electron density as a prediction target to compute the electrostatic properties of the system.

### B. Simulation of ionic capacitor

We now report the results of the data-driven molecular dynamics, which is run following the simulation workflow described in Sec. IV. Finite-field simulations associated with applied cell potentials of  $\Delta V = 0$  V and  $-1$  V are performed at room temperature in timesteps of 2 fs using a Nosé-Hoover thermostat.<sup>40</sup> Predictions of the electrostatic atomic forces are obtained from a SALTED model trained on  $N = 2000$  configurations, as validated from the discussion previously carried out. By parallelizing the calculation over an Intel i9-12900 CPU, we find that each step of the dynamics takes about 0.5 s, corresponding to a speedup  $> 10^3$  with respect to first-principles QM/MM calculations. This substantial increase in computational efficiency results to be decisive for achieving timescales of the order of nanoseconds, thus making it possible to access the slow dynamics of the electrolyte while preserving a highly accurate representation of electrode charge density. For both applied voltages, in particular, we are able to reach long simulation times of  $\approx 3$  ns.

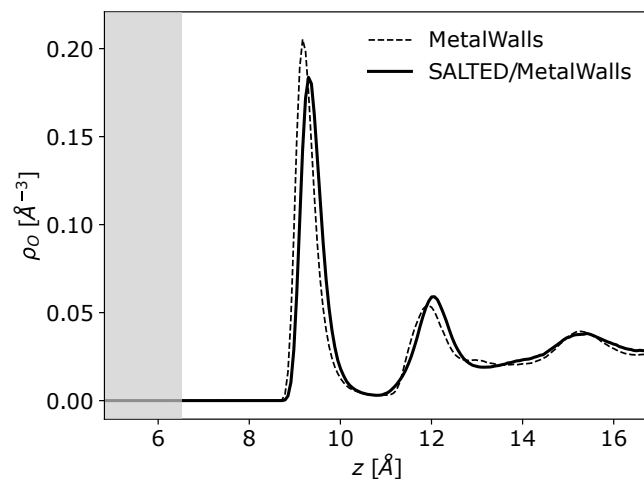


FIG. 4. Longitudinal density profiles of oxygen atoms on the right hand side of the ionic capacitor studied in this work, subject to a null cell potential of  $\Delta V = 0$  V. Dashed line: classical MetalWalls simulation. Full line: SALTED/MetalWalls simulation. Shaded gray area indicates the gold electrode.

It is instructive to compare our results against those of a classical MetalWalls simulation. In this case, the electronic polarization of the gold electrode is approximated through a charge-equilibration scheme that makes use of fluctuating Gaussian charges.<sup>35</sup> Specifically, we represent the partial charges on the gold atoms through a Gaussian width of  $\sigma_{MW} = 1.06$  Å, and simulate classical trajectories for a total time that is, once again, of about 3 ns. We note that similar simulations are only 50 times faster than those performed through our method. This is remarkable if considering that, in addition to the SALTED predictions, quantum electrostatic forces are computed from the charge density generated by 182 spherical harmonics functions per gold atom, rather than from a single isotropic Gaussian function.

Regardless from the applied potential, we find that the thermal distribution of water molecules obtained through our method displays a systematic shift with respect to the result of a classical simulation. In particular, we observe a distancing of about 0.5 Å of the first adsorption peak, which comes along with overall smoother density profiles. This is shown in Fig 4 in the case of a vanishing applied voltage, indicating that a classical electrode model tends to give a stiffer potential of mean force with respect to what expected from a first-principles approach.

More striking differences between the classical and data-driven simulations emerge when looking at the results obtained under an applied potential  $\Delta V = 1$  V. As an example, we report in Fig. 5-a the ionic density profiles of  $\text{Na}^+$  and  $\text{Cl}^-$  associated with the screening of the negatively charged metal surface. While the distribution of chlorine shows an overall agreement between the two methods, a qualitative difference is found in the adsorption of sodium cations. In particular, a double pick is produced as a result of the classical simulation, which does not



