

# Static and dynamic water structures at interfaces: A case study with focus on Pt(111)

Cite as: J. Chem. Phys. **155**, 194702 (2021); <https://doi.org/10.1063/5.0067106>

Submitted: 13 August 2021 • Accepted: 22 October 2021 • Published Online: 18 November 2021

Alexandra C. Dávila López,  Thorben Eggert,  Karsten Reuter, et al.

## COLLECTIONS

Paper published as part of the special topic on [The Chemical Physics of the Electrode-Electrolyte Interface](#)



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

[Structure and dynamics of aqueous NaCl solutions at high temperatures and pressures](#)

The Journal of Chemical Physics **155**, 194506 (2021); <https://doi.org/10.1063/5.0067166>

[Is the water/Pt\(111\) interface ordered at room temperature?](#)

The Journal of Chemical Physics **155**, 224701 (2021); <https://doi.org/10.1063/5.0077580>

[Ab initio study of nuclear quantum effects on sub- and supercritical water](#)

The Journal of Chemical Physics **155**, 194107 (2021); <https://doi.org/10.1063/5.0071857>

The Journal  
of Chemical Physics

**SPECIAL TOPIC:** Low-Dimensional  
Materials for Quantum Information Science

Submit Today!



# Static and dynamic water structures at interfaces: A case study with focus on Pt(111)

Cite as: J. Chem. Phys. 155, 194702 (2021); doi: 10.1063/5.0067106

Submitted: 13 August 2021 • Accepted: 22 October 2021 •

Published Online: 18 November 2021



View Online



Export Citation



CrossMark

Alexandra C. Dávila López,<sup>1</sup> Thorben Eggert,<sup>1,2</sup>  Karsten Reuter,<sup>1</sup>  and Nicolas G. Hörmann<sup>1,2a)</sup> 

## AFFILIATIONS

<sup>1</sup>Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

<sup>2</sup>Chair of Theoretical Chemistry and Catalysis Research Center, Technische Universität München, Lichtenbergstr. 4, 85748 Garching, Germany

**Note:** This paper is part of the JCP Special Topic on The Chemical Physics of the Electrode–Electrolyte Interface.

<sup>a)</sup>Author to whom correspondence should be addressed: [hoermann@fhi-berlin.mpg.de](mailto:hoermann@fhi-berlin.mpg.de)

## ABSTRACT

An accurate atomistic treatment of aqueous solid–liquid interfaces necessitates the explicit description of interfacial water ideally via *ab initio* molecular dynamics simulations. Many applications, however, still rely on static interfacial water models, e.g., for the computation of (electro)chemical reaction barriers and focus on a single, prototypical structure. In this work, we systematically study the relation between density functional theory-derived static and dynamic interfacial water models with specific focus on the water–Pt(111) interface. We first introduce a general construction protocol for static 2D water layers on any substrate, which we apply to the low index surfaces of Pt. Subsequently, we compare these with structures from a broad selection of reference works based on the Smooth Overlap of Atomic Positions descriptor. The analysis reveals some structural overlap between static and dynamic water ensembles; however, static structures tend to overemphasize the in-plane hydrogen bonding network. This feature is especially pronounced for the widely used low-temperature hexagonal ice-like structure. In addition, a complex relation between structure, work function, and adsorption energy is observed, which suggests that the concentration on single, static water models might introduce systematic biases that are likely reduced by averaging over consistently created structural ensembles, as introduced here.

© 2021 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0067106>

## I. INTRODUCTION

Solid–liquid interfaces (SLIs) are ubiquitous in nature, and their study is highly relevant for understanding the (electro)chemical transformation processes that lead to natural corrosion or (electro)catalytic applications, such as in batteries, electrolyzers, or fuel cells.<sup>1–6</sup> While liquid properties at distances larger than  $\sim 1$  nm from the interface are already bulk-like<sup>7,8</sup> and approximately described with standard continuum models,<sup>9,10</sup> the structure and composition of the first (few) solvent layers in contact with a solid surface typically show a significant dependence on, e.g., the solid substrate and the thermodynamic conditions, e.g., applied electrode potentials in electrochemistry contexts.<sup>11–19</sup> In aqueous solutions, an accurate atomistic treatment of the substrate–water interface necessitates the explicit inclusion of at least the first water layer for a wide range of properties such as adsorption energies or the

potential of zero charge (PZC).<sup>15,20–23</sup> On the other hand, the sensitive dependence of these properties on the interfacial water structure<sup>24–27</sup> necessitates, in principle, an appropriate sampling of these to obtain reliable thermodynamic averages.<sup>11,21,26,28,29</sup> Indeed, as the accurate description of electronic degrees of freedom and the chemical reactivity of the substrate are important, *ab initio* molecular dynamics (AIMD) results based on density functional theory (DFT) are the only reliable benchmarks to date. The long relaxation times of interfacial water<sup>30–32</sup> pose, however, a significant challenge to AIMD simulations and restrict such studies to only a few selected model systems.<sup>14–18,21,26,33–36</sup> As a result, many studies, e.g., on adsorption and solvation energies across different substrates and adsorbates<sup>27,37–41</sup> or on electrochemical reaction barriers<sup>42–49</sup> still rely on simplified and quasi-static interfacial water models. The water–Pt(111) interface is most likely the best studied of such systems and has been addressed by AIMD simulations<sup>13,14,21,29,50,51</sup>

and static low-temperature water models.<sup>25,50,52–54</sup> Typical static interfacial water models are comprised of ice-like, hexagonal interfacial layers,<sup>20,48–50,52,54</sup>  $H^{\text{down}}$ ,  $H^{\text{up}}$ , and chain- $H^{\text{down}}$ , and also involve more complex phases<sup>52–55</sup> such as  $\sqrt{37} \times \sqrt{37}R25.3^\circ$  and  $\sqrt{39} \times \sqrt{39}R16.1^\circ$ . The latter structure comprises five, six, and seven water rings.<sup>52–54</sup> Such structures are typically less prominent in room temperature AIMD studies<sup>16,50,56</sup> where only (predominant) *local* water configurations can be identified.<sup>16</sup> This puts forward the question about the relation between theoretical results from static and dynamically sampled extended interfacial water structures and possible structural biases due to the dominant use of hexagonal water (bilayer) arrangements in applied works.<sup>42,48,49</sup> Furthermore, as no “standard” static water models seem available for other than the prototypical fcc-(111) surface, there is a dramatic lack of knowledge of interfacial water structures on other interfaces and substrates.

In order to fill this void, we systematically study and compare ensembles of static and dynamic water structures. We introduce first our construction protocol for obtaining a consistent set of static water structures at solid interfaces for any given surface termination, supercell size, and crystal structure. Here, we start by creating a dataset of selected, topologically different 2D water layers in vacuum. The created 2D water layers can be adsorbed on any given substrate, leveraging optimal lattice matching algorithms and consecutive geometric relaxation via DFT. Subsequently, we apply the algorithm to construct a dataset of 2D static water (2DSW) structures on low index Pt surfaces and compare the obtained structural ensemble of interfacial water with the structure of water layers from other reference works. In particular, we investigate structural similarities within quasi-2D water layers in bulk liquid water and ice, at (liquid) water–vacuum interfaces, and at Pt(111)–water interfaces using the local structure descriptor SOAP (Smooth Overlap of Atomic Positions).<sup>57</sup> One central aspect in this respect is the analysis of the relations between structure, adsorption energy, and work function reduction for Pt(111) with a single adsorbed water layer and as obtained from AIMD and our 2DSW ensemble. In general, the analysis reveals a bias of 2DSW structures toward three intermolecular hydrogen bonds per water and excess hydrogen bonds pointing to the substrate, which yield a maximization of the overall number of bonds within the interface in contrast to the AIMD configurations (see Secs. III B and III C). Nevertheless, we find a decent overlap between AIMD and 2DSW ensemble averages, while single selected static models exhibit large variations in their predictive quality.

## II. CONSTRUCTION PROTOCOL FOR WATER LAYERS ON SUBSTRATES

Our starting point for the proposed protocol is the common understanding that the first water layer is extremely important for a variety of properties at metal–water interfaces and that  $H_2O$ – $H_2O$  interactions dominate the total energy, at least on coinage metal low index surfaces.<sup>8,58,59</sup> An intuitive approach to model adsorbed water is by using water layers in vacuum as a starting guess, which directly motivates the following procedure:

- Creation of a dataset of 2D water phases in vacuum.
- Creation of the target substrate [(*hkl*) surface and supercell size].

