Universal interatomic potential for perovskite oxides

Jing Wu (武静),[1,*] Jiyuan Yang (杨季元),[1,*] Yuan-Jinsheng Liu (刘袁今生),[1,*] Duo Zhang (张铎)[2,3] Yudi Yang (杨雨迪),[1] Yuzhi Zhang (张与之),[2,3] Linfeng Zhang (张林峰),[2,3] and Shi Liu (刘仕)[1,4,†]

[1]*Department of Physics, Key Laboratory for Quantum Materials of Zhejiang Province, School of Science and Research Center for Industries of the Future, Westlake University, Hangzhou, Zhejiang 310030, China*
[2]*DP Technology, Beijing 100080, China*
[3]*AI for Science Institute, Beijing 100080, China*
[4]*Institute of Natural Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang 310024, China*

With their celebrated structural and chemical flexibility, perovskite oxides have served as a highly adaptable material platform for exploring emergent phenomena arising from the interplay between different degrees of freedom. Molecular dynamics (MD) simulations leveraging classical force fields, commonly depicted as parametrized analytical functions, have made significant contributions in elucidating the atomistic dynamics and structural properties of crystalline solids including perovskite oxides. However, the force fields currently available for solids are rather specific and offer limited transferability, making it time-consuming to use MD to study new materials systems since a new force field must be parametrized and tested first. The lack of a generalized force field applicable to a broad spectrum of solid materials hinders the facile deployment of MD in computer-aided materials discovery (CAMD). Here, by utilizing a deep-neural network with a self-attention scheme, we have developed a unified force field (UniPero) that enables MD simulations of perovskite oxides involving 14 metal elements and conceivably their solid solutions with arbitrary compositions. Notably, isobaric-isothermal ensemble MD simulations with this model potential accurately predict the experimental temperature-driven phase transition sequences for several markedly different ferroelectric oxides, including a six-element ternary solid solution $Pb(In_{1/2}Nb_{1/2})O_3$–$Pb(Mg_{1/3}Nb_{2/3})O_3$–$PbTiO_3$. We believe the universal interatomic potential along with the training database, proposed regression tests, and the auto-testing workflow, all released publicly, will pave the way for a systematic improvement and extension of a unified force field for solids, potentially heralding a new era in CAMD.

Classical force fields, including CHARMM [1], AMBER [2], and GROMOS [3], which are capable of accurately describing the interatomic interactions within key biological molecules such as amino acids and nucleic acids, have played a pivotal role in expanding the use of molecular dynamics (MD) simulations in computer-aided drug design (CADD). These force fields enabled large-scale simulations of atomic and molecular movements in biological systems with remarkable efficiency and precision, greatly facilitating detailed studies of drug-receptor interactions [4], protein conformational changes [5,6], binding affinities [7], and the search for lead compounds [8]; they are now indispensable tools in the pursuit of new drugs and therapeutics [9,10]. In contrast, although MD simulations have been successful in elucidating the properties of various types of solids by offering atomic-level insights, their integration into computer-aided materials discovery (CAMD) has not yet achieved the same level of industry-wide adoption as in CADD. This disparity is primarily due to the lack of a generalized force field applicable to a wide range of solid materials.

Many force fields for crystalline solids are tailored for a specific material or a narrow group of materials, covering a limited chemical space [11–18]. While the reactive force-field (ReaxFF) interatomic potential allows for element-specific parametrization and is commonly viewed as possessing strong transferability [19], its developers still caution against arbitrarily combining parameter sets in anticipation of consistent predictive power [20]. Moreover, the usage of a highly sophisticated energy function containing a large number of empirical parameters (as seen in ReaxFF) makes it challenging to parametrize force fields for new materials. In this work, taking ABO_3-type perovskite oxides as an example, we demonstrate that a deep neural network-based model, namely deep potential (DP) [21] with a novel attention mechanism, referred to as DPA-1 [22], exhibits the required representability and transferability for the development of a universal interatomic potential. By adopting the strategy of "modular development of deep potential" (ModDP) [23] and relying solely on *ab initio* training data, we have achieved a unified force field that enables MD simulations for diverse perovskite oxides comprising 14 different metal elements and is conceivably adaptable for their solid solutions of any composition. For instance, we report the first large-scale MD simulations of a six-element ternary

---

solid solution $Pb(In_{1/2}Nb_{1/2})O_3$–$Pb(Mg_{1/3}Nb_{2/3})O_3$–$PbTiO_3$ (PIN-PMN-PT), which is a relaxor ferroelectric possessing superior piezoelectric properties [24,25].

