Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue



Check for updates

Cite this: *Digital Discovery*, 2022, **1**, 703

# Neural-network-backed evolutionary search for $SrTiO_3(110)$ surface reconstructions†

Ralf Wanzenböck, Marco Arrigoni, Sebastian Bichelmaier, Florian Buchner, Jesús Carrete and Georg K. H. Madsen *

Received 4th July 2022
Accepted 23rd August 2022

DOI: 10.1039/d2dd00072e

rsc.li/digitaldiscovery

The determination of atomic structures in surface reconstructions has typically relied on structural models derived from intuition and domain knowledge. Evolutionary algorithms have emerged as powerful tools for such structure searches. However, when density functional theory is used to evaluate the energy the computational cost of a thorough exploration of the potential energy landscape is prohibitive. Here, we drive the exploration of the rich phase diagram of $TiO_x$ overlayer structures on $SrTiO_3(110)$ by combining the covariance matrix adaptation evolution strategy (CMA-ES) and a neural-network force field (NNFF) as a surrogate energy model. By training solely on $SrTiO_3(110)$ 4×1 overlayer structures and performing CMA-ES runs on 3×1, 4×1 and 5×1 overlayers, we verify the transferability of the NNFF. The speedup due to the surrogate model allows taking advantage of the stochastic nature of the CMA-ES to perform exhaustive sets of explorations and identify both known and new low-energy reconstructions.

## 1 Introduction

The determination of surface structure is vital in gaining a better understanding of the properties and possible applications of materials. Traditionally, being able to identify the true atomic structure of materials from experimental findings is an acquired skill, based on experience and domain knowledge.[1–4] Due to its semiconducting nature and possible applications in electronic devices, strontium titanate ($SrTiO_3$) has increasingly been the focus of experimental and theoretical studies.[5–10] Since many possible applications of $SrTiO_3$ are realized in the form of nanoscale thin films, its surfaces are of particular interest.

Here, we are interested in exploring the phase diagram of $TiO_x$ overlayer structures on $SrTiO_3(110)$ under Ti-poor conditions.[11] The polar nature of the bulk terminated $SrTiO_3(110)$ surfaces is balanced by a $Ti_{n+2}O_{3n+4}$ overlayer with a $n\times1$ unit cell.[2] The $n$ increases systematically with Sr chemical potential[11] and an early model, based on simple geometric considerations and DFT calculations, proposed a systematic set of overlayer structures consisting of a ring of six Ti–O tetrahedra and an additional ring of increasing length so that a $n = 2$ unit cell would exhibit a 6–6 overlayer, a $n = 3$ unit cell a 6–8 overlayer and so on and so forth.[2] Subsequently, DFT investigations refined the picture[3,12–14] and showed, *e.g.*, how an 8–10 overlayer structure is stable in the 5×1 cell.[12] However, despite the decade-long interest, the understanding of the phase diagram is still based on the investigation of specific structures derived from experience and geometric considerations.

The wide availability of first-principles computations based on density functional theory (DFT), combined with the steadily increasing computational power, have greatly improved the feasibility of global geometry optimization. Evolutionary algorithms[15,16] combined with these first-principles calculations have proven to be powerful tools for structure searches,[17–22] including the investigation of surface reconstructions.[23] However, the stochastic nature of the corresponding algorithms means that the computational cost and run-time associated with a thorough exploration of the potential energy surface (PES) quickly strain available computational resources. With the growing availability of machine-learned (ML) approximations of the PES[24] various incarnations have been successfully adapted for use with evolutionary algorithms in structure prediction[25,26] and surface reconstructions.[27,28] Due to the large variety of structures that can potentially be visited during an evolutionary search, the construction of a suitable ML PES approximation is very challenging. Adaptive approaches have been used where the ML approximation is updated every time a point of high uncertainty is visited.[28] However, the estimation of uncertainty is still a relatively open question in ML.[29] Alternatively, the ML PES should be trained on a highly diverse set of structures.[27] In most cases though, this is not practically feasible and gathering training data from the entire PES *a priori* would also somewhat defy the objective of making an ML PES approximation.

In the present work we perform an exploration of the $TiO_x$ overlayer structures by combining the covariance matrix adaptation evolution strategy (CMA-ES)[30] and a fully automatically

*Institute of Materials Chemistry, TU Wien, 1060 Vienna, Austria. E-mail: georg.madsen@tuwien.ac.at*

† Electronic supplementary information (ESI) available. See https://doi.org/10.1039/d2dd00072e

differentiable neural-network force field (NNFF).[31] First, we discuss how an implementation of the CMA-ES is adapted to deal with surface structures. We use the initial DFT based CMA-ES run to set up a diverse set of training data and construct an NNFF. We train the initial NNFF solely on $SrTiO_3(110)$ 4×1 overlayer structures and perform CMA-ES runs on 3×1, 4×1 and 5×1 overlayers. We verify the transferability of the NNFF in a diagnostic approach by constructing NNFFs including data from the overlayers not part of the original training dataset and utilizing similarity in the local environments. With this we are able to investigate larger systems, while performing the majority of expensive DFT calculations on more accessible structures. Furthermore, the NNFFs make it possible to perform sets of CMA-ES runs, taking advantage of the stochastic nature of the algorithm to more fully explore the PES. We show how the procedure reproduces known structures and also finds new stable structures for all three overlayers considered.

## 2 Methods

### 2.1 Slab setup

We use a slab setup starting from a bulk-terminated structure composed of five SrTiO and six $O_2$ alternating layers with vacuum separating the surface atoms from the periodic recurrence of the slab along the surface normal. To both sides of the slab roughly evenly spaced $Ti_2O_5$ and $TiO_2$ units were added according to the stoichiometry $Ti_{n+2}O_{n+4}$ for unit cells with $n$×1 periodicity. With this the $SrTiO_3(110)$ 3×1, 4×1 and 5×1 unit cells contain 105, 136 and 167 atoms each in total. For the subsequent CMA-ES runs we refer to these structures as the founder structures, Fig. 1 (see also deposited data[32]). The outermost $O_2$ layer and added Ti–O units comply with the target composition $Ti_{n+2}O_{3n+4}{}^{2}$ and are referred to as the $TiO_x$ overlayers. The atoms in the innermost region (red background in Fig. 1) are fixed to act as an anchor and the opposite sides of the slab are symmetrical. The mirror atoms' positions $\boldsymbol{r}_i^m$ are obtained from the original positions $\boldsymbol{r}_i^m$ by a point inversion about the center of the slab. The width of the fixed bulk layer (three SrTiO and two $O_2$ layers in Fig. 1) is controlled by a computational parameter (the depth $d$) discussed below.