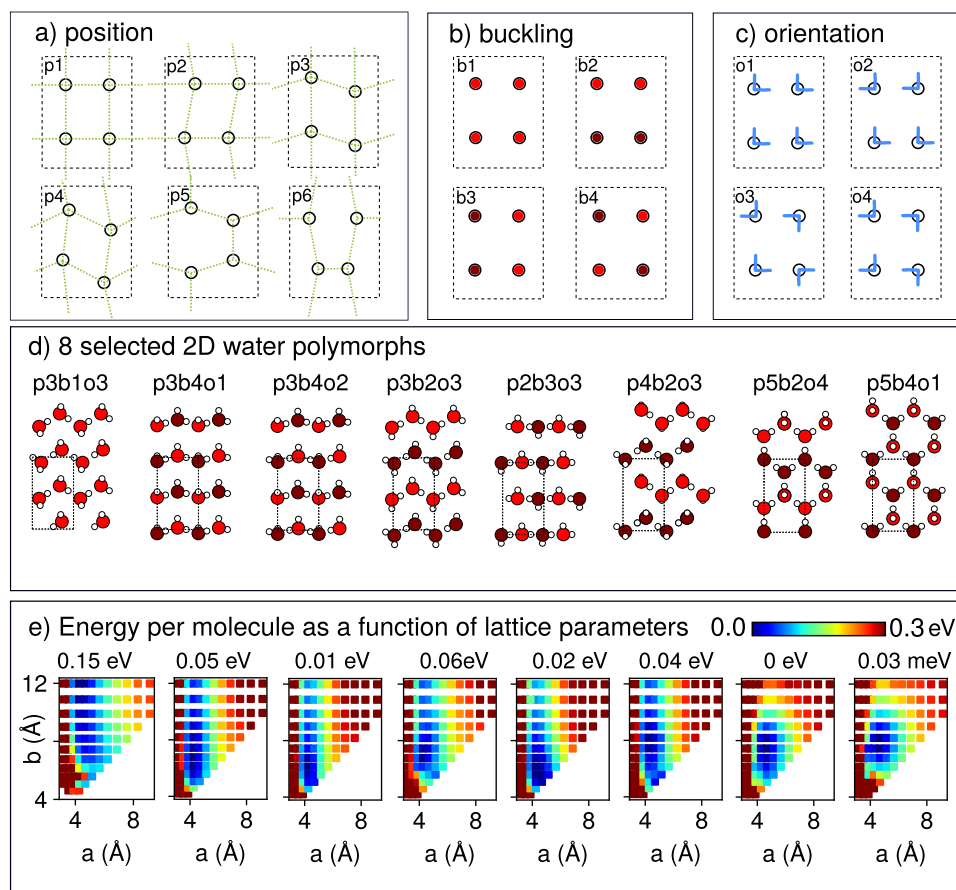
- Adsorption of 2D water phases that match the substrate supercell geometry.
- Relaxation of the water adsorbate layer.

### A. 2D water structures in vacuum

Based on DFT calculations, we initially investigated 15 different polymorphs of 2D water in vacuum consisting of 1–8 water molecules in the primitive cell as reported in the literature.<sup>60–63</sup> We varied lattice constants and unit cell shapes in order to get an overview of the typical 2D water orderings in vacuum. 2D water is highly flexible, and typical literature studies are in the context of confined water, e.g., in between two sheets of graphene, rationalizing size or pressure restrictions in the out-of-plane direction.<sup>61</sup> As these are not present here, a compression mainly leads to a change in the degree of buckling and stacking of water molecules into several layers, an expansion typically to the formation of 1D water chains with varying degrees of separation. In general, energetic differences per molecule between the considered systems with different numbers of water are small, indicating that complex, large-scale water patterns are not necessary to describe the approximate energetic landscape. Furthermore, a wide variety of observed water patterns can already be achieved with only four water molecules<sup>62</sup> in a rectangular primitive lattice, featuring water molecules arranged in rectangles, parallelograms, and hexagons (see Fig. 1). Note that known prototypical motives such as squares, rectangles, parallelograms, and hexagons can be understood within rectangular cells by a variation in the relative position of the molecules within the cell, as shown in Fig. 1(a). Furthermore, all observed buckling patterns (water molecules not in one plane) fall in four classes, where the water network is in a plane, exhibits horizontal or vertical chains, or forms a 2D checkerboard arrangement [see Fig. 1(b)]. In addition, the relative orientation of the water molecules can follow several patterns; the most symmetric (in-plane) representations where each water molecule can maximize the number of hydrogen bonds are drawn in Fig. 1(c).

After these observations that a wide variety of different topologies are, indeed, already possible for four molecules in rectangular cells, we decided to limit ourselves to polymorphs, which obey these restrictions, as they have several advantages: On the one hand, the relatively small water unit cells allow for a more exhaustive representation of the structural landscape due to the reduced number of degrees of freedom. On the other hand, lattice matching can be achieved by smaller increments in the (water) supercell size. As a result, it is more likely to find small substrate supercells, which is beneficial, e.g., in terms of applying the protocol to many different systems. In addition, the choice of equal lattices for all water polymorphs ensures the consistent inclusion of all the considered water topologies on any substrate.

Our final set of 2D water structures consists of eight different topologies [Fig. 1(d)] that are observed in the studied selection of 15 low-temperature polymorphs and include a large variety of position (p), buckling (b), and orientation (o) patterns. We furthermore scanned the lattice constants on a logarithmically spaced grid (99 grid points), for which we report the relative stability with respect to the optimum lattice constant in Fig. 1(e). This clarifies the small energetic differences of any of these topological variants. The first six topologies are chains in the direction of the lattice



**FIG. 1.** (a)–(c) Basic motives of high symmetry water structures with four  $\text{H}_2\text{O}$  molecules in a rectangular cell can be discriminated by the following: (a) Different high symmetry oxygen positions in the unit cell, which yields water arrangements in rectangles, squares, parallelograms (and mixed versions), and hexagons. (b) Out-of-plane buckling patterns. Dark red is used for oxygen atoms in a lower plane. (c) Relative water orientations (fulfilling water rules). (d) Our selected set of 2D water polymorphs, named according to the topological descriptors in (a)–(c). (e) Relative energies per water molecule for different lateral dimensions scanned on a grid. Within the eight selected polymorphs, p5b2o4/p3b1o3 yields the most/least stable water structure (cf. minimum energies reported above the energy landscapes).

vector  $a$  such that their energy is highly sensitive to changes in that direction and largely invariant for changes in the direction of the lattice vector  $b$ . In contrast, the energy of the latter two configurations with their hexagonal geometry is sensitive to changes in both directions. A logarithmic lattice spacing is chosen as it leads to identical relative offsets of neighboring lattice vectors (here 10%), which ensures lattice matching (see below) without overlaps. Finally, we reduced this set to 20 identical lattice constant combinations for all eight polymorphs, which cover the lowest energy regions in order to construct our final dataset of 160 2D static water polymorphs.

## B. Choice of substrate supercells

Choosing an appropriate substrate supercell evidently depends on the quantities and systems of interest. As surface cells with maximum distances to periodic images in the lateral directions are expected to yield the least periodic boundary artifacts, we think all studies on surfaces should ideally be performed in such

*maximally isotropic supercells*. For this purpose, we implemented an algorithm that yields (with minimal user intervention) slabs in vacuum, which are *maximally isotropic* in the in-plane directions for any given substrate material,  $(hkl)$  surface, and target surface area. The algorithm is outlined in the [supplementary material](#), and a Python implementation is provided alongside this work.

## C. 2D lattice matching of 2D water phases

With our given dataset of 2D water prototypes at the varying lattice constant and chosen substrate supercell, we use (a slightly modified version of) pymatgen<sup>64–66</sup> to adsorb 2D water on the substrate via leveraging its implemented lattice matching<sup>67</sup> functionality similarly as implemented in the pymatgen.analysis.interface module. The logarithmic lattice spacing in our dataset ensures consistency with the lattice vector mismatch constraints in pymatgen (relative length differences) and enables non-overlapping matches for neighboring lattice constants.

In addition, we include possible lateral shifts of the adsorbed water layer (scanned on a regular grid within the primitive surface unit cell and a lateral grid spacing of  $\sim 2$  Å) while also taking both possible water layer orientations into account in case the water prototype is polar in the out-of-plane direction. A self-contained Python implementation of the outlined 2DSW construction protocol is provided alongside this publication.

#### D. The 2DSW protocol applied to low-index Pt surfaces

Using the outlined approach, we created lattice-matched 2DSW water structures on *maximally isotropic substrate supercells* for supercell sizes between 6 and 36 surface atoms for the low index surfaces (100), (110), and (111) of fcc-Pt. While a wide variety of possible combinations for the cell size and number of water molecules is possible, interfaces with 8 and 16 water molecules take a special role as they yield matches for all three (*hkl*) terminations. Here, we restrict our analysis to minimal supercell sizes with only eight water molecules. For the Pt(111) termination, which we will analyze in more detail subsequently, the algorithm automatically chooses the well-known ( $\sqrt{12} \times \sqrt{12}$ )R30° supercell and yields 80 different initial configurations for interfacial water. These are relaxed via DFT, and we compute the adsorption energetics, as well as geometric and electronic properties (eight initial structures ran into convergence problems, which were not further analyzed and discarded). The computational parameters are reported in Subsection 2 of the Appendix, and the results for Pt(100) and Pt(110) surfaces are shown in the [supplementary material](#).