The ABO_3-type perovskite oxide is chosen as the model system because of its role as an exceptionally versatile material platform. The adaptability in chemical, structural, and compositional variations grants a unique capability to fine-tune interactions such as electron-phonon coupling, crystal field, and spin-orbit interaction, the competitions among which often dictate the functional properties of perovskite oxides [26,27]. Given these intricacies, it becomes challenging, if not outright impractical, to rely on a force field with a fixed analytical structure and a limited set of parameters to create a generalized force field for perovskite oxides spanning a broad chemical and compositional range. To address this issue, we adopt the DP model and its extension augmented with a self-attention scheme.

The DP model, based on a deep neural network with the number of learnable parameters on the order of 10^6, offers a robust mathematical structure to represent highly nonlinear and complex interatomic interactions while bypassing the need to handcraft descriptors that represent local atomic environments [21]. Specifically, the DP model features a symmetry-preserving embedding network that maps an atom's local environment to inputs for a fitting neural network which then outputs the atomic energy; the sum of atomic energies yields the total energy. This approach is a manifestation of the embedded atom concept, capturing the many-body character of interatomic interactions. Therefore, the DP model trained using density functional theory (DFT) total energies implicitly accounts for various types of interactions including the long-range interactions. More recently, the DPA-1 model [22] has introduced an element-type embedding net and integrated a self-attention mechanism. This mechanism excels in modeling the importance of neighboring atoms and reweighting the interactions among them based on both distance and angular information, thus allowing for the mixing and communication within the latent space of the embedding parameters of elements and structures. As a result, the DPA-1 model achieves satisfactory transferability to unseen systems with varied elemental compositions [22], fulfilling the essential criteria for constructing a universal force field applicable to solids comprising multiple elements.

To develop a DPA-1 model capable of representing diverse element types and compositions that the perovskite structure can support, the construction of a comprehensive training database becomes pivotal. For simple materials systems comprising two or three elements, the DP-GEN scheme [28,29] for data generation is streamlined and largely autonomous, requiring little human involvement. This concurrent learning procedure iteratively explores the configuration space (by running MD simulations) using one of the four models trained with existing data. It then labels only those MD-sampled configurations that exhibit high uncertainty levels, determined by the maximum standard deviation of predictions from the four models. The energies and forces of these newly labeled configurations, computed using first-principles DFT, are subsequently incorporated into the training database for the next learning cycle. The iteration stops when all configurations sampled from MD simulations meet a
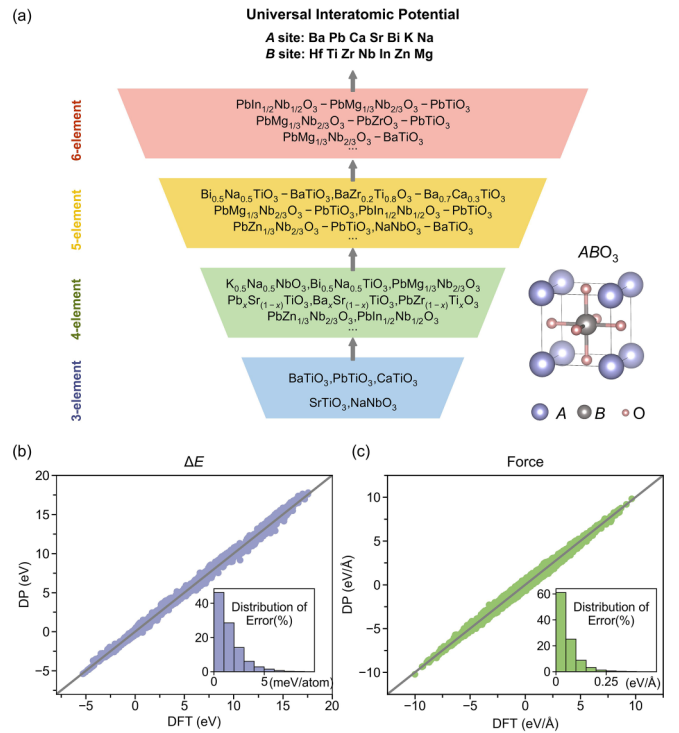