### 2.2 CMA-ES

The CMA-ES[30,33] draws a sample of $\lambda$ individuals $\boldsymbol{x}_k^{(g)}$, $k = 1, \ldots, \lambda$ at each generation $g$ from a multivariate normal distribution:

$$\mathbf{x}_k^{(g)} \sim \mathcal{N}\left(\boldsymbol{m}^{(g-1)}, \left[\sigma^{(g-1)}\right]^2 \boldsymbol{C}^{(g-1)}\right), \tag{1}$$

where $\boldsymbol{m}$ represents the distribution mean, $\sigma$ the step size and $\boldsymbol{C}$ the covariance matrix. The distribution is iteratively adapted by the algorithm after each generation through the use of so-called evolution paths,[33] which are able to take into account the evolutionary history in the parameters update.

We employ the standard hyperparameter settings,[33] including the initialization of the covariance matrix to identity and choosing the population size *via* $\lambda = 4 + \lfloor 3 \log D \rfloor$, where $D$ are the degrees of freedom. The initial step size, which controls the width of the initial distribution, is set to $\sigma^{(0)} = 0.12$ Å.

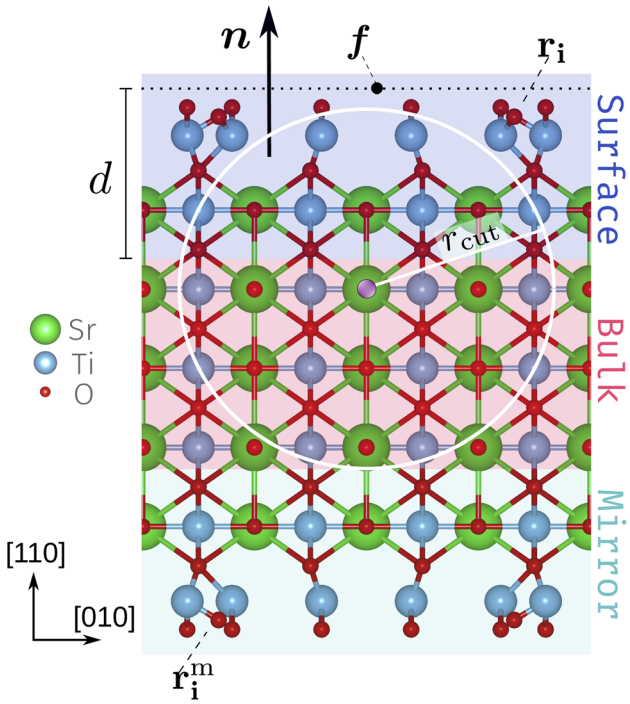


Fig. 1 Side view of an $SrTiO_3(110)$ 4×1 founder slab. The schematic diagram includes the surface normal $\boldsymbol{n}$, defect focus $\boldsymbol{f}$, depth $d$ and the surface atoms $\boldsymbol{r}_i$ with their mirrors images $\boldsymbol{r}_i^m$. The white circle and radius $r_{cut}$ represent a 2D slice of the local environment of the atom highlighted in purple. The fixed bulk-like anchor (red), the surface layers that can be accessed by the CMA-ES algorithm (blue) and their mirror layers (green) highlighted. The highlighting corresponds to a choice of $d = 6$ Å.

We recently implemented the CMA-ES for point defect structural exploration.[26] Here, we extend the implementation to efficiently work on surface structures. The CMA-ES directly determines the positions of only the surface atoms (overlayer and outermost bulk, highlighted in blue in Fig. 1) $\boldsymbol{r}_i$; the mirror atoms $\boldsymbol{r}_i^m$ on the opposite side (highlighted in green) are adjusted accordingly. To specify which atoms the algorithm has access to, the surface normal $\boldsymbol{n}$, a defect focus $\boldsymbol{f}$ (on a plane perpendicular to $\boldsymbol{n}$) and a depth d[Å] need to be defined. In this work we set the defect focus such that it marks the lower edge of the surface slab along the c-axis and the surface normal as $\boldsymbol{n} = (0, 0, 1)$. With these settings constant, solely the depth $d$ controls which atoms the algorithm has access to.

### 2.3 DFT

We utilize GPAW[34] in linear-combination-of-atomic-orbitals (LCAO)[35] mode as the energy backend for the first CMA-ES runs and for the local gradient based optimizations and, furthermore, for the generation of training data and evaluation of structures gathered from NNFF-backed CMA-ES runs. All GPAW calculations are performed employing the Perdew–Burke–Ernzerhof (PBE)[36] functional. Simulation boxes with periodic boundary conditions perpendicular to the surface normal are used for the GPAW calculations and the $k$-point grid is set to (2, 2, 1). For reference some structures were also

 

optimized locally using VASP[37] with periodic boundary conditions applied along all axes. Several individuals are optimized locally *via* the FIRE[38] algorithm, built into the atomic simulation environment (ASE).[39]

To investigate the diversity of the training data, we calculated the net atomic charges (NAC) using the Chargemol[40] program, which implements the DDEC6 approach.[41,42]

### 2.4 NNFF

We train NNFFs following the NeuralIL methodology.[31] The implementation is based on JAX,[43] which offers just-in-time compilation and end-to-end automatic differentiation allowing us to train on all 3 $n_{\text{atoms}}$ forces within each structure. The Cartesian coordinates are encoded using the power spectrum of element-specific atom-centered spherical Bessel descriptors[31,44,45] and atom types are factored in *via* an embedding layer.[31,46] The importance of encoding element specific environments for multicomponent systems was previously pointed out.[31] In the present work, these were encoded with a resolution set by $n_{\max} = 5$ within a cutoff radius of $r_{\text{cut}} = 5.5$ Å, resulting in 126 descriptors per atom. For all models we use a fully connected pyramidal architecture with hidden layers consisting of 256:128:64:32:16:16:16 nodes.