### III. PROPERTIES OF WATER LAYERS

#### A. Methodology

In order to analyze the water ordering of quasi-2D water layers, we used a variety of sources, including *ab initio*-determined bulk ice<sup>68</sup> and bulk liquid water structures,<sup>72</sup> liquid water slabs in vacuum,<sup>72</sup> liquid water on Pt(111),<sup>15,21</sup> and static ice-like-layers on Pt(111) from Ref. 54 and from our 2DSW construction protocol. Table I lists the according datasets together with our chosen naming convention and the most important associated properties. Note that

**TABLE I.** Overview of the datasets used in this paper and the correspondingly used exchange–correlation functionals (plus vdW correction). The column “Setup” discriminates between the properties of the original dataset—bulk and symmetric/asymmetric slab calculations. All datasets are publicly available (see original publications), apart from the AIMD(P/R) trajectories, which were provided by the (co-)authors of the respective publications, Le *et al.*<sup>21</sup> and Heenen *et al.*<sup>15</sup>

| Name                             | Reference | Setup | Functional(-vdW)                          |
|----------------------------------|-----------|-------|---|
| H <sub>2</sub> O ice             | 68        | Bulk  | revPBE <sup>69,70</sup> -D3 <sup>71</sup> |
| H <sub>2</sub> O AIMD            | 72        | Bulk  | rVV10 <sup>73</sup>                       |
| H <sub>2</sub> O/vacuum AIMD     | 72        | Sym.  | rVV10 <sup>73</sup>                       |
| H <sub>2</sub> O@Pt(111) AIMD(P) | 21        | Sym.  | PBE <sup>74</sup> -D3 <sup>75,76</sup>    |
| H <sub>2</sub> O@Pt(111) AIMD(R) | 15        | Asym. | RPBE <sup>77</sup> -D3 <sup>75,76</sup>   |
| H <sub>2</sub> O@Pt(111) ice     | 54        | Asym. | optPBE-vdW <sup>78</sup>                  |
| H <sub>2</sub> O@Pt 2DSW         | This work | Asym. | PBE <sup>74</sup>                         |

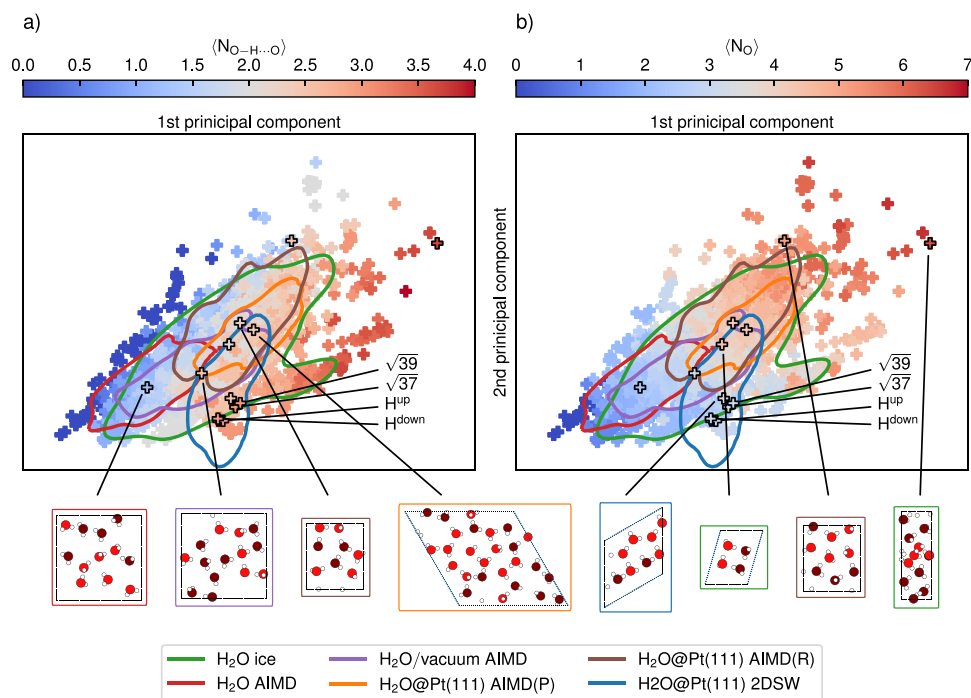
only the latter two sets of structures [H<sub>2</sub>O@Pt(111) ice and H<sub>2</sub>O@Pt 2DSW] include natively well-defined quasi-2D water layers, while all other sources also incorporate in parts or in full bulk-like water regions. However, water layers are also evidently present in 3D bulk water structures (H<sub>2</sub>O ice, H<sub>2</sub>O AIMD) and are, in particular, distinct and non-bulk-like for water–vacuum (H<sub>2</sub>O/vacuum AIMD) or water–Pt interfaces [H<sub>2</sub>O@Pt(111) AIMD(P), H<sub>2</sub>O@Pt(111) AIMD(R)].

In order to study the structure within these layers, we create from the latter datasets first systems with water–vacuum boundaries, if needed, and subsequently extract water layers at water–vacuum interfaces. In particular, we selected snapshots from the bulk H<sub>2</sub>O AIMD trajectory and introduced vacuum layers at random positions along the *z* direction while enforcing the integrity of H<sub>2</sub>O molecules. Similarly, for H<sub>2</sub>O ice (54 ice phases from Ref. 68), we created water slabs up to a maximum Miller index of 3 using `pymatgen`<sup>64–66</sup> again while enforcing molecular integrity. For the water–Pt AIMD trajectories [H<sub>2</sub>O@Pt(111) AIMD(P), H<sub>2</sub>O@Pt(111) AIMD(R)], we removed all Pt-atoms from the system, thus creating a vacuum void. Interfacial water layers were then extracted from the so-preprocessed structures by determining all interfacial water molecules only for the respective vacuum–water interfaces of interest (e.g., the original H<sub>2</sub>O@Pt interface) via the algorithm described in Subsection 1 of the Appendix.

The so-created structural datasets for water layers in bulk, at vacuum–water, and at Pt–water interfaces are analyzed via the use of an abstract representation for atomic environments, in particular, based on the local structure descriptor SOAP<sup>57</sup> as implemented in the `Dscribe` package<sup>79</sup> and as similarly performed for 3D periodic water structures.<sup>68</sup> Local environments are represented for each water molecule by computing a SOAP vector centered on the oxygen atom and using a cutoff radius of 4 Å, consistent with the geometrical characteristics of water. Interfacial water structures are characterized by the average over all SOAP vectors taken from each O atom in a configuration,<sup>79</sup> which thus also allows (global) for a structure comparison across structures that differ in the number of H<sub>2</sub>O molecules (cf. also Subsection 3 of the Appendix).

#### B. The structure of extended water layers

In order to gain an intelligible, e.g., visual, representation of the similarity between and within the different structural ensembles of extended water layers, we apply principal component analysis (PCA) to the complete dataset of average SOAP vectors. The first and second PCA components are depicted in Fig. 2, and the solid lines indicate approximate locations for each dataset as obtained via kernel density estimation.<sup>80</sup> Additional dataset-specific plots are provided in the [supplementary material](#). Figures 2(a) and 2(b) and the color-coding illustrate how PCA discriminates between the average number of intermolecular hydrogen bonds (upper left to the lower right) and the average number of oxygens within the SOAP radius (lower left to the upper right). Hydrogen bonds between two water molecules are defined as in Ref. 16, where the oxygens are closer than 3.5 Å to each other and the angle between O–O–H is less than 35°. There is a broad distribution of structures in the analyzed data, ranging from rather isolated water molecules to highly coordinated ones with four hydrogen bonds to its neighbors and the bulk-ice-derived layers (green line) covering the whole range, with selected



**FIG. 2.** PCA map for the structure within water layers based on the average SOAP vector (one point corresponds to one structure; cf. Subsection 3 of the Appendix). The datasets are those of Table I, and the color-codings in (a) and (b) correspond to the average number of hydrogen bonds  $\langle N_{O-H...O} \rangle$  as defined in Ref. 16 and the number of nearby oxygen atoms  $\langle N_O \rangle$  within a cutoff radius of 4 Å, respectively. The contour lines illustrate the locations of each dataset as obtained via kernel density estimation. For each dataset, we show an exemplary structure as indicated by the frame color.

examples plotted in Fig. 2. Bulk liquid water (H<sub>2</sub>O AIMD, red) is mainly located in the low density region at the lower left corner, while water at water–vacuum interfaces (H<sub>2</sub>O/vacuum AIMD, violet) and [H<sub>2</sub>O@Pt(111) AIMD(P/R), orange/brown] move upward along the diagonal, toward higher density structures. Note that the increase in density (number of nearby oxygens  $\langle N_O \rangle$ ) does not dramatically alter the number of H-bonds  $\langle N_{O-H...O} \rangle$  within the water layer, as can be rationalized from the color codings in Figs. 2(a) and 2(b) and from the reported average values provided in Table II.