This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0218379

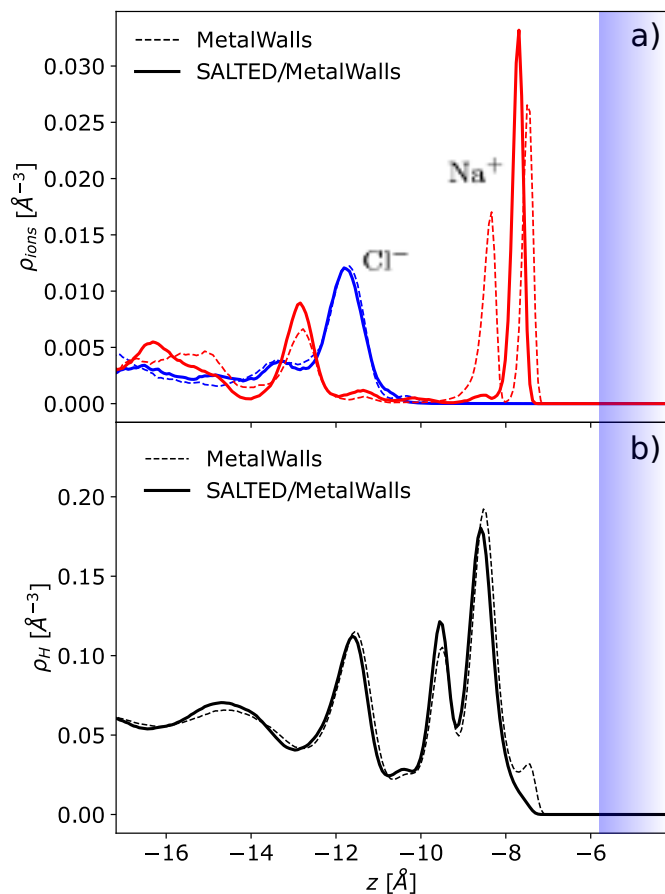


FIG. 5. Longitudinal density profiles of sodium chloride (a), and hydrogen atoms (b) at the negatively charged electrode surface of the ionic capacitor studied in this work, subject to an applied cell potential of  $\Delta V = 1$  V. Dashed lines: classical MetalWalls simulation. Full lines: SALTED/MetalWalls simulation. Red and blue color codes refer to  $\text{Na}^+$  and  $\text{Cl}^-$  ions, respectively. Shaded blue area indicates the gold electrode.

appear when using SALTED to predict the quantum-mechanical response of the electrode. This artifact of the classical model can be associated with the rigidity by which the distribution of the metal charge is represented, which mostly affects the accuracy of the electric field in the proximity of the gold electrode. In fact, it has already been shown that, depending on the chosen value of  $\sigma_{\text{MW}}$ , MetalWalls simulations can yield substantial variations in the distribution of  $\text{Na}^+$  at the Au(100) surface.<sup>41</sup>

In addition to the ionic profiles, a clear-cut discrepancy is also observed in the distribution of hydrogen atoms; this is illustrated in Fig. 5-b. Interestingly, we find that a prepeak is produced by the classical simulation at about 2 Å from the gold slab. This prepeak is associated with an occasional reorientation of the water molecules that expose the hydrogen atoms towards the electrode, a phenomenon that is however not predicted by our method. This is confirmed by an analysis of the orientation of the water dipoles with respect to the normal to the electrode surface, as reported in the Supplementary Material.

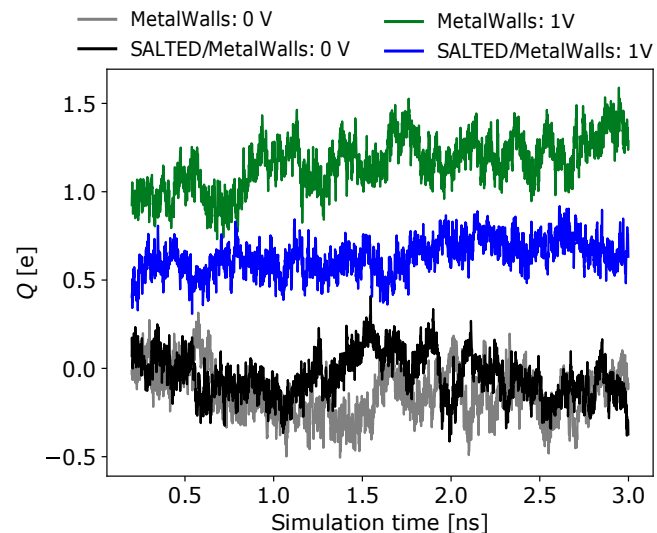


FIG. 6. Evolution of the electrode surface charge in the ionic capacitor studied in this work. Black and blue lines refer to SALTED/MetalWalls simulations at 0 and 1 V, respectively. Gray and light blue lines refer to classical MetalWalls simulations at 0 and 1 V, respectively.