The training is performed over 350 epochs with a one-cycle learning rate schedule[47] varying the learning rate linearly from $3\times10^{-4}$ to $3\times10^{-3}$ and back, and switching to $3\times10^{-5}$ for the last 10% of each epoch, with the data being split into minibatches of eight structures each. To reduce the influence of outliers, the log-cosh loss function[48] is employed, with a characteristic scale parameter of 0.1 eV Å$^{-1}$, effectively clipping very large forces.

## 3 Results and discussion

We first performed three CMA-ES runs with GPAW as the fitness backend. All runs started from the same founder structure with the depth $d$ being switched from 3 to 6 and to 9 Å, see Fig. 1. This allowed the CMA-ES to manipulate the positions of atoms in only the $TiO_x$ overlayer for $d = 3$ Å, additionally one SrTiO layer and one $O_2$ layer for $d = 6$ Å (as depicted in Fig. 1) and all but the central SrTiO layer for $d = 9$ Å. This corresponded to totals of 22, 42 and 62 atoms, respectively. Each run produced 500 generations of 22 individuals, totaling 33 000 structures among the three runs.

Fig. 2 shows the trajectory of the average energy of the populations and the CMA-ES step size. The step size, $\sigma$, was automatically adjusted by the algorithm, narrowing or widening the underlying distribution [eqn. (1)] as necessary. The inset in Fig. 2 displays the overlayer reconstruction of a $SrTiO_3(110)$ 4×1 surface that was identified by local optimization of the lowest energy structure generated by the CMA-ES for depth $d = 6$ Å. The structure reproduces the expected overlayer with rings of corner-sharing $TiO_4$ consisting of six and ten members, respectively.[2,3] However, it does not represent the ground-state overlayer structure, as will be discussed later.

To perform sets of multiple evolutionary runs within reasonable time frames and, thereby, utilize the stochastic nature of the CMA-ES algorithm to explore more distinct local minima of the PES, a NNFF model was trained on selected $SrTiO_3(110)$ 4×1 structures. The training data was obtained from the three DFT-based CMA-ES trajectories described above. From these we used the individual structure with the lowest energy and a randomly chosen structure of each generation of each run. For the validation data, an individual from each generation was chosen at random. With this we arrived at 3000 configurations in the training data and 500 in the validation data. Finally, 5000 structures were chosen completely at random from the available DFT calculations as test data, only

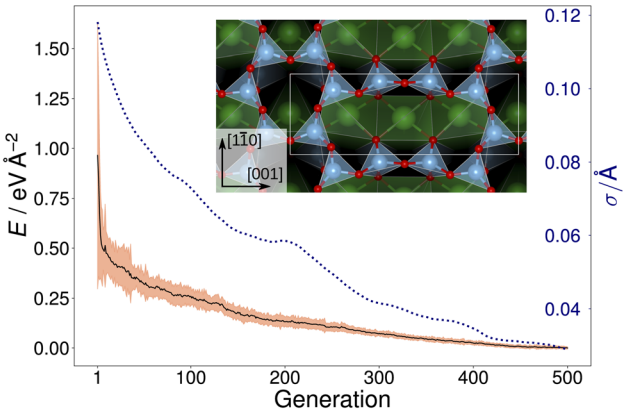


Fig. 2 Trajectory (black solid line) of the average energy of the population over 500 generations of a GPAW-backed CMA-ES run on a $SrTiO_3(110)$ 4×1 structure with $d = 6$ Å. The three times the standard deviation of energies within each generation is highlighted in orange. The area highlighted in orange has a height equal to three times the standard deviation of the energies within each generation. The dotted blue line shows the CMA-ES parameter step-size $\sigma$. In the inset, the local optimization of the resulting overlayer consisting of rings of corner-sharing $TiO_4$ tetrahedra is shown; the unit cell is indicated by a white rectangle.

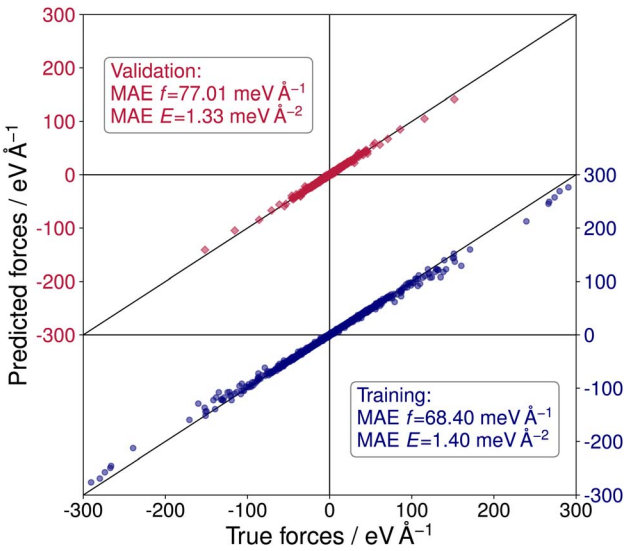


Fig. 3 Predicted *vs.* true forces for the NeuralIL model $NNFF_1$ on training (blue dots) and validation data (red squares). MAEs are given for forces and energies.

Table 1 Performance of three trained NeuralIL models on distinct test data sets (5000 4×1 structures, 1100 5×1 structures, 149 GPAW-optimized 5×1 structures), comparing the MAEs of the force components and of the energies per surface area

| Test set | MAE | $NNFF_1$ | $NNFF_2$ | $NNFF_3$ |
|---|---|---|---|---|
| 4×1 | $f$/meV Å$^{-1}$ | 77.19 | 84.22 | 79.76 |
| | $E$/meV Å$^{-2}$ | 1.35 | 2.07 | 1.61 |
| 5×1 | $f$/meV Å$^{-1}$ | 278.55 | 123.46 | 126.37 |
| | $E$/meV Å$^{-2}$ | 9.89 | 1.92 | 2.21 |
| Opt | $f$/meV Å$^{-1}$ | 175.40 | 105.35 | 107.66 |
| | $E$/meV Å$^{-2}$ | 8.95 | 7.72 | 6.59 |

excluding previously selected data. The resulting model, labeled $NNFF_1$, achieved a mean absolute error (MAE) of 1.79 meV per atom or in terms of surface energy 1.40 meV Å$^{-2}$ for training energies and 68.40 meV Å$^{-1}$ for training forces. The NNFF performance on training and validation data is shown in Fig. 3, which does not show any indication of overfitting. Similar MAEs were also found for the test set, labeled 4×1 in Table 1.