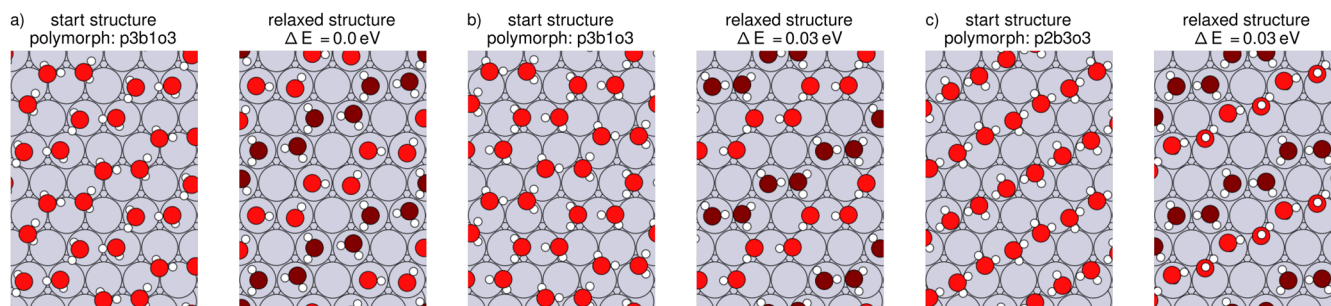
**TABLE II.** Average number of hydrogen bonds  $\langle N_{O-H...O} \rangle$  and number of nearby oxygens  $\langle N_O \rangle$  within the cutoff radius of 4 Å per H<sub>2</sub>O molecule for each dataset and respective standard deviations in brackets.

| Dataset  | $\langle N_{O-H...O} \rangle$ | $\langle N_O \rangle$ |
|--|-------------------------------|-----------------------|
| H <sub>2</sub> O ice <sup>68</sup>             | 1.86(0.88)                    | 3.41(1.38)            |
| H <sub>2</sub> O AIMD <sup>72</sup>            | 1.29(0.45)                    | 2.30(0.51)            |
| H <sub>2</sub> O/vacuum AIMD <sup>72</sup>     | 2.20(0.38)                    | 3.14(0.46)            |
| H <sub>2</sub> O@Pt(111) AIMD(P) <sup>21</sup> | 2.21(0.21)                    | 4.17(0.41)            |
| H <sub>2</sub> O@Pt(111) AIMD(R) <sup>15</sup> | 2.19(0.33)                    | 4.17(0.62)            |
| H <sub>2</sub> O@Pt(111) ice <sup>54</sup>     | 3.00(0.00)                    | 3.00(0.03)            |
| H <sub>2</sub> O@Pt(111) 2DSW                  | 2.56(0.45)                    | 3.07(0.47)            |

For comparison, we also marked in Fig. 2 specifically the prototypical low-temperature structures<sup>54</sup>—(hexagonal) H<sup>up/down</sup> and  $\sqrt{37}/\sqrt{39}$ —which fall in the region of our static water dataset H<sub>2</sub>O@Pt 2DSW, which includes natively hexagonal water layers. It is worth mentioning that the widely used H<sup>up/down</sup> structures are observed at the extreme boundaries of the configurational space, whereby on average oxygen atoms in H<sup>up/down</sup> structures form 0.8 more hydrogen bonds than in H<sub>2</sub>O@Pt(111) AIMD, as shown in Table II. While there is no overlap between these and the AIMD sampled interface structures on Pt(111) (orange and brown lines), our 2DSW protocol yields at least some overlap. In particular, structures with less hydrogen bonds, which mainly consist of chain-like water arrangements, fall in the region of AIMD results (see also discussion below). These results indicate that ensemble averages based on 2DSW static water structures might approximate AIMD averages more accurately than using only selected low-temperature (e.g., hexagonal) structural models. Interestingly, Le *et al.* pointed out that beyond the first solvation layer, the hydrogen bonding is mainly affected by the water–water interaction rather than the water–metal interaction.<sup>16</sup> This suggests that better results might be achieved by including more water layers in the system.

### C. Water layers on Pt(111)

While all 2DSW water structures are provided alongside this work, we show a selection of three obtained structures in Fig. 3,



**FIG. 3.** Selected water–Pt(111) configurations initialized from the p3b1o3 and p2b3o3 polymorphs and their corresponding relaxed structures. (a) The most stable relaxed structure of the whole 2DSW dataset. [(b) and (c)] Two slightly less stable configurations ( $\Delta E = 0.03$  eV/ $H_2O$ ). Oxygen atoms corresponding to flat  $H_2O$  molecules are colored by dark red in the relaxed structures.

namely, the lowest energy configuration after relaxation (a) and two local minima [(b) and (c)].

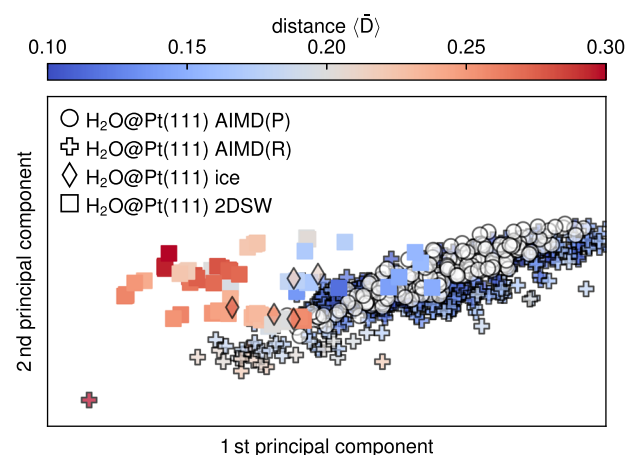
Interestingly, the starting structure of the least stable polymorph p3b1o3 [Fig. 3(a)] leads to the most stable relaxed structure in our set, which is characterized by hexagonal  $H_2O$  rings, in which the hydrogen network is distinguished by two chains: one constituted of only  $H^{\text{down}}$  oriented water molecules and the other chain comprises  $H_2O$  molecules parallel to the surface. Here, the structure is stabilized by forming the three hydrogen bonds per oxygen atom as in other ice-like water layers. Furthermore, this structure was also reported by Clabaut *et al.*<sup>54</sup> as the most stable among the hexagonal ice adlayers.

The configuration in Fig. 3(b) is also initialized from the p3b1o3 polymorph, but the start configurations differ from each other in the atomic positions of the water molecules, while Fig. 3(c) started from chain-like water structures. Both relaxed structures in Figs. 3(b) and 3(c) form chains of hexagonal  $H_2O$  rings, however, with differing water orientations. Regarding energy, both structures are only 0.03 eV/ $H_2O$  less stable relative to the lowest energy configuration [Fig. 3(a)]. Similar chain-structures have been reported for (110) metal surfaces.<sup>53</sup>

In general, our 2DSW protocol yields a variety of hexagonal ring structures comprised of  $H^{\text{down}}$ ,  $H^{\text{up}}$  oriented water molecules and water molecules parallel to the surface, similarly as shown in Fig. 1(a) at 11 ps in Ref. 51, supporting their importance for simulations in small unit cells.

For the analysis of water layers on the substrate Pt(111), we follow the previous approach and leverage the average SOAP vector as the structural descriptor, this time, however, including specifically the Pt substrate atoms in the atomic environments (see Subsection 3 of the Appendix). Furthermore, we use a SOAP-based average distance measure ( $\bar{D}$ ) to evaluate the structural similarity of a single structure with the AIMD(P) ensemble of interfacial water structures, which we took as our reference (cf. Subsection 3 of the Appendix). The PCA map in Fig. 4 yields similar results as before and illustrates how ( $\bar{D}$ ) (cf. color) is able to measure structural similarity to the reference ensemble. As expected, both AIMD simulations remain close to each other. For the  $H_2O@Pt(111)$  ice layers (diamonds), we find a clear trend in distances  $\bar{D}$ , namely,  $H^{\text{up}}$  (0.27) >  $H^{\text{down}}$  ≈ chain- $H^{\text{down}}$  >  $\sqrt{37}$  >  $\sqrt{39}$  (0.21), where the rather large dissimilarity of the

hexagonal structures is likely ascribed to the lack of disorder. Similarly as before, our 2DSW ensemble shows better agreement with the AIMD results, although some large dissimilarities ( $\bar{D}$ ) are observed. Further analysis is reported in the supplementary material, where we also analyze other commonly used descriptors of (interfacial) water, such as the atomic density distribution as a function of the distance to the surface and the radial distribution functions within the water adlayers. In general, we find good qualitative agreement in the vertical density profile for the 2DSW ensemble, which is able to reproduce the bimodal oxygen peak structure at the approximately right vertical distance  $z_{\text{Pt-O}}$  between slab and bottom-most oxygen atoms {2.2 Å (2DSW) vs 2.0 Å [AIMD(P)]}. However, some differences are observed in the oxygen density profile, e.g., the first peak for AIMD data is wider than the peak for 2DSW, as shown in Fig. S4 of the supplementary material. This can be ascribed to the absence of the dynamic interchange of interfacial water molecules within the solvation layer and with the bulk liquid water.



**FIG. 4.** PCA map for interfacial water structures on Pt(111) with the atomic positions of the substrate Pt atoms included in the SOAP descriptors. Datasets from other sources are highlighted by black edges, and the colors indicate the average SOAP distance ( $\bar{D}$ ) to the AIMD(P) reference dataset, which consists of the colorless circles. (For more details, see the text and Subsection 3 of the Appendix.)