FIG. 1. (a) Workflow for the development of a universal force field of perovskite oxides following the ModDP protocol. Comparison of (b) relative energies ΔE and (c) atomic forces predicted using the DPA-1 model with reference DFT results for $\approx$20 000 configurations in the final training database. The inset shows the distribution of the absolute error.

certain standard of accuracy across all four models. More details regarding DP-GEN can be found in previous studies [29–31]. However, the robustness of the DP-GEN scheme is limited when constructing the training database for complex oxides with diverse element types and compositions.

In previous work, we demonstrated that the ModDP protocol facilitates a systematic development and improvement of DP models for complex oxide solid solutions [23]. At its core, ModDP embodies the concepts of "data reusing" and "divide and conquer": the converged training database associated with an end-member material is treated as an independent module, and is then reused to train the DP model of solid solutions derived from these end members. Following the same spirit, we devise a procedure that progressively introduces perovskites with increasing complexity (number of element types) to enhance the capability of DPA-1 level by level. The workflow of the force field development procedure that leads to a universal interatomic potential is depicted in Fig. 1(a). The initial database contains $\approx$1000 configurations of 26 different types of perovskites, involving $\approx$200 compositions and 14 metal elements. It is noted that some data for $PbTiO_3$ and $SrTiO_3$ were taken from a published database [23], and we did *not* intentionally design the initial database, aiming to minimize human intervention. We first use the standard DP-GEN scheme to converge a DPA-1 model suited for three-element (including oxygen) perovskite oxides such as $PbTiO_3$, $BaTiO_3$, $CaTiO_3$, $SrTiO_3$, and $NaNbO_3$. Subsequently, the converged training database

serves as the starting point for DP-GEN to improve DPA-1 for four-element perovskite systems like $PbZr_{1-x}Ti_xO_3$ and $Pb(Mg_{1/3}Nb_{2/3})O_3$. Ultimately, we achieve a converged DPA-1 model for six-element perovskite systems, including the ternary solid solution of PIN-PMN-PT. As detailed below, the final DPA-1 model serves as a universal interatomic potential (named as UniPero), capable of modeling a wide variety of perovskites with MD simulations.

All DFT calculations are performed with Atomic-orbital Based *Ab-initio* Computation at UStc (ABACUS) package [32,33] using numerical atomic orbitals (NAOs) and SG15-optimized norm-conserving Vanderbilt (ONCV) pseudopotentials [34]. The Perdew-Burke-Ernzerhof revised for solids (PBEsol) [35] within the generalized gradient approximation is chosen as the exchange-correlation functional, and the double-ζ plus polarization functions (DZP) with a plane-wave cutoff energy of 100 Ry is employed as the NAO basis set. The energies and forces are computed with a $0.1\,\text{Bohr}^{-1}$ k-point spacing and an energy convergence threshold of 1.0×10^{-7} Ry. All isobaric-isothermal (NPT) ensemble MD simulations are performed using LAMMPS [36], with a time step of 2 fs and the temperature controlled via the Nosé-Hoover thermostat and the pressure controlled by the Parrinello-Rahman barostat. When simulating temperature-driven phase transitions, the perovskite systems are modeled with a $10 \times 10 \times 10$ supercell (assuming a five-atom unit cell) consisting of 5000 atoms, except for PIN-PMN-PT which is modeled with a $20 \times 20 \times 20$ supercell consisting of 40 000 atoms. The smooth edition of the DP model (with a cutoff radius of 6 Å and the smoothing starting at 0.5 Å) with an attention scheme DPA-1 is adopted. The DEEPMD-KIT package [37] is used in the training cycle of DP-GEN. The basic networks of DPA-1 comprise the embedding and fitting networks, both of which utilize ResNet architectures with default settings. Additionally, DPA-1 also incorporates a one-hot type embedding network with a size of 8, as well as a two-layer attention mechanism possessing 128 hidden dimensions during the scaled-dot product attention computation. To enhance reproducibility, we have made our final training database and hyperparameters accessible via a public repository [38].