A further benefit of driving the dynamics of the electrochemical interface through electron-density predictions is that of giving access to the total charge  $\pm Q$  accumulated at the electrode surfaces. In a finite-field simulation setup,  $Q$  can be computed by integrating  $\Delta n_e$  from the middle of the metal slab to the classical electrolyte region.<sup>18</sup> In Fig. 6, we report the time evolution of  $Q$ , computed every picosecond, for both null and finite cell potentials. When considering the results at  $\Delta V = 0$  V, we observe a fluctuation of the electrode charge around 0 for both classical and SALTED-based simulations. This is to be expected from the symmetry of the isolated metallic slab, implying that no net polarization can be found at thermal equilibrium without an externally applied electric bias. Conversely, simulations run at  $\Delta V = 1$  V lead to a finite average charge  $\langle Q \rangle$ . In particular, we find that the computed classical value of  $\langle Q \rangle$  is about twice as large than what predicted through our method, i.e., 1.16e versus 0.63e, respectively. This difference in the system's electric capacitance is due to a substantial overestimation of the electronic polarization of the metallic slab obtained through the classical model, a result that was already hinted in Ref. 18 when comparing predictions of  $\Delta n_e$  performed on a test classical trajectory. Although a better agreement might be obtained by tuning the value of the classical Gaussian width  $\sigma_{\text{MW}}$ , the observed discrepancy suggests that disposing of a physically sound electrode model is preferable to obtain reliable results.

In addition to  $\langle Q \rangle$ , having access to the thermal fluctuations of  $Q$  allows us to provide an estimate of the differential capacitance associated with the formation of the electrical double layer:<sup>42</sup>

$$C_{\text{diff}}^{\text{EDL}} = \beta \langle (Q - \langle Q \rangle)^2 \rangle. \quad (24)$$

At  $\Delta V = 0$  V, a similar differential capacitance is obtained for both type of simulations, resulting in a value of  $C_{\text{diff}}^{\text{EDL}} = 8.9 \mu\text{F}/\text{cm}^2$  for the classical model, and a slightly reduced value of  $C_{\text{diff}}^{\text{EDL}} = 7.4 \mu\text{F}/\text{cm}^2$  for the data-driven model. This discrepancy is magnified when looking at the results at  $\Delta V = 1$  V. In line with what already computed in Ref. 18, the classical simulation is associated with a value of  $C_{\text{diff}}^{\text{EDL}} = 10.1 \mu\text{F}/\text{cm}^2$ . Conversely, our SALTED/MetalWalls simulation comes along with much more attenuated charge fluctuations, yielding a differential capacitance that is about twice as small, i.e.,  $C_{\text{diff}}^{\text{EDL}} = 4.6 \mu\text{F}/\text{cm}^2$ . While simulations longer than 3 ns would be required to provide a tighter convergence of  $C_{\text{diff}}^{\text{EDL}}$ , these results remark, once again, the need of introducing an accurate representation of the electrode in the description of the electrochemical interface.

## VI. CONCLUSIONS

The presented method represents a relevant example of how to incorporate a machine-learning approach that addresses the prediction of electronic-structure properties into the atomistic study of complex materials. By providing a formally rigorous integration between the SALTED and MetalWalls programs, we have shown how to extend the reach of applications of QM/MM approaches that aim at preserving a quantum description of metal electrodes in the simulation of electrochemical interfaces. Most notably, the possibility of driving the dynamics of the electrolyte over nanoseconds timescales represents a great leap forward towards obtaining a thermally relaxed, yet accurate, description of the electrical double layer.