To this point, we have compared structural similarity to the AIMD(P) reference trajectory; however, we have not yet analyzed how far this influences derived quantities such as average adsorption energies and work function changes, which show a strong structural sensitivity, e.g., to the distance between the metal surface and the water adlayer, as pointed out by Tripkovic *et al.*<sup>25</sup> Similarly, AIMD simulations are comprised of a variety of interfacial water structures due to fluctuating water orientations and ad- and desorption events, which strongly influences the work function due to interfacial charge transfer from the first water layer.<sup>17,18,36</sup> In this regard, Surendralal *et al.* reported values of absolute potentials ranging from 2.5 to roughly 7.5 eV for H<sub>2</sub>O@Pt(111) AIMD simulations, which result in an average of 4.86 eV (see the supplementary material in Ref. 81), and it is of interest to see how far 2DSW ensemble averages relate to such AIMD-based reference results. Here, we evaluate the adsorption energy per water molecule as

$$E_{\text{ads}} = \frac{1}{N_{\text{H}_2\text{O}}} \left[ E_{\text{H}_2\text{O-Pt}}^{\text{DFT}} - \left( E_{\text{Pt}}^{\text{DFT}} + N_{\text{H}_2\text{O}} E_{\text{H}_2\text{O}}^{\text{DFT}} \right) \right], \quad (1)$$

where  $N_{\text{H}_2\text{O}}$  is the number of water molecules at the interface and  $E_{\text{H}_2\text{O-Pt}}^{\text{DFT}}$ ,  $E_{\text{Pt}}^{\text{DFT}}$ , and  $E_{\text{H}_2\text{O}}^{\text{DFT}}$  are the total energy of the surface with the adsorbed water layer, the energy of a clean surface, and the energy of an isolated water molecule, respectively. The work function change  $\Delta\Phi$  due to the adsorption of a water film is determined by the difference in the work function between water-covered and clean slab  $\Delta\Phi = \Phi_{\text{H}_2\text{O-Pt}} - \Phi_{\text{Pt}}$ , both of which are readily accessible as the position of the Fermi level relative to the electrostatic potential in the vacuum region. We computed the respective quantities for our 2DSW dataset but reanalyzed in the same way the AIMD(P) trajectory with all waters removed except for the first interfacial water layer. Gratifyingly, the so-computed work function reduction  $\langle\Delta\Phi\rangle$  for the AIMD(P) trajectory by only the first water layer is identical to the results reported in the original study by Le *et al.*<sup>21</sup> ( $-1.11$  eV here vs  $-1.1$  eV in Ref. 21). This supports once more the major role of only the first water layer.<sup>22,23</sup>

In order to understand better the impact of different possible choices of static interfacial water structures, we studied two specific subsets of our raw 2DSW dataset, which include (i) only configurations with an energy difference  $\Delta E \leq 0.05$  eV/H<sub>2</sub>O relative to the lowest energy configuration and (ii) configurations with  $\langle\bar{D}\rangle < 0.21$ . These two subsets are inspired by two typical assumptions in theoretical works, namely, to focus (i) on low energy structures and (ii) on structures that are close to some accurate reference (e.g., experiments or as here more accurate theoretical

simulations). Table III summarizes the central properties of the considered structural datasets.

Structurally, the low energy subset (i) exhibits the highest observed number of hydrogen bonds  $\langle N_{\text{O-H}\cdots\text{O}} \rangle$ , and the number of nearby oxygens  $\langle N_{\text{O}} \rangle$  is close to 3, which likely derives from the fact that structural optimization leads to a maximized number of hydrogen bonds within the water layer and maximized number of bonds with the metal surface [cf. most stable configuration in Fig. 3(a)]. Therefore, the structures mainly consist of the hexagonal ( $N_{\text{O-H}\cdots\text{O}} = 3$ ) and some two-dimensional chain structures ( $N_{\text{O-H}\cdots\text{O}} = 2.5$ ). In contrast, the latter subset (ii), which is optimized for structural similarity with the AIMD(P) dataset, is characterized by chain structures, as shown in Fig. 3(b), where the average number of hydrogen bonds (2.23) and the average number of nearby oxygen atoms (3.41) are close to the values of the reference set (cf. Table III and Fig. 4).

For all static water ensembles, we find  $E_{\text{ads}}$  values that are significantly lower than the AIMD(P) average, as the latter structures were dynamically sampled with additional bonding partners in the second water layer in the original simulations, which are (artificially) removed in our analysis. Within our dataset,  $E_{\text{ads}}$  of the structurally most similar configuration [see Fig. 5(a)] is roughly 0.1 eV higher than the lowest energy configuration. On the other hand, our average adsorption energies fall close to the reported adsorption energies for static water models.<sup>20,52,58,82</sup> While the numerical differences to the AIMD averages are thus evident, a diverse set of local minima configurations as starting points for further investigations, e.g., based on short AIMD simulations or nudged elastic band calculations, might be beneficial for a better understanding of the effects of local water environments on other processes, such as solvation, diffusion, or electrochemical reactions.

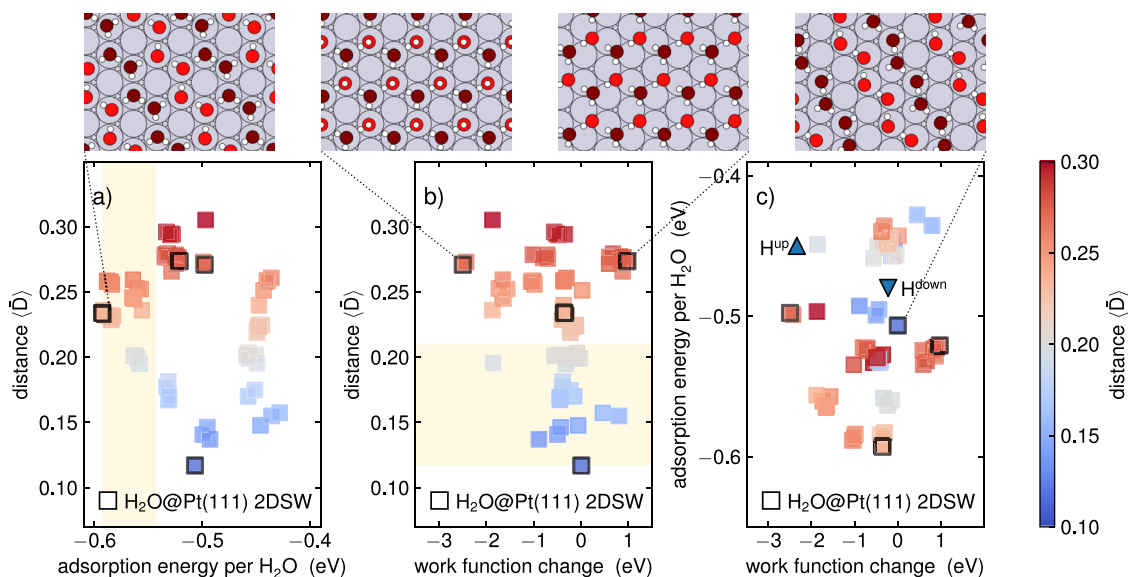
In terms of the work function reduction, the better structural agreement of subset (ii) with the AIMD(P) reference dataset does not lead to better agreement in the work function reduction [ $\langle\Delta\Phi\rangle = -0.43/-0.30$  eV for 2DSW/subset (ii) vs  $-1.11$  eV for AIMD(P)]. Indeed, the low energy subset (i) performs best with  $\langle\Delta\Phi\rangle = -0.83$  eV, which lies within the margin of errors of published AIMD reference results ( $-1.1$  to  $-0.55$  eV<sup>14,21,81</sup>). Note that accurate work functions and work function reductions are relevant for applications in electrochemical contexts as they are directly related to the potential of zero charge of the electrode on an absolute potential scale.<sup>83</sup>

In order to get a better understanding of the correlations between the structural distances  $\langle\bar{D}\rangle$ , the adsorption energy  $E_{\text{ads}}$ ,

**TABLE III.** Central properties of the (reevaluated) H<sub>2</sub>O@Pt(111) AIMD(P) and the 2DSW dataset and of two analyzed subsets: (i) ( $\Delta E < 0.05$  eV/H<sub>2</sub>O with respect to the lowest energy configuration) and (ii) ( $\langle\bar{D}\rangle < 0.21$ ). We report the average number of hydrogen bonds ( $\langle N_{\text{O-H}\cdots\text{O}} \rangle$ ), the number of nearby oxygens ( $\langle N_{\text{O}} \rangle$ ), the adsorption energies ( $E_{\text{ads}}$ ), the work functions ( $\Phi_{\text{Pt}}$ ) of the bare surfaces, and the work function change ( $\Delta\Phi$ ) due to one water adlayer. The standard deviations as obtained from the study of  $N_{\text{conf}}$  configurations are reported in parentheses, as well as the estimated uncertainty (standard error of the mean).

|   | $N_{\text{conf}}$ | $\langle N_{\text{O-H}\cdots\text{O}} \rangle$ | $\langle N_{\text{O}} \rangle$ | $\langle E_{\text{ads}} \rangle$ (eV/H <sub>2</sub> O) | $\langle \Phi_{\text{Pt}} \rangle$ (eV) | $\langle \Delta\Phi \rangle$ (eV) |
|---|-------------------|--|--------------------------------|--|---|-----------------------------------|
| H <sub>2</sub> O@Pt(111) AIMD(P) <sup>21</sup>  | 176               | 2.21(0.21)                                     | 4.17(0.41)                     | $-0.34 \pm 0.00(0.03)$                                 | $5.75 \pm 0.00(0.01)$                   | $-1.11 \pm 0.03(0.33)$            |
| H <sub>2</sub> O@Pt(111) 2DSW   | 72                | 2.56(0.45)                                     | 3.07(0.47)                     | $-0.51 \pm 0.01(0.05)$                                 | $5.77 \pm 0.00(0.01)$                   | $-0.43 \pm 0.09(0.76)$            |
| H <sub>2</sub> O@Pt(111) 2DSW <sub>(i)</sub> [ $\Delta E < 0.05$ eV/H <sub>2</sub> O] | 21                | 2.83(0.24)                                     | 2.93(0.17)                     | $-0.58 \pm 0.00(0.01)$                                 | $5.76 \pm 0.00(0.00)$                   | $-0.84 \pm 0.13(0.60)$            |
| H <sub>2</sub> O@Pt(111) 2DSW <sub>(ii)</sub> [ $\langle\bar{D}\rangle < 0.21$ ]      | 22                | 2.23(0.29)                                     | 3.41(0.44)                     | $-0.49 \pm 0.01(0.05)$                                 | $5.77 \pm 0.00(0.01)$                   | $-0.30 \pm 0.10(0.48)$            |