By gathering the data from CMA-ES runs, a large variety of structures were included in the data sets, with a standard deviation of 201.11 meV per atom or 150.08 meV Å$^{-2}$ in the training energies and 2.35 eV Å$^{-1}$ in the training force components, with maximal training forces of up to 290.53 eV Å$^{-1}$. The diversity of the training data is illustrated with the net atomic charges[41,42] on titanium. Fig. 4 compares the resulting charges to known titanium oxidation states. It is seen that the charges within the surface slabs cover an even larger range than the illustrated reference systems. The trained $NNFF_1$ did not show any issues handling the different oxidation states.

With the goal of a broader exploration of the respective PES, a set of 50 CMA-ES evolution runs was performed on $SrTiO_3$(110) 3×1, 4×1 and 5×1 surfaces utilizing $NNFF_1$. We selected $d = 6$ Å thus allowing the CMA-ES to manipulate two

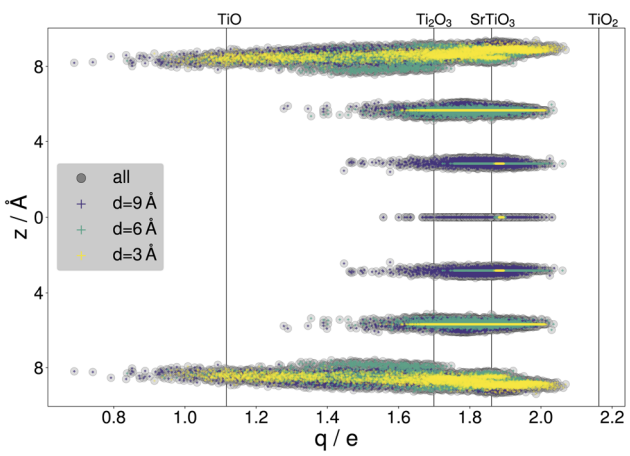


Fig. 4 Net atomic charges of titanium atoms within the $SrTiO_3$(110) 4×1 training structures gathered from all three DFT-backed searches. The Ti charges are highlighted according to the depth parameter, with results for $d = 3$ Å in yellow, $d = 6$ Å in green and $d = 9$ Å in purple. $z = 0$ Å marks the SrTiO layer in the center of the surface slabs. The vertical lines indicate charges calculated for bulk systems with titanium in different oxidation states: TiO, $Ti_2O_3$, $TiO_2$ (rutile) and $SrTiO_3$.

$SrTiO_3$ layers. Within a set the same founder structure was used for all runs, whereas the initial random seed of the CMA-ES was varied. For all runs the maximum number of generations was set to at least 500, with the additional stopping criterion of the standard deviation of the individuals' energies within a generation going below 50 meV per unit cell. Every generation of 4×1 and 5×1 structures consisted of 22 individuals, while this number was automatically set to 21 by the algorithm for 3×1 due to the lower number of atoms in the structure. Otherwise, all CMA-ES parameters were identical for all investigated systems. This led the algorithm to arrive at a diverse collection of individuals belonging to a number of different local minima on the PES for each system. The best individual of each CMA-ES run within a set was optimized locally using first $NNFF_1$ and subsequently GPAW as the backend.

For the $SrTiO_3$(110) 3×1 surface slab, 49 of the 50 candidate structures fell on three distinct energy levels corresponding to the overlayer reconstructions shown in Fig. 5(a)–(c). The corresponding CMA-ES energy trajectories are shown in Fig. 6, with four runs arriving at structure (b) and the rest evenly distributed between (a) and (c). All three 3×1 overlayers include six- and eight-membered rings of corner-sharing $TiO_4$ tetrahedra, with the smaller ring taking triangular or rhombic shape. The $TiO_x$ overlayers in (a) and (c) are related by a shift of half a lattice unit along the [1$\bar{1}$0] direction, which results in a different connectivity to the substrate. As expected from the literature,[2] reconstruction (a) arrived at the lowest energy. Additionally, one structure for each of the three reconstructions pictured in Fig. 5(a)–(c) was further optimized using VASP. The overlayers did not change through the VASP optimization and the relative order of the energies was the same: with the minimum VASP energy of (a) set to zero, (b) arrived at 4.8 meV Å$^{-2}$ and (c) at 5.9 meV Å$^{-2}$.

The same strategy was applied to $SrTiO_3$(110) 4×1 surfaces. Here, the GPAW-backed optimization of the 50 best individuals identified two distinct low-energy overlayers presented in Fig. 5(d) and (e), with the rest of the structures spanning an energy range of 100 meV Å$^{-2}$. The configuration with the third-lowest energy, shown in Fig. 5(f) interestingly reproduces the result of the DFT-backed CMA-ES, Fig. 2, which underlines the need for running multiple sets of PES explorations. All three overlayers, Fig. 5(d–f), showed corner-sharing $TiO_4$ tetrahedra arranged in rings of six and ten members, with the smaller ring taking the shape of a rhombus (d) or a triangle (e) and (f), respectively. Structure (e) is the expected stable structure.[2,3] However, overlayer (d) represents the computational energy minimum and suggests an alternative organization of the overlayer, exhibiting a $p2$ symmetry in contrast to the mirror symmetry in (e) and (f). The rhombohedral structural motif of the six-membered ring is similar to that of the 3×1 candidate in Fig. 5(b). Notably, regions exibiting $p2$ symmetry can be identified in STM images published by Wang *et al.* as Fig. 1(a),[14] near Type-II vacancies. To illustrate this observation a simulated STM image of the structure in Fig. 5(d) was generated utilizing the Tersoff–Hamann approximation.[49] The result can be seen in Fig. 1 of the ESI together with the experimental STM image and indeed a very good agreement is observed. Configurations

 