By taking an ionic capacitor as an example, we have found that disposing of a quantum treatment of the electrode charge density represents a critical aspect for predicting the structural properties of the electrolyte at the interface. In particular, our results remark how a classical electrode model cannot reproduce the complex nature of the system's electrostatics close to the metal surface. This is not surprising considering that strong electric fields are known to be encountered whenever an electrolyte atom penetrates the electronic cloud of the metal electrode,<sup>43</sup> an effect that is especially difficult to include through the use of classical atomic charges. In this regard, an interesting question will be that of understanding how our method compares against data-driven charge equilibration schemes based on partial charges, where a ML representation of atomic electronegativities,<sup>44</sup> electric dipoles,<sup>45</sup> and/or polarizabilities,<sup>46</sup> is adopted to predict the electrostatics of the system. We mention, in particular, a recent ML interface with MetalWalls that can be used to refine the atomic Gaussian charges of the electrode by relying on a linear density-response formalism.<sup>47</sup>

In perspective, training SALTED on the electron density of different types of (possibly defected) metal surfaces will make it possible to exploit the transferability of the ML model to predict the electrical double layer structure

of realistic energy materials. Moreover, a notable extension of our method could involve promoting redox active molecular species to the QM region, thus enabling the simulation of electrochemical reactions in the presence of an explicit ionic solution. In this context, an important aspect will be that of integrating our approach with already existing ML interatomic potentials that purely rely on the local structural information around the atoms of the system.<sup>48,49</sup> This possibility would be especially attractive both for studying the charging mechanisms in pseudocapacitive materials,<sup>50</sup> as well as for rationalizing the role of the electrolyte in electrocatalytic processes.<sup>51</sup> Furthermore, disposing of a three-dimensional map of the electronic charge density will constitute an added value of our method in view of predicting the regioselectivity of the catalyst to electron-transfer phenomena.

## SUPPLEMENTARY MATERIAL

See the Supplementary Material for a complete derivation of the electronic electric field, together with complementary results.

## AUTHOR DECLARATIONS

The authors have no conflicts to disclose.

## SOURCE CODE

Printing of density coefficients and 2-center auxiliary integrals is made available under the official trunk version of CP2K.<sup>52</sup> The latest SALTED implementation used to produce the results of this work can be accessed from GitHub.<sup>53</sup> A version of MetalWalls suitable to be interfaced with SALTED can be downloaded from GitLab.<sup>54</sup>

## DATA AVAILABILITY

The data that support the findings of this study are openly available from Zenodo.<sup>55</sup>

## ACKNOWLEDGMENTS

We thank Alessandra Serva, Laura Scalfi and Augustin Bussy for useful discussion. This work was supported by the French National Research Agency under the France 2030 program (Grant ANR-22-PEBA-0002).

<sup>1</sup>J. M. Dawlaty, S. Perkin, M. Salanne, and A. P. Willard, *The Journal of Chemical Physics* **159**, 150401 (2023).

<sup>2</sup>K. Schwarz and R. Sundararaman, *Surface Science Reports* **75**, 100492 (2020).

<sup>3</sup>A. Gro, *Current Opinion in Electrochemistry* **40**, 101345 (2023).