**FIG. 5.** (a)–(c) The correlations between the average SOAP distance  $\langle \bar{D} \rangle$ , adsorption energy per  $\text{H}_2\text{O}$ , and work function change for the 2DSW dataset reveal a complex structure–property relationship for interfacial water on Pt(111). The beige areas in (a) and (b) correspond to our selection criteria for subsets (i) and (ii) of the 2DSW dataset, respectively. Triangular data points correspond to hexagonal  $\text{H}^{\text{down}}$  and  $\text{H}^{\text{up}}$  ice layers from Ref. 50.

and the work function change  $\Delta\Phi$ , we plot an according analysis of the complete 2DSW dataset in Fig. 5. The beige areas in Figs. 5(a) and 5(b) correspond to our selection criteria for subsets (i) and (ii), respectively. Consistent with the described lower stability of dynamical water structures, we also observe here that the best matching structures (blue) are higher in energy [ $\sim 0.1$  eV, Fig. 5(a)]. At the same time, at these  $E_{\text{ads}}$  values, the distance measure  $\langle \bar{D} \rangle$  seems to exhibit a bimodal distribution, thus also incorporating a significant amount of structures with maximal  $\langle \bar{D} \rangle$  values. The observed work function change  $\Delta\Phi$  exhibits a rather large spread [Fig. 5(b)], which is slightly larger at higher  $\langle \bar{D} \rangle$  values. Indeed, structures with different  $\langle \bar{D} \rangle$  values can still exhibit similar energies and work function change [Fig. 5(c)], indicating a high degree of complexity for the studied structure–property relationships, providing renewed evidence for the difficulty in treating SLIs, in particular water–metal surfaces with simplified, e.g., individual, static interfacial water models. As energies and work functions are not included in the PCA, the prediction can be a challenging task in some cases as here shown.<sup>84</sup>

#### IV. DISCUSSION AND CONCLUSIONS

The systematic study of static interfacial water layers on Pt(111) revealed that prototypical, low energy interfacial water structures exhibit quite special orderings and lie rather at the boundaries of the observed configuration space of AIMD simulations. At variance, our 2DSW ensemble, which includes a wide variety of water topologies, does exhibit some more overlap. The main structural discrepancy is the overemphasis of a high number of in-layer  $\text{H}_2\text{O}$ – $\text{H}_2\text{O}$  bonds for static, single layer water models, which is most prominent for the lowest energy structures (cf. Table III). These

differences are, on the one hand, linked to the conceptual differences between local minimum structures and finite temperature MD trajectories and, on the other hand, to the neglect of more water layers in the present work, which can influence the configuration of interfacial water through providing donors and acceptors for hydrogen bonds.<sup>51</sup> Some better structural agreement might be obtained already by extending/substituting the dataset of 2D water start configurations with multilayer water models. Whether an implicit solvent environment<sup>22,23,47,85–87</sup> can appropriately emulate the bonding to the second water layer and whether AIMD simulations of minimal explicit–implicit hybrid models can reproduce results from fully explicit AIMD simulations remain an open question. In particular, similar results would only be expected if the implicit model would stabilize stronger non-H-bond-saturated water structures at the interfaces, e.g.,  $\text{H}^{\text{up}}$ , which clarifies the need for such models.

Furthermore, the observed overstructuring of interfacial water might also be influenced by the choice of the substrate supercell, as the studied  $\sqrt{12} \times \sqrt{12}$  Pt(111) substrate allows for commensurate hexagonal water orderings. Our protocol thus also provides a starting point for a consistent analysis of such (artificial) periodic boundary effects, which are often discussed, but not yet well understood.

In general, while the work function change and the adsorption energy can vary dramatically for different static water structures, our 2DSW ensemble averages can provide relatively robust averages that show decent correlations with AIMD results, at least trend-wise. As many applications still necessitate the study of quasi-static water models, e.g., for evaluating reaction barriers via nudged elastic band methods, these results suggest that ensemble averages over a consistent set of static water structures might provide more robust

predictions. Indeed, the application of clustering methods to our 2DSW dataset reveals a total number of 11 topologically different prototype structures (see Sec. S.5.A of the [supplementary material](#)), which indicates that such an approach would still remain computationally much more feasible than brute-force AIMD-based methods.

Finally, our approach, in particular the implementation as a Python package provided alongside this work, might serve as a good starting point for the study of low-temperature water structures (or other solvents) on other surface supercell sizes and substrate materials or as uncorrelated starting points for AIMD simulations.

## SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for (i) a discussion about the choice of supercell in the construction protocol, (ii) a description of the in-house algorithm to find water molecules at the interface, (iii) PCA maps for the water configurations by using the local SOAP descriptors, (iv) density profile and O–O radial distribution function plots for H<sub>2</sub>O@Pt(111) systems listed in [Table I](#), (v) an overview of the prototype structures found in H<sub>2</sub>O@Pt(111) 2DSW, and (vi) further results for Pt(100) and Pt(110) surfaces. For access to the relevant data of this work, see the Data Availability section.

## ACKNOWLEDGMENTS

We would like to thank Professor Dr. Jun Cheng and Dr. Hendrik H. Heenen who provided the AIMD(P) and AIMD(R) trajectories, which are related to Refs. [21](#) and [15](#), respectively. Furthermore, we want to thank Ambrosio *et al.*, Monserrat *et al.*, and Clabaut *et al.* who made the AIMD trajectories and atomistic structures publicly available alongside their publications, namely, Refs. [72](#), [68](#), and [54](#).

N.G.H. acknowledges financial support through the EuroTech Postdoc Programme, which is co-funded by the European Commission under its framework program Horizon 2020 and Grant Agreement No. 754462, and support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—Grant No. EXC 2089/1-390776260. This work was supported by a grant from the Gauss Centre for Supercomputing e.V. ([www.gauss-centre.eu](http://www.gauss-centre.eu)) by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS<sup>88</sup> at the Juelich Supercomputing Centre (JSC).

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to declare.

### Author Contributions

A.C.D.L and T.E. contributed equally to this work.

## DATA AVAILABILITY

The data that support the findings of this study are available within the article and its [supplementary material](#) or from the original publications (cf. Acknowledgments). The newly generated H<sub>2</sub>O@Pt 2DSW structural dataset and Python codes (analysis and

determination of interface water molecules) are openly available at <https://dx.doi.org/10.17617/3.7c>. A Python implementation of the 2DSW protocol is available as an open-source Python package at <https://github.com/computationalelectrochemistrygroup/WaterStructureCreator>.

## APPENDIX: METHODS

### 1. Determination of interfacial water molecules

Interfacial water layers at (artificially created or naturally present) water–vacuum interfaces are selected by first determining all interfacial O atoms and then adding all chemically bound H atoms (bond cutoff = 1.2 Å). Interfacial O atoms are determined via the following algorithm that is inspired from the definition of quantum mechanical cavities in the context of implicit solvation methods:<sup>89,90</sup>

- Define the water-filled region  $\Omega$  by superimposing soft spheres localized on the atoms with given radius  $r_0$  and that smooth decay from a value of 1 within the sphere to 0 outside it. Evaluate the so-obtained filling function  $\Omega$  on a numerical grid and renormalize all values  $\Omega > 1$  to 1.
- Estimate the surface of  $\Omega$  by computing the numerical derivative  $\mathbf{S} = \nabla\Omega$ , and define a grid point  $i$  as an interfacial grid points  $i_s$  whenever

$$\|\mathbf{S}_{i_s}\| \geq \tau \max_j \|\mathbf{S}_j\|, \quad \tau < 1. \quad (\text{A1})$$

- Interface atoms are given by the set of closest oxygen atoms to the real space positions  $\{\mathbf{r}_{i_s}\}$  of all interface grid points.

More details, e.g., on the numerical implementation in periodic boundary conditions and chosen numerical parameters (e.g.,  $r_0$ ,  $\tau$ ) are provided in the [supplementary material](#).

### 2. Computational details

Periodic DFT calculations are performed with the Quantum ESPRESSO package<sup>91</sup> (PWscf) and the Perdew–Burke–Ernzerhof-generalized gradient approximation (PBE-GGA)<sup>74</sup> for the exchange–correlation energy functional. All atoms are represented by ultrasoft pseudopotentials from the open-source GBRV library<sup>92</sup> with density and wave function cutoffs of 360 and 45 Ry, respectively. The water films, modeled with eight H<sub>2</sub>O molecules, are placed on side of the surface, and a dipole correction as implemented in ENVIRON<sup>85</sup> is applied. The Pt surfaces are simulated by using a (3 × 3) supercell with three layers for Pt(100), a (3 × 2) supercell with four layers for Pt(110), and a ( $\sqrt{12} \times \sqrt{12}$ )R30° supercell with three layers for Pt(111). The two bottommost metal layers are fixed at their ideal bulk positions. Slabs are separated by  $\approx 17$  Å. Geometry optimizations are carried through until the forces on all relaxed atoms were smaller than 0.1 eV Å<sup>-1</sup>. Brillouin zone integrations are performed using  $\Gamma$ -centered Monkhorst–Pack meshes of (4 × 4 × 1) and a Marzari–Vanderbilt smearing of 0.02 Ry.