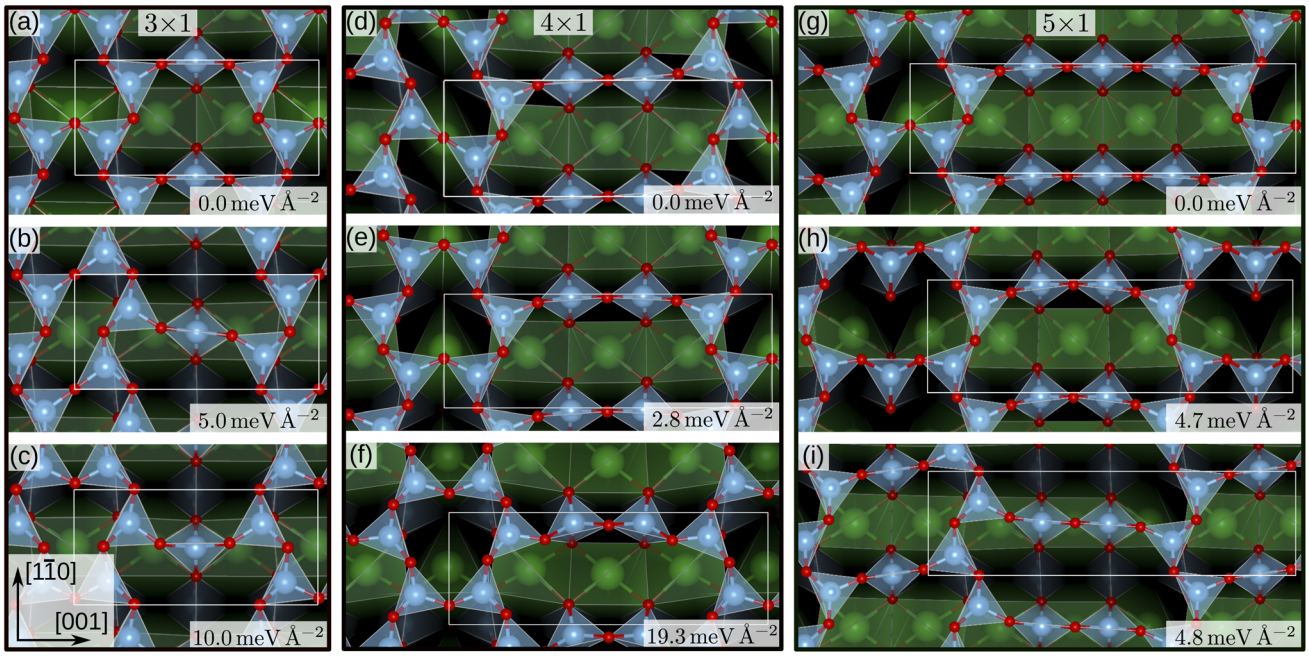


Fig. 5 $SrTiO_3$(110) 3×1, 4×1 and 5×1 reconstruction overlayers identified by performing sets of NNFF-backed CMA-ES runs and further refined by two subsequent optimizations, driven by the NNFFs and GPAW. All structures show corner-sharing $TiO_4$ tetrahedra in different arrangements of six-, eight-, ten- or twelve-membered rings. For each of the system sizes the calculated energy minimum is set to zero. Structures (a), (e) and (g) reproduce overlayers in agreement with literature.[2] The energies shown were evaluated using GPAW and offset to the computational minimum per reconstruction.

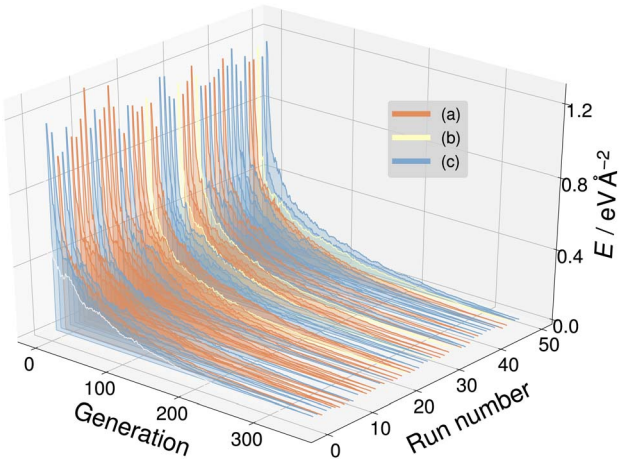


Fig. 6 The energy trajectories of the 50 CMA-ES runs on the $SrTiO_3$(110) 3×1 surface, with the calculated energy minimum set to zero. The labels (a) to (c) correspond to the overlayers shown in Fig. 5(a)–(c). Run number zero (white) is an outlier.

representing each minimum were optimized with VASP. The overlayer structures remained unchanged. The energy differences became smaller, with (d) and (e) differing by only 0.08 meV Å$^{-2}$ and (f) lying 9.7 meV Å$^{-2}$ higher, but the order remained the same.

Finally, we performed a set of 50 CMA-ES runs for the $SrTiO_3$(110) 5×1 surfaces. The GPAW-optimized CMA-ES results identified a number of overlayers consisting of corner-sharing $TiO_4$ tetrahedra arranged in rings with structures representing variations of six- and twelve-membered rings or eight- and ten-membered rings. The lowest-energy structures are shown in Fig. 5(g), a six-twelve structure proposed in literature,[2] and in (i), an eight–ten structure with $p2$ symmetry with a 4.8 meV Å$^{-2}$ higher energy than (g). To the best of our knowledge the $p2$ structure in (i) has not been proposed before and we have deposited the calculated STM image in Fig. 2 of the ESI†.

The predictions of ML models can be expected to have the largest uncertainty for data with a low similarity to the training data.[29] Intuitively, the structures with the lowest similarity to the 4×1 training data would be found among the structures visited in the 5×1 searches. To investigate the reliability of the NNFF we trained a new NNFF by adding 5×1 structures to the training data. To that end, 3500 out of the 550 000 available 5×1 structures generated by the 50 $NNFF_1$ CMA-ES runs were selected randomly for evaluation with GPAW. We then gathered 4000 structures as training data and 800 for validation, evenly distributed in 4×1 and 5×1 data points. The remaining 1100 new 5×1 DFT data were collected in a second test set, labeled 5×1 in Table. 1. The resulting model was labeled "$NNFF_2$" and achieved an MAE on training forces of 83.53 meV Å$^{-1}$ and 93.96 meV Å$^{-1}$ for validation. The force and energy MAEs for $NNFF_1$ and $NNFF_2$ for the two test sets are given in Table. 1. As expected, the performance on a test set exclusively containing 5×1 structures improved significantly, with $NNFF_2$ achieving an MAE of 123.5 meV Å$^{-1}$ on the forces, in comparison to 278.6 meV Å$^{-1}$ for $NNFF_1$. The performance on the 4×1 test set deteriorates somewhat, from 77.2 meV Å$^{-1}$ to 84.2 meV Å$^{-1}$.