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0218379

- <sup>4</sup>K. Goloviznina, J. N. Canongia Lopes, M. Costa Gomes, and A. A. H. Pdua, *Journal of Chemical Theory and Computation* **15**, 5858 (2019).
- <sup>5</sup>C.-Y. Li, J.-B. Le, Y.-H. Wang, S. Chen, Z.-L. Yang, J.-F. Li, J. Cheng, and Z.-Q. Tian, *Nature Materials* **18**, 697 (2019).
- <sup>6</sup>R. Khatib, A. Kumar, S. Sanvito, M. Sulpizi, and C. S. Cucinotta, *Electrochimica Acta* **391**, 138875 (2021).
- <sup>7</sup>A. Chen, J.-B. Le, Y. Kuang, and J. Cheng, *The Journal of Chemical Physics* **157**, 094702 (2022).
- <sup>8</sup>G. Jeanmairat, B. Rotenberg, and M. Salanne, *Chemical Reviews* **122**, 10860 (2022).
- <sup>9</sup>L. Scalfi, T. Dufils, K. G. Reeves, B. Rotenberg, and M. Salanne, *The Journal of Chemical Physics* **153**, 174704 (2020).
- <sup>10</sup>K. Goloviznina, J. Fleischhaker, T. Binninger, B. Rotenberg, H. Ers, V. Ivanistsev, R. Meissner, A. Serva, and M. Salanne, *arXiv*, 204.07848 (2024).
- <sup>11</sup>J. D. Elliott, M. Chiricotto, A. Troisi, and P. Carbone, *Carbon* **207**, 292 (2023).
- <sup>12</sup>S.-J. Shin, J. W. Gittins, M. J. Golomb, A. C. Forse, and A. Walsh, *Journal of the American Chemical Society* **145**, 14529 (2023), pMID: 37341453.
- <sup>13</sup>Z. Wei, J. D. Elliott, A. A. Papaderakis, R. A. Dryfe, and P. Carbone, *Journal of the American Chemical Society* **146**, 760 (2024).
- <sup>14</sup>S.-J. Shin, H. Choi, S. Ringe, D. H. Won, H.-S. Oh, D. H. Kim, T. Lee, D.-H. Nam, H. Kim, and C. H. Choi, *Nature Communications* **13**, 5482 (2022).
- <sup>15</sup>S.-J. Shin, D. H. Kim, G. Bae, S. Ringe, H. Choi, H.-K. Lim, C. H. Choi, and H. Kim, *Nature Communications* **13**, 174 (2022).
- <sup>16</sup>N. Abidi and S. N. Steinmann, *ACS Appl. Mater. Interfaces* **15**, 25009 (2023).
- <sup>17</sup>K. Takahashi, H. Nakano, and H. Sato, *The Journal of Chemical Physics* **157**, 234107 (2022).
- <sup>18</sup>A. Grisafi, A. Bussy, M. Salanne, and R. Vuilleumier, *Phys. Rev. Mater.* **7**, 125403 (2023).
- <sup>19</sup>A. Grisafi and A. Lewis, “Salted: Symmetry-adapted learning of three-dimensional electron densities,” (2023).
- <sup>20</sup>A. Coretti, C. Bacon, R. Berthin, A. Serva, L. Scalfi, I. Chubak, K. Goloviznina, M. Haefele, A. Marin-Lafèche, B. Rotenberg, S. Bonella, and M. Salanne, *J. Chem. Phys.* **157**, 184801 (2022).
- <sup>21</sup>T. Dufils, G. Jeanmairat, B. Rotenberg, M. Sprik, and M. Salanne, *Phys. Rev. Lett.* **123**, 195501 (2019).
- <sup>22</sup>A. Grisafi, A. M. Lewis, M. Rossi, and M. Ceriotti, *Journal of Chemical Theory and Computation* **19**, 4451 (2023).
- <sup>23</sup>A. M. Lewis, A. Grisafi, M. Ceriotti, and M. Rossi, *Journal of Chemical Theory and Computation* **17**, 7203 (2021).
- <sup>24</sup>T. D. Khne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schtt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Bortnik, M. Taillefumier, A. S. Jakobovits, A. Lazaro, H. Pabst, T. Mller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Gl, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, and J. Hutter, *The Journal of Chemical Physics* **152**, 194103 (2020).
- <sup>25</sup>J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- <sup>26</sup>J. VandeVondele and J. Hutter, *The Journal of Chemical Physics* **127**, 114105 (2007).