Single point calculations of the interfacial water on Pt(111) for the AIMD simulation from the work of Le *et al.*<sup>21</sup> were performed with the same settings, except that a (2 × 2 × 1) k-point grid was used. The symmetric setup of the original AIMD was split in two

asymmetric setups with three layers of platinum each. The bottom half was rotated by 180° around the y axis to have the same orientation for both sides. No further manipulations were performed on the cell or positions of the atoms.

### 3. Structural similarity measures

We selected the local structure descriptor SOAP<sup>57</sup> as implemented in the Dscribe package<sup>79</sup> and as similarly performed for 3D periodic water structures<sup>68</sup> for the representation of the atomic environment. The local environment  $\mathcal{X}$  around an atom is modeled by summing over Gaussians centered on each atom  $i$  in the structure within a cutoff, yielding the density  $\rho_{\mathcal{X}}(r)$ . After expansion of the density in a basis of orthonormal radial functions  $g_b(|r|)$  and spherical harmonics  $Y_{lm}(\hat{r})$ , the power spectrum  $\mathbf{p}(\mathcal{X})_{b_1 b_2 l}$  can be expressed in terms of expansion coefficients  $c_{blm}$  as

$$\rho_{\mathcal{X}}(r) = \sum_{i \in \mathcal{X}} \exp\left(-\frac{(x_i - r)^2}{2\sigma^2}\right), \quad (\text{A2})$$

$$\rho_{\mathcal{X}}(r) = \sum_{blm} c_{blm} g_b(|r|) Y_{lm}(\hat{r}), \quad (\text{A3})$$

$$\mathbf{p}(\mathcal{X})_{b_1 b_2 l} = \pi \sqrt{\frac{8}{2l+1}} \sum_m (c_{b_1 l m})^\dagger c_{b_2 l m}. \quad (\text{A4})$$

The  $c_{blm}$  coefficients form the components of an abstract SOAP descriptor vector, which is conveniently used for measuring structural similarity throughout this work.

For the analysis of the internal structure of water layers, we evaluate the SOAP vectors centered on the oxygen atoms, including neighboring H and O atoms up to a cutoff radius of 4 Å. For the analysis of water layers on top of Pt(111), we also include the substrate Pt positions in the environment within the sphere defined by the cutoff. In both cases, crossover terms between different atomic species in the power spectrum were allowed to include the information of hydrogen and, in the later case, platinum as well. A description of the entire system was obtained by using the inner average of the SOAP vectors.<sup>79</sup> The hyperparameters for computing the SOAP descriptors used throughout this work are  $n_{\max} = 8$ ,  $l_{\max} = 8$ ,  $\sigma = 0.3$ , and  $r_{\text{cutoff}} = 4$  Å.

For comparing the structural similarity of single structures A with a reference structure B, we use the average distance  $\bar{D}^{93}$  as given by the average kernel  $\bar{K}^{93}$  and defined by

$$\bar{D}(A, B) = \sqrt{2 - 2 \bar{K}(A, B)}, \quad (\text{A5})$$

$$\bar{K}(A, B) = \left[ \frac{1}{N} \sum_i \mathbf{p}(\mathcal{X}_i^A) \right] \cdot \left[ \frac{1}{M} \sum_j \mathbf{p}(\mathcal{X}_j^B) \right], \quad (\text{A6})$$

where  $N$  and  $M$  are the number of atoms in structures A and B, respectively. The so-obtained distance measure  $\bar{D}$  is subsequently averaged over all structures in the reference dataset to obtain an average structural dissimilarity measure  $\langle \bar{D} \rangle$  with the reference ensemble, which we call shortly *average SOAP distance*, here. In the present work, the reference ensemble was created by taking snapshots every 0.1 ps from the AIMD(P) trajectory<sup>21</sup> that was provided by Cheng.

### REFERENCES

- X. Guo, S. Gin, and G. S. Frankel, *npj Mater. Degrad.* **4**, 34 (2020).
- J. H. Stenlid, E. C. Dos Santos, A. Bagger, A. J. Johansson, J. Rossmeisl, and L. G. M. Pettersson, *J. Phys. Chem. C* **124**, 469 (2020).
- J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard, and H. Jónsson, *J. Phys. Chem. B* **108**, 17886 (2004).
- J. K. Nørskov, T. Bligaard, J. Rossmeisl, and C. H. Christensen, *Nat. Chem.* **1**, 37 (2009).
- N. G. Hörmann, M. Jäckle, F. Gossenberger, T. Roman, K. Forster-Tonigold, M. Naderian, S. Sakong, and A. Groß, *J. Power Sources* **275**, 531 (2015).
- Z. W. Seh, J. Kibsgaard, C. F. Dickens, I. Chorkendorff, J. K. Nørskov, and T. F. Jaramillo, *Science* **355**, eaad4998 (2017).
- P. Fenter and N. C. Sturchio, *Prog. Surf. Sci.* **77**, 171 (2004).
- O. Björneholm, M. H. Hansen, A. Hodgson, L.-M. Liu, D. T. Limmer, A. Michaelides, P. Pedevilla, J. Rossmeisl, H. Shen, G. Tocci, E. Tyrone, M.-M. Walz, J. Werner, and H. Bluhm, *Chem. Rev.* **116**, 7698 (2016).
- O. Andreussi, F. Nattino, and N. G. Hörmann, in *Atomic-Scale Modelling of Electrochemical Systems*, edited by M. Melander, T. Laurila, and K. Laasonen (John Wiley & Sons Ltd., 2020).
- K. Schwarz and R. Sundararaman, *Surf. Sci. Rep.* **75**, 100492 (2020).
- L. S. Pedroza, A. Poissier, and M.-V. Fernández-Serra, *J. Chem. Phys.* **142**, 034706 (2015).
- S. K. Natarajan and J. Behler, *Phys. Chem. Chem. Phys.* **18**, 28704 (2016).
- L. Bellarosa, R. García-Muelas, G. Revilla-López, and N. López, *ACS Cent. Sci.* **2**, 109 (2016).
- S. Sakong and A. Groß, *J. Chem. Phys.* **149**, 084705 (2018).
- H. H. Heenen, J. A. Gauthier, H. H. Kristoffersen, T. Ludwig, and K. Chan, *J. Chem. Phys.* **152**, 144703 (2020).
- J.-B. Le, A. Cuesta, and J. Cheng, *J. Electroanal. Chem.* **819**, 87 (2018).
- J.-B. Le, Q.-Y. Fan, J.-Q. Li, and J. Cheng, *Sci. Adv.* **6**, eabb1219 (2020).
- J.-B. Le, A. Chen, L. Li, J.-F. Xiong, J. Lan, Y.-P. Liu, M. Iannuzzi, and J. Cheng, *JACS Au* **1**, 569 (2021).
- Z. K. Goldsmith, M. F. Calegari Andrade, and A. Selloni, *Chem. Sci.* **12**, 5865 (2021).
- A. Hodgson and S. Haq, *Surf. Sci. Rep.* **64**, 381 (2009).
- J.-B. Le, M. Iannuzzi, A. Cuesta, and J. Cheng, *Phys. Rev. Lett.* **119**, 016801 (2017).
- L. Blumenthal, J. M. Kahk, R. Sundararaman, P. Tangney, and J. Lischner, *RSC Adv.* **7**, 43660 (2017).
- N. G. Hörmann, Z. Guo, F. Ambrosio, O. Andreussi, A. Pasquarello, and N. Mazhari, *npj Comput. Mater.* **5**, 100 (2019).
- R. Jinnouchi and A. B. Anderson, *Phys. Rev. B* **77**, 245417 (2008).
- V. Tripkovic, M. E. Björketun, E. Skúlason, and J. Rossmeisl, *Phys. Rev. B* **84**, 115452 (2011).
- T. Ludwig, J. A. Gauthier, K. S. Brown, S. Ringe, J. K. Nørskov, and K. Chan, *J. Phys. Chem. C* **123**, 5999 (2019).
- J. Park and L. T. Roling, *AIChE J.* **66**, e17036 (2020).
- S. Schnur and A. Groß, *Catal. Today* **165**, 129 (2011).
- S. Sakong, K. Forster-Tonigold, and A. Groß, *J. Chem. Phys.* **144**, 194701 (2016).
- D. T. Limmer, A. P. Willard, P. Madden, and D. Chandler, *Proc. Natl. Acad. Sci. U. S. A.* **110**, 4200 (2013).
- A. P. Willard, D. T. Limmer, P. A. Madden, and D. Chandler, *J. Chem. Phys.* **138**, 184702 (2013).
- S. N. Steinmann, R. Ferreira De Morais, A. W. Götz, P. Fleurat-Lessard, M. Iannuzzi, P. Sautet, and C. Michel, *J. Chem. Theory Comput.* **14**, 3238 (2018).
- S. Sakong and A. Groß, *Phys. Chem. Chem. Phys.* **22**, 10431 (2020).
- H. H. Kristoffersen, T. Vegge, and H. A. Hansen, *Chem. Sci.* **9**, 6912 (2018).
- A. Bouzid and A. Pasquarello, *J. Phys. Chem. Lett.* **9**, 1880 (2018).
- X.-Y. Li, A. Chen, X.-H. Yang, J.-X. Zhu, J.-B. Le, and J. Cheng, *J. Phys. Chem. Lett.* **12**, 7299 (2021).
- G. S. Karlberg, *Phys. Rev. B* **74**, 153414 (2006).