This, however, is expected because of the reduced number of 4×1 structures in the training data.

Utilizing $NNFF_2$, we performed a second set of 50 CMA-ES runs on the 5×1 surface slab, which again resulted in the low-energy structures (g) and (i) in Fig. 5. This can be seen as a confirmation of the transferability of the $NNFF_1$ model. However, it is also clear that the energy differences between the structures (g) and (i) are on the same scale as the MAE of both $NNFF_1$ and $NNFF_2$. When evaluating the energies of the two structures we also see that $NNFF_1$ swaps their order, while with $NNFF_2$ they become almost energetically degenerate, as seen in Fig. 7. So while both NNFFs are able to distinguish the two basins, it is an open question how much the energy uncertainty influences the CMA-ES exploration of the PES.

To investigate this question, a third data set was created by pooling all previously used 4×1 and 5×1 training and validation structures and augmenting the data with 350 5×1 structures representing the minima found. These 350 additional structures were randomly selected from the GPAW optimization trajectories of seven CMA-ES results, two from set one ($NNFF_1$) and five from set two ($NNFF_2$), representing various energy basins. The resulting data was then randomly split into 5000 training data and 1250 validation data and the model, labeled "$NNFF_3$", trained over 500 epochs with all other parameters unchanged, achieving an MAE on training forces of 76.90 meV $Å^{-1}$ and 87.69 meV $Å^{-1}$ for validation. By pooling the whole collection of 4×1 and 5×1 data, the model performed well on both the 4×1 and 5×1 test sets, Table. 1. We also constructed a new test set, denoted as "Opt" in Table. 1, consisting of the GPAW converged minima of all three CMA-ES sets that were performed on 5×1 structures. Not surprisingly, $NNFF_3$ performed well for this test set. Importantly, $NNFF_3$ was able to correctly reproduce the order of the lowest energy structures found for the 5×1 reconstruction, Fig. 7. Subsequently, a third set of 50 CMA-ES runs was performed using $NNFF_3$. These runs again found variations of the known overlayer configurations but also added Fig. 5(h) as the overall second lowest energy minimum to the results. The structure is similar to the *pm* eight–ten overlayer structure found by Li *et al.*[12] It does, however, not fully reproduce the earlier structure where the bridging Ti tetrahedra of the eight ring shares an edge with the substrate whereas the same tetrahedra in (h) shares a corner.

That neither the (h) structure nor the structure proposed by Li were found in the $NNFF_1$ CMA-ES run raises the question whether this is due to these structures not being identified as energy minima using the $NNFF_1$ or simply a result of the increasing dimensionality of search space with cell size and the stochastic nature of the CMA-ES. To answer this question, we performed a local relaxation with $NNFF_1$ starting from the non-optimized (h) structure originally acquired using $NNFF_2$ and a reconstructed version of the lowest energy structure proposed by Li *et al.*[12] We verified that $NNFF_1$ correctly identifies the corresponding energy basins on the surrogate PES. From this follows that out of a sufficiently large set of randomly initialized $NNFF_1$-backed CMA-ES runs, some would arrive at the structures, without the need for data augmentation. These structures being missed in the original 50 CMA-ES runs we interpret as a result of the dimensionality of the PES being very large for the 5×1 reconstruction. This is also evident in the results obtained for the 3×1 CMA-ES run, where 49 out of 50 runs result in the (a)–(c) structures, as compared to the 4×1 CMA-ES run, where only 12 out of 50 runs result in the (d)-(f) structures. This procedure also verifies that the structure proposed by Li *et al.*[12] is indeed the lowest energy structure, with a GPAW surface energy 3.2 meV $Å^{-2}$ lower than that of the (g) structure.

Finally, we also re-optimized the three low-energy 5×1 structures with VASP which also identified (g) as the energy minimum, with (h) and (i) having surface energies only around 1 meV $Å^{-2}$ higher.

Constructing transferable NNFFs, such that the model can be trained on smaller structures and applied to larger structures or different systems and system sizes is a prominent topic[50,51] and also part of the original vision behind introducing atomic-descriptor-based NNFFs.[52] Although a conventional narrative has developed that the performance of ML models is good when they are interpolating between points of the training set and poor when they are used to extrapolate, that informal line of reasoning has not survived rigorous examination. It has been

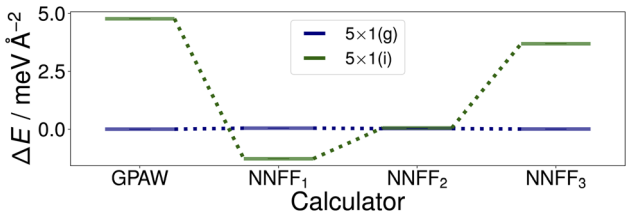


Fig. 7 Surface energies of the (g) and (i) structures, Fig. 5, identified by the first two sets of evolution runs performed on the $SrTiO_3$(110) 5×1 surface. The energies of the optimized structures GPAW structures (first column) were evaluated with the three trained models. The energies shown are offset to that of structure (g).

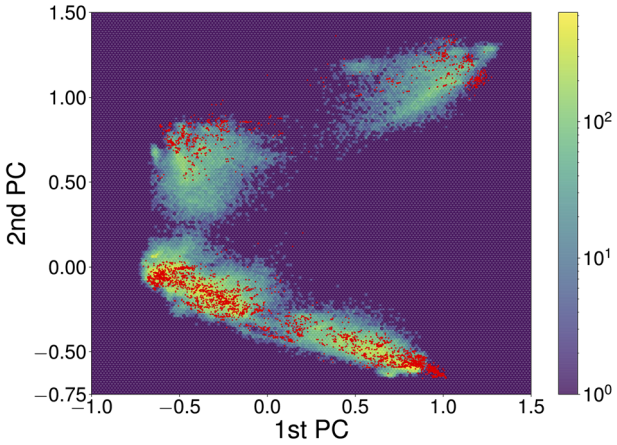


Fig. 8 The first two principal components of descriptors of local atomic environments within $SrTiO_3$(110) 4×1 training data. The projections of the training data and "Opt" test set are depicted as the hexagonally binned background and red dots, respectively. The principal component analysis (PCA) was performed utilizing the scikit-learn library.[54] To correct for symmetry and recurring environments, the PCA was limited to local environments on one side of each slab, excluding the environments of atoms within fixed bulk-like layers.