- <sup>27</sup>S. Goedecker, M. Teter, and J. Hutter, *Phys. Rev. B* **54**, 1703 (1996).
- <sup>28</sup>M. Guidon, J. Hutter, and J. VandeVondele, *Journal of Chemical Theory and Computation* **5**, 3010 (2009), pMID: 26609981.
- <sup>29</sup>A. Bussy, O. Schtt, and J. Hutter, *The Journal of Chemical Physics* **158**, 164109 (2023).
- <sup>30</sup>K. R. Briling, A. Fabrizio, and C. Corminboeuf, *The Journal of Chemical Physics* **155**, 024107 (2021).
- <sup>31</sup>A. Grisafi and M. Ceriotti, *J. Chem. Phys.* **151**, 204105 (2019).
- <sup>32</sup>A. Grisafi, J. Nigam, and M. Ceriotti, *Chem. Sci.* **12**, 2078 (2021).
- <sup>33</sup>K. K. Huguenin-Dumittan, P. Loche, N. Haoran, and M. Ceriotti, *The Journal of Physical Chemistry Letters* **14**, 9612 (2023).
- <sup>34</sup>G. Fraux, P. Loche, S. Kliavinek, K. K. Huguenin-Dumittan, D. Tisi, and A. Goscinski, “rascaline,” .
- <sup>35</sup>A. Marin-Lafèche, M. Haefele, L. Scalfi, A. Coretti, T. Dufils, G. Jeanmairat, S. K. Reed, A. Serva, R. Berthin, C. Bacon, S. Bonella, B. Rotenberg, P. A. Madden, and M. Salanne, *Journal of Open Source Software* **5**, 2373 (2020).
- <sup>36</sup>D. Marx and J. Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods* (Cambridge University Press, 2009).
- <sup>37</sup>J. L. F. Abascal and C. Vega, *The Journal of Chemical Physics* **123**, 234505 (2005).
- <sup>38</sup>T. Yagasaki, M. Matsumoto, and H. Tanaka, *Journal of Chemical Theory and Computation* **16**, 2460 (2020).
- <sup>39</sup>G. L. Stoychev, A. A. Auer, and F. Neese, *Journal of Chemical Theory and Computation* **13**, 554 (2017).
- <sup>40</sup>S. Nosé, *J. Chem. Phys.* **81**, 511 (1984).
- <sup>41</sup>A. Serva, L. Scalfi, B. Rotenberg, and M. Salanne, *The Journal of Chemical Physics* **155**, 044703 (2021).
- <sup>42</sup>L. Scalfi, D. T. Limmer, A. Coretti, S. Bonella, P. A. Madden, M. Salanne, and B. Rotenberg, *Phys. Chem. Chem. Phys.* **22**, 10480 (2020).
- <sup>43</sup>W. Schmickler, *Chemical Reviews* **96**, 3177 (1996).
- <sup>44</sup>T. W. Ko, J. A. Finkler, S. Goedecker, and J. Behler, *Nature Communications* **12**, 398 (2021).
- <sup>45</sup>C. G. Staacke, S. Wengert, C. Kunkel, G. Csányi, K. Reuter, and J. T. Margraf, *Machine Learning: Science and Technology* **3**, 015032 (2022).
- <sup>46</sup>Y. Shao, L. Andersson, L. Knijff, and C. Zhang, *Electronic Structure* **4**, 014012 (2022).
- <sup>47</sup>T. Dufils, L. Knijff, Y. Shao, and C. Zhang, *Journal of Chemical Theory and Computation* **19**, 5199 (2023).
- <sup>48</sup>S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, *Nature Communications* **13**, 2453 (2022).
- <sup>49</sup>S. Klawohn, J. P. Darby, J. R. Kermode, G. Csnyi, M. A. Caro, and A. P. Bartk, *The Journal of Chemical Physics* **159**, 174108 (2023).
- <sup>50</sup>C. Choi, D. S. Ashby, D. M. Butts, R. H. DeBlock, Q. Wei, J. Lau, and B. Dunn, *Nature Reviews Materials* **5**, 5 (2020).
- <sup>51</sup>S. Ringe, E. L. Clark, J. Resasco, A. Walton, B. Seger, A. T. Bell, and K. Chan, *Energy Environ. Sci.* **12**, 3001 (2019).
- <sup>52</sup><https://github.com/cp2k/cp2k>.
- <sup>53</sup><https://github.com/andreagrisafi/SALTED>.
- <sup>54</sup>[https://gitlab.com/andreagrisafi/metalwalls/-/tree/salted\\_interface?ref\\_type=heads](https://gitlab.com/andreagrisafi/metalwalls/-/tree/salted_interface?ref_type=heads).
- <sup>55</sup><https://zenodo.org/doi/10.5281/zenodo.11175494>.