The final training database contains 19 288 configurations. We note that DPA-1 aims to reproduce DFT total energies, which inherently vary significantly between different materials. For configurations in the database, the DFT total energies span a range of an astonishing $\approx 400\,\text{eV/atom}$ (including an element-specific constant energy shift). Impressively, the DPA-1 model has a mean absolute error (MAE) of 1.75 meV/atom for the total energy, highlighting its remarkable representability. To better illustrate the fitting performance, we compare the energy differences (ΔE) predicted by DPA-1 with DFT values. For a specific composition, a configuration is (arbitrarily) chosen as the reference; the ΔE values are then computed relative to the total energy of this chosen reference configuration. As shown in Fig. 1(b), DPA-1 well reproduces DFT values of ΔE. The predictions from DPA-1 for atomic force are in satisfactory agreement with DFT results as well, achieving a MAE of 0.054 eV/Å [Fig. 1(c)].

Employing a concurrent learning approach like DP-GEN allows the training database to be updated efficiently and iteratively. As a result, we do not have the typical validation

or testing datasets commonly associated with conventional supervised learning. We propose two sets of tasks to evaluate DPA-1's applicability for perovskite oxides. For the first set of testing jobs (denoted as Task I), *ab initio* molecular dynamics (AIMD) simulations are performed at three specific temperatures (300, 450, 900 K) for a few picoseconds. DPA-1 is then employed to evaluate the energies for configurations along these AIMD trajectories. The objective is to ascertain DPA-1's accuracy in predicting energies for configurations sampled by AIMD at finite temperatures. As illustrated in Fig. 2(a) which uses PIN-PMN-PT as an example, the energy evolution trajectory obtained with DPA-1 closely follows that from AIMD at three different temperatures, confirmed by Pearson correlation coefficients of 0.93, 0.95, and 0.99, respectively. The second set of evaluations (Task II) is to generate classical DPMD trajectories using DPA-1 at varying temperatures, the energies of configurations along which are subsequently determined with DFT. This task is designed to validate whether classical MD trajectories from DPA-1 properly sample the relevant configuration spaces at finite temperatures. The applicability of DPA-1 is gauged by the MAEs of energies and atomic forces for all the configurations generated from these two sets of tasks. We have selected 10 representative perovskite oxides for testing: $BaTiO_3$, $PbTiO_3$, $SrTiO_3$, $Ba_{0.5}Sr_{0.5}TiO_3$, $K_{0.5}Na_{0.5}NbO_3$, $NaNbO_3$–$BaTiO_3$, $PbZr_{0.5}Ti_{0.5}O_3$, $BaZr_{0.2}Ti_{0.8}O_3$–$Ba_{0.3}Ca_{0.7}TiO_3$, PMN–PT, and PIN–PMN–PT. It is evident from Fig. 2(c) (for Task I) and Fig. 2(d) (for Task II) that the DPA-1 model well reproduces the energy evolution from AIMD and performs appropriate sampling at finite temperatures for markedly different perovskite systems. The MAE for energy is lower than 1.45 meV/atom, and that for atomic force is below 0.052 eV/Å.