- <sup>38</sup>J. Rossmeisl, J. Greeley, and G. Karlberg, "Electrocatalysis and catalyst screening from density functional theory calculations," in *Fuel Cell Catalysis* (John Wiley & Sons Ltd., 2008), Chap. 3, pp. 57–92.
- <sup>39</sup>M. E. Björketun, Z. Zeng, R. Ahmed, V. Tripkovic, K. S. Thygesen, and J. Rossmeisl, *Chem. Phys. Lett.* **555**, 145 (2013).
- <sup>40</sup>Z.-D. He, S. Hanselman, Y.-X. Chen, M. T. M. Koper, and F. Calle-Vallejo, *J. Phys. Chem. Lett.* **8**, 2243 (2017).
- <sup>41</sup>Q. Zhang and A. Asthagiri, *Catal. Today* **323**, 35 (2019).
- <sup>42</sup>E. Skúlason, G. S. Karlberg, J. Rossmeisl, T. Bligaard, J. Greeley, H. Jónsson, and J. K. Nørskov, *Phys. Chem. Chem. Phys.* **9**, 3241 (2007).
- <sup>43</sup>S. A. Akhade, N. J. Bernstein, M. R. Esopi, M. J. Regula, and M. J. Janik, *Catal. Today* **288**, 63 (2017).
- <sup>44</sup>J. Rossmeisl, E. Skúlason, M. E. Björketun, V. Tripkovic, and J. K. Nørskov, *Chem. Phys. Lett.* **466**, 68 (2008).
- <sup>45</sup>K. Chan and J. K. Nørskov, *J. Phys. Chem. Lett.* **6**, 2663 (2015).
- <sup>46</sup>J. A. Gauthier, C. F. Dickens, H. H. Heenen, S. Vijay, S. Ringe, and K. Chan, *J. Chem. Theory Comput.* **15**, 6895 (2019).
- <sup>47</sup>G. Kastlunger, P. Lindgren, and A. A. Peterson, *J. Phys. Chem. C* **122**, 12771 (2018).
- <sup>48</sup>J. Hussain, H. Jónsson, and E. Skúlason, *ACS Catal.* **8**, 5240 (2018).
- <sup>49</sup>J. Hussain, H. Jónsson, and E. Skúlason, *Faraday Discuss.* **195**, 619 (2016).
- <sup>50</sup>S. Schnur and A. Groß, *New J. Phys.* **11**, 125003 (2009).
- <sup>51</sup>T. Roman and A. Groß, *Catal. Today* **202**, 183 (2013).
- <sup>52</sup>S. Meng, E. G. Wang, and S. Gao, *Phys. Rev. B* **69**, 195404 (2004).
- <sup>53</sup>J. Carrasco, A. Hodgson, and A. Michaelides, *Nat. Mater.* **11**, 667 (2012).
- <sup>54</sup>P. Clabaut, R. Staub, J. Galiana, E. Antonetti, and S. N. Steinmann, *J. Chem. Phys.* **153**, 054703 (2020).
- <sup>55</sup>S. Standop, A. Redinger, M. Morgenstern, T. Michely, and C. Busse, *Phys. Rev. B* **82**, 161412 (2010).
- <sup>56</sup>A. Groß, F. Gossenberger, X. Lin, M. Naderian, S. Sakong, and T. Roman, *J. Electrochem. Soc.* **161**, E3015 (2014).
- <sup>57</sup>A. P. Bartók, R. Kondor, and G. Csányi, *Phys. Rev. B* **87**, 184115 (2013).
- <sup>58</sup>K. Tonigold and A. Gross, *J. Comput. Chem.* **33**, 695 (2012).
- <sup>59</sup>A. Massey, F. McBride, G. R. Darling, M. Nakamura, and A. Hodgson, *Phys. Chem. Chem. Phys.* **16**, 24018 (2014).
- <sup>60</sup>T. Roman and A. Groß, *J. Phys. Chem. C* **120**, 13649 (2016).
- <sup>61</sup>J. Chen, G. Schusteritsch, C. J. Pickard, C. G. Salzmann, and A. Michaelides, *Phys. Rev. Lett.* **116**, 025501 (2016).
- <sup>62</sup>F. Corsetti, P. Matthews, and E. Artacho, *Sci. Rep.* **6**, 18651 (2016).
- <sup>63</sup>F. Corsetti, J. Zubeltzu, and E. Artacho, *Phys. Rev. Lett.* **116**, 085901 (2016).
- <sup>64</sup>S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, *Comput. Mater. Sci.* **68**, 314 (2013).
- <sup>65</sup>R. Tran, Z. Xu, B. Radhakrishnan, D. Winston, W. Sun, K. A. Persson, and S. P. Ong, *Sci. Data* **3**, 160080 (2016).
- <sup>66</sup>W. Sun and G. Ceder, *Surf. Sci.* **617**, 53 (2013).
- <sup>67</sup>A. Zur and T. C. McGill, *J. Appl. Phys.* **55**, 378 (1984).
- <sup>68</sup>B. Monserrat, J. G. Brandenburg, E. A. Engel, and B. Cheng, *Nat. Commun.* **11**, 5757 (2020).
- <sup>69</sup>Y. Zhang and W. Yang, *Phys. Rev. Lett.* **80**, 890 (1998).
- <sup>70</sup>C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).
- <sup>71</sup>S. Grimme, A. Hansen, J. G. Brandenburg, and C. Bannwarth, *Chem. Rev.* **116**, 5105 (2016).
- <sup>72</sup>F. Ambrosio, Z. Guo, and A. Pasquarello, *J. Phys. Chem. Lett.* **9**, 3212 (2018).
- <sup>73</sup>G. Miceli, S. de Gironcoli, and A. Pasquarello, *J. Chem. Phys.* **142**, 034501 (2015).
- <sup>74</sup>J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- <sup>75</sup>S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *J. Chem. Phys.* **132**, 154104 (2010).
- <sup>76</sup>S. Grimme, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **1**, 211 (2011).
- <sup>77</sup>B. Hammer, L. B. Hansen, and J. K. Nørskov, *Phys. Rev. B* **59**, 7413 (1999).
- <sup>78</sup>J. Klimeš, D. R. Bowler, and A. Michaelides, *J. Phys.: Condens. Matter* **22**, 022201 (2009).
- <sup>79</sup>L. Himanen, M. O. J. Jäger, E. V. Morooka, F. Federici Canova, Y. S. Ranawat, D. Z. Gao, P. Rinke, and A. S. Foster, *Comput. Phys. Commun.* **247**, 106949 (2020).
- <sup>80</sup>F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- <sup>81</sup>S. Surendralal, M. Todorova, and J. Neugebauer, *Phys. Rev. Lett.* **126**, 166802 (2021).
- <sup>82</sup>J. Carrasco, J. Klimeš, and A. Michaelides, *J. Chem. Phys.* **138**, 024708 (2013).
- <sup>83</sup>S. Trasatti, *Electrochim. Acta* **35**, 269 (1990).
- <sup>84</sup>B. Cheng, R.-R. Griffiths, S. Wengert, C. Kunkel, T. Stenzel, B. Zhu, V. L. Deringer, N. Bernstein, J. T. Margraf, K. Reuter, and G. Csányi, *Acc. Chem. Res.* **53**, 1981 (2020).
- <sup>85</sup>O. Andreussi, I. Dabo, and N. Marzari, *J. Chem. Phys.* **136**, 064102 (2012).
- <sup>86</sup>O. Andreussi and G. Fiscaro, *Int. J. Quantum Chem.* **119**, e25725 (2019).
- <sup>87</sup>N. G. Hörmann, O. Andreussi, and N. Marzari, *J. Chem. Phys.* **150**, 041730 (2019).
- <sup>88</sup>Jülich Supercomputing Centre, *J. Large-Scale Res. Facil.* **5**, A135 (2019).
- <sup>89</sup>G. Fiscaro, L. Genovese, O. Andreussi, S. Mandal, N. N. Nair, N. Marzari, and S. Goedecker, *J. Chem. Theory Comput.* **13**, 3829 (2017).
- <sup>90</sup>O. Andreussi, N. G. Hörmann, F. Nattino, G. Fiscaro, S. Goedecker, and N. Marzari, *J. Chem. Theory Comput.* **15**, 1996 (2019).
- <sup>91</sup>P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. Otero-de-la-Roza, L. Paulatto, S. Poncè, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, *J. Phys.: Condens. Matter* **29**, 465901 (2017).
- <sup>92</sup>K. F. Garrity, J. W. Bennett, K. M. Rabe, and D. Vanderbilt, *Comput. Mater. Sci.* **81**, 446 (2014).
- <sup>93</sup>S. De, A. P. Bartók, G. Csányi, and M. Ceriotti, *Phys. Chem. Chem. Phys.* **18**, 13754 (2016).