 

shown that in the high-dimensional spaces commonly explored by ML the probability of a new point falling in the convex hull of the training set is negligible and extrapolation happens constantly.[53] Whether those new points will be correctly described is therefore more closely related to the functional form and flexibility of the model and the diversity of the data used for training. The prediction of the NNFF can still be expected to perform best for structures with a degree of similarity to the training structures. In our specific case, the good performance of the original 4×1 NNFF, $NNFF_1$, can thus be attributed to the good coverage of the space of atomic descriptors afforded by the training structures. To illustrate this, Fig. 8 shows the first two principal components of the descriptors for the 4×1 training data and the same projections of the descriptors for the "Opt" 5×1 test set. The "Opt" data points do fall within or close to the area covered by the local atomic environments represented in the 4×1 training data.

## 4 Summary and conclusions

We successfully combine the covariance matrix adaptation evolution strategy (CMA-ES) and a fully automatically differentiable, high-dimensional neural-network force field (NNFF) to explore the phase diagram of $TiO_x$ overlayer structures on $SrTiO_3$(110) 3×1, 4×1 and 5×1 surfaces. This allows exploiting the stochastic nature of the CMA-ES to perform sets of evolution runs. Thereby the potential energy surfaces (PES) of the systems of interest can be explored more efficiently and to a greater degree than would be possible with DFT. We have thereby arrived at a diverse collection of local energy minima, both reproducing known structures and proposing new stable candidates.

The transferability of the neural-network force field trained solely on the 4×1 surface unit cell structures is demonstrated by comparing to results obtained with neural networks trained on structures from the 5×1 surface unit cell. We attribute the transferability to the diversity of structures generated by the CMA-ES search procedure.

The presented method proves to be capable of thoroughly and efficiently exploring configuration space while greatly reducing the needed computation time. Our formulation of the CMA-ES in a machine-learning context naturally suggests possible extensions to even more complex energy landscapes. Machine-learning models can be used to relax the assumption of Gaussian distribution of candidates and to include intrinsic metrics of the quality of the surrogate model.

## Code availability

A compatible version of NeuralIL, including example scripts for training and evaluation are available.[55]

## Data availability

A dataset containing the CMA-ES founder structures, the overlayers shown in Fig. 5, the trained models and all associated training, validation and test data, is available on Zenodo.[32]

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We thank Michele Riva and Alexander Imre for many helpful discussions, Zhiming Wang for providing us with the STM image used in Fig. 1 of the ESI† and an anonymous referee for their insights and suggestions. This work was supported by the Austrian Science Fund (FWF) (SFB F81 TACO).

## References

1 K. Takayanagi, Y. Tanishiro, M. Takahashi and S. Takahashi, *J. Vac. Sci. Technol. A*, 1985, **3**, 1502–1506.
2 J. A. Enterkin, A. K. Subramanian, B. C. Russell, M. R. Castell, K. R. Poeppelmeier and L. D. Marks, *Nat. Mater.*, 2010, **9**, 245–248.
3 M. Riva, M. Kubicek, X. Hao, G. Franceschi, S. Gerhold, M. Schmid, H. Hutter, J. Fleig, C. Franchini, B. Yildiz and U. Diebold, *Nat. Commun.*, 2018, **9**, 3710.
4 G. Franceschi, M. Schmid, U. Diebold and M. Riva, *J. Mater. Chem. A*, 2020, **8**, 22947–22961.
5 B. C. Russell and M. R. Castell, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2008, **77**, 245414.
6 E. Breckenfeld, R. Wilson, J. Karthik, A. R. Damodaran, D. G. Cahill and L. W. Martin, *Chem. Mater.*, 2012, **24**, 331–337.
7 Z. Zhong and P. Hansmann, *Phys. Rev. X*, 2017, **7**, 011023.
8 A. Chikina, F. Lechermann, M.-A. Husanu, M. Caputo, C. Cancellieri, X. Wang, T. Schmitt, M. Radovic and V. N. Strocov, *ACS Nano*, 2018, **12**, 7927–7935.
9 S. Zeng, X. Yin, T. Herng, K. Han, Z. Huang, L. Zhang, C. Li, W. Zhou, D. Wan, P. Yang, J. Ding, A. Wee, J. Coey, T. Venkatesan, A. Rusydi and A. Ariando, *Phys. Rev. Lett.*, 2018, **121**, 146802.
10 R. He, H. Wu, L. Zhang, X. Wang, F. Fu, S. Liu and Z. Zhong, *Phys. Rev. B*, 2022, **105**, 064104.
11 M. Riva, G. Franceschi, Q. Lu, M. Schmid, B. Yildiz and U. Diebold, *Phys. Rev. Mater.*, 2019, **3**, 043802.
12 F. Li, Z. Wang, S. Meng, Y. Sun, J. Yang, Q. Guo and J. Guo, *Phys. Rev. Lett.*, 2011, **107**, 036103.
13 Z. Wang, F. Li, S. Meng, J. Zhang, E. Plummer, U. Diebold and J. Guo, *Phys. Rev. Lett.*, 2013, **111**, 056101.
14 Z. Wang, H. Xianfeng, S. Gerhold, M. Schmid, C. Franchini and U. Diebold, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2014, **90**, 035436.
15 H.-G. Beyer and H.-P. Schwefel, *Nat. Comput.*, 2002, **1**, 3–52.
16 E. S. Henault, M. H. Rasmussen and J. H. Jensen, *PeerJ Phys. Chem.*, 2020, **2**, e11.
17 T. S. Bush, C. R. A. Catlow and P. D. Battle, *J. Mater. Chem.*, 1995, **5**, 1269–1272.
18 A. R. Oganov and C. W. Glass, *J. Chem. Phys.*, 2006, **124**, 244704.
19 A. R. Oganov, A. O. Lyakhov and M. Valle, *Acc. Chem. Res.*, 2011, **44**, 227–237.