With DPA-1's capability to faithfully capture the complex and highly nonlinear potential energy surfaces involving diverse element types, we expect developing a DPA-1 model with further enhanced transferability and accuracy can be achieved by introducing more data from new materials systems into the current training database. Therefore, the proposed tasks, which are readily automatable, serve as suitable *regression tests* (a common practice in software development) for the continuous improvement of the DPA-1 model. We have implemented an auto-testing workflow based on an online platform (Bohrium Notebook) [38] that allows real-time access to our DPA-1 model, cloud-based computing resources, and data analysis scripts. This initiative aims to streamline force field development and foster collaborative contributions from the community.

A definitive test of the applicability of the developed DPA-1-based universal interatomic potential is to examine temperature-driven phase transitions in ferroelectric perovskites. As summarized in Fig. 3, a single DPA-1 model, without any tuning, successfully reproduces the experimental sequences of phase transitions for lead-based ferroelectrics ($PbTiO_3$, $Pb_{0.5}Sr_{0.5}TiO_3$, and $PbZr_{0.5}Ti_{0.5}O_3$), lead-free ferroelectrics ($BaTiO_3$, $KNbO_3$, and $K_{0.5}Na_{0.5}NbO_3$), quantum paraelectric $SrTiO_3$, and tertiary solid solutions (0.29PIN–0.45PMN–0.26PT and 0.36PIN–0.36PMN–0.28PT). For example, the temperature-dependent lattice constants obtained from NPT MD simulations reveal a tetragonal (T, space group $P4mm$) to cubic (C, space group $Pm\bar{3}m$) transition at
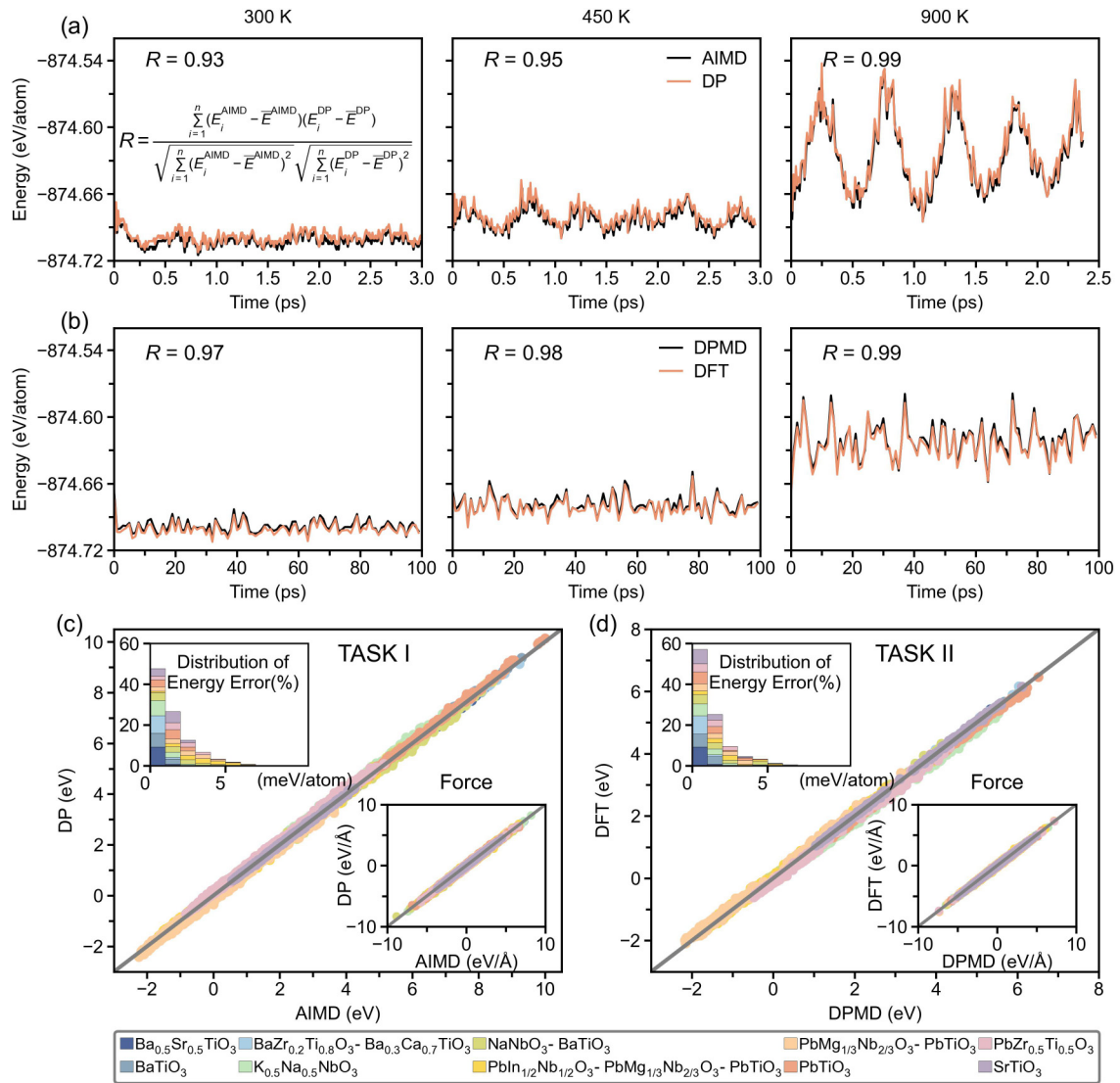
FIG. 2. Proposed tasks to evaluate the applicability of UniPero using PIN–PMN–PT as an example. (a) Task I. The energies of configurations sampled by AIMD at three different temperatures are checked by DPA-1. (b) Task II. The energies of configurations sampled by classical MD with DPA-1 are checked by DFT. Comparison of relative energies and atomic forces for configurations of 10 representative perovskite oxides generated from (c) Task I and (d) Task II.