20 L. B. Vilhelmsen and B. Hammer, *Phys. Rev. Lett.*, 2012, **108**, 126101.
21 L. B. Vilhelmsen and B. Hammer, *J. Chem. Phys.*, 2014, **141**, 044711.
22 F. Curtis, X. Li, T. Rose, A. Vazquez-Mayagoitia, S. Bhattacharya, L. M. Ghiringhelli and N. Marom, *J. Chem. Theory Comput.*, 2018, **14**, 2246–2264.
23 L. R. Merte, M. S. Jørgensen, K. Pussi, J. Gustafson, M. Shipilin, A. Schaefer, C. Zhang, J. Rawle, C. Nicklin, G. Thornton, R. Lindsay, B. Hammer and E. Lundgren, *Phys. Rev. Lett.*, 2017, **119**, 096102.
24 O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko and K.-R. Müller, *Chem. Rev.*, 2021, **121**, 10142–10186.
25 A. R. Oganov, C. J. Pickard, Q. Zhu and R. J. Needs, *Nat. Rev. Mater.*, 2019, **4**, 331–348.
26 M. Arrigoni and G. K. H. Madsen, *npj Comput. Mater.*, 2021, **7**, 1–13.
27 J. Timmermann, F. Kraushofer, N. Resch, P. Li, Y. Wang, Z. Mao, M. Riva, Y. Lee, C. Staacke, M. Schmid, C. Scheurer, G. S. Parkinson, U. Diebold and K. Reuter, *Phys. Rev. Lett.*, 2020, **125**, 206101.
28 L. R. Merte, M. K. Bisbo, I. Sokolovicć, M. Setvin, B. Hagman, M. Shipilin, M. Schmid, U. Diebold, E. Lundgren and B. Hammer, *Angew. Chem. Int. Ed.*, 2022, **61**, e202204244; *Angew. Chem.*, 2022, **134**, e202204244.
29 L. Hirschfeld, K. Swanson, K. Yang, R. Barzilay and C. W. Coley, *J. Chem. Inf. Model.*, 2020, **60**, 3770–3780.
30 N. Hansen and A. Ostermeier, *Evol. Comput.*, 2001, **9**, 159–195.
31 H. Montes-Campos, J. Carrete, S. Bichelmaier, L. M. Varela and G. K. H. Madsen, *J. Chem. Inf. Model.*, 2022, **62**, 88–101.
32 R. Wanzenböck, M. Arrigoni, S. Bichelmaier, F. Buchner, J. Carrete and G. K. H. Madsen, *Neural-network-backed evolutionary search for SrTiO3(110) surface reconstructions*, 2022, DOI: **10.5281/zenodo.6782465**.
33 N. Hansen, 2016, arXiv:1604.00772 [cs.LG]oref.
34 J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Dułak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. A. Hansen, H. H. Kristoffersen, M. Kuisma, A. H. Larsen, L. Lehtovaara, M. Ljungberg, O. Lopez-Acevedo, P. G. Moses, J. Ojanen, T. Olsen, V. Petzold, N. A. Romero, J. Stausholm-Møller, M. Strange, G. A. Tritsaris, M. Vanin, M. Walter, B. Hammer, H. Häkkinen, G. K. H. Madsen, R. M. Nieminen, J. K. Nørskov, M. Puska, T. T. Rantala, J. Schiøtz, K. S. Thygesen and K. W. Jacobsen, *J. Phys.: Condens. Matter*, 2010, **22**, 253202.
35 A. H. Larsen, M. Vanin, J. J. Mortensen, K. S. Thygesen and K. W. Jacobsen, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2009, **80**, 195112.
36 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
37 G. Kresse and J. Furthmüller, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 11169–11186.
38 E. Bitzek, P. Koskinen, F. Gähler, M. Moseler and P. Gumbsch, *Phys. Rev. Lett.*, 2006, **97**, 170201.
39 A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng and K. W. Jacobsen, *J. Phys.: Condens. Matter*, 2017, **29**, 273002.
40 T. A. Manz and N. G. Limas, *Chargemol program for performing DDEC analysis*, Version 3.15, 2017, ddec.sourceforge.net.
41 T. A. Manz and N. G. Limas, *RSC Adv.*, 2016, **6**, 47771–47801.
42 N. G. Limas and T. A. Manz, *RSC Adv.*, 2016, **6**, 45727–45747.
43 J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne and Q. Zhang, *JAX: composable transformations of Python+NumPy programs*, 2018, **http://github.com/google/jax**.
44 A. P. Bartók, R. Kondor and G. Csányi, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **87**, 184115.
45 E. Kocer, J. Mason and H. Erturk, *AIP Adv.*, 2020, **10**, 015021.
46 S. Gupta, T. Kanchinadam, D. Conathan and G. Fung, *Front. Appl. Math. Stat.*, 2020, **5**, 67.
47 L. N. Smith, 2018, arXiv:1803.09820 [cs.LG]oref.
48 Q. Wang, Y. Ma, K. Zhao and Y. Tian, *Ann. Data Sci.*, 2022, **9**, 187–212.
49 J. Tersoff and D. R. Hamann, *Phys. Rev. Lett.*, 1983, **50**, 1998–2001.
50 C. Mangold, S. Chen, G. Barbalinardo, J. Behler, P. Pochet, K. Termentzidis, Y. Han, L. Chaput, D. Lacroix and D. Donadio, *J. Appl. Phys.*, 2020, **127**, 244901.
51 Y. Shaidu, E. Kücükbenli, R. Lot, F. Pellegrini, E. Kaxiras and S. de Gironcoli, *npj Comput. Mater.*, 2021, **7**, 1–13.
52 J. Behler and M. Parrinello, *Phys. Rev. Lett.*, 2007, **98**, 146401.
53 R. Balestriero, J. Pesenti and Y. LeCun, 2021, oref.
54 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, *JMLR*, 2011, **12**, 2825–2830.
55 R. Wanzenböck, M. Arrigoni, S. Bichelmaier, F. Buchner, J. Carrete and G. K. H. Madsen, *Madsen-s-research-group/neuralil-public-releases: CMA-SrTiO3-surfaces*, 2022, DOI: **10.5281/zenodo.6779823**.

 