$\approx$550 K in PbTiO$_3$ [Fig. 3(a)], whereas the transition temperature T_c drops to $\approx$280 K in the A-site doped solid solution, Pb$_{0.5}$Sr$_{0.5}$TiO$_3$ [Fig. 3(d)]. In comparison, MD simulations using DPA-1 predict that the B-site doped solid solution PbZr$_{0.5}$Ti$_{0.5}$O$_3$ adopts a rhombohedral phase (R, space group $R3m$) at low temperatures, and it undergoes rhombohedral-tetragonal-cubic (R-T-C) phase transitions as the temperature increases [Fig. 3(g)], consistent with experimental observations [39,40]. SrTiO$_3$ is a well-known quantum paraelectric that exhibits a temperature-driven tetragonal-cubic transition ($I4/mcm \rightarrow Pm\bar{3}m$) characterized by a decrease in the TiO$_6$ octahedral tilt angle. This phase transition, corresponding to a tiny energy barrier of $\approx$1 meV/atom, is reproduced by DPA-1 [Fig. 3(b)], and the theoretical value of T_c is 150 K, which is comparable to the experimental value of 105 K [41]. The same model also correctly predicts the phase transition sequence, rhombohedral-orthorhombic-tetragonal-cubic (R-O-T-C), in BaTiO$_3$ and K$_{0.5}$Na$_{0.5}$NbO$_3$ [see Figs. 3(c)

and 3(h)]. Moreover, even for the highly complex six-element tertiary solid solutions such as PIN-PMN-PT, DPA-1 not only reproduces the T–C phase transition observed in experiments, but also predicts the correct trend in the change of T_c due to the compositional variation. As shown in Figs. 3(f) and 3(i), the T_c rises by about 50 K from 0.29PIN–0.45PMN–0.26PT to 0.36PIN–0.36PMN–0.28PT, similar to the T_c change of $\approx$60 K found in experiments due to the increased percentage of PIN from 29% to 36% [42,43]. It is observed that theoretical T_C values for ferroelectric transitions are typically lower than experimental values by $100 - 200$ K. Such discrepancies appear to be a common feature for force fields developed by fitting to results computed with the PBEsol density functional [23,44,45]. Given the high fidelity of DPA-1 as demonstrated by its excellent fitting to DFT results, the underestimation of T_C is more likely a reflection of PBEsol's limitations in forecasting finite-temperature properties rather than an issue with DPA-1's precision.
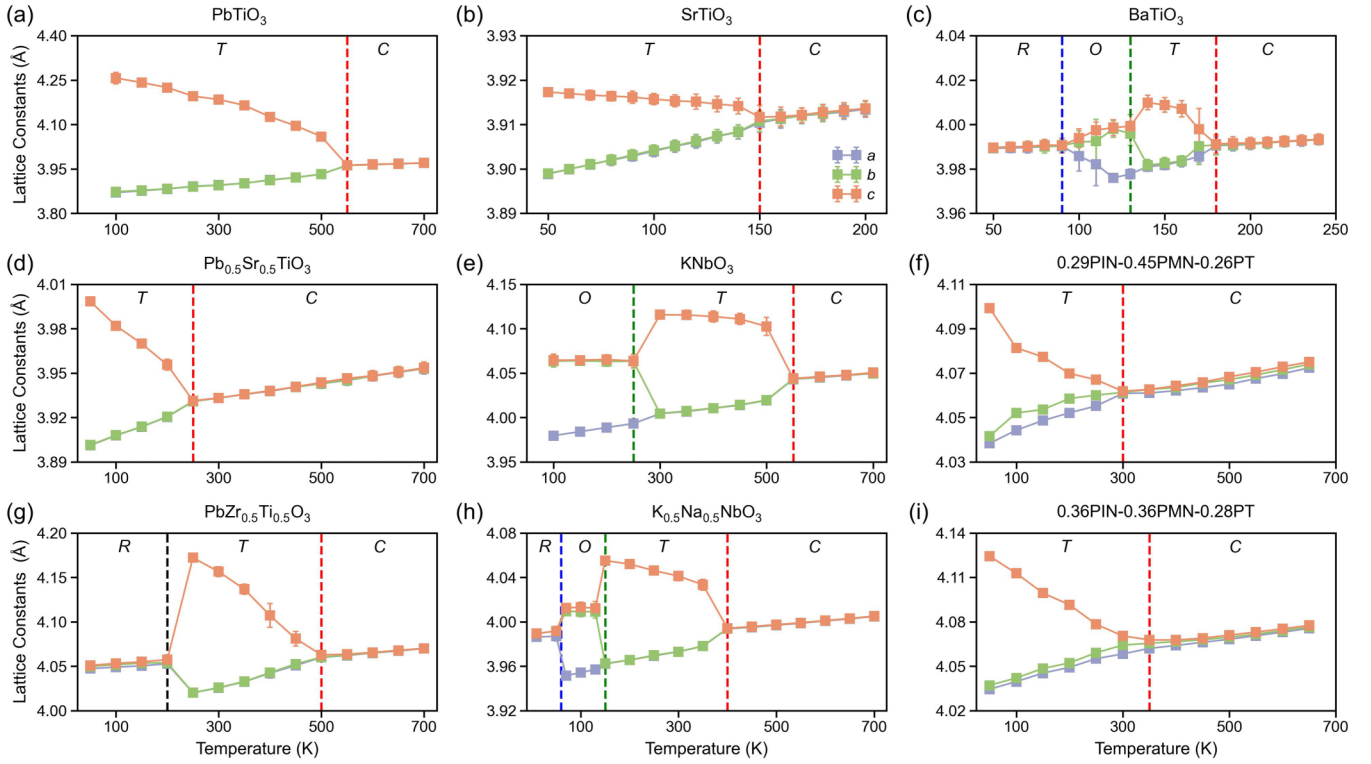
FIG. 3. Temperature-dependent lattice constants simulated with *NPT* DPMD using the universal interatomic potential for (a) $PbTiO_3$, (b) $SrTiO_3$, (c) $BaTiO_3$, (d) $Pb_{0.5}Sr_{0.5}TiO_3$, (e) $KNbO_3$, (f) 0.29PIN–0.45PMN–0.26PT, (g) $PbZr_{0.5}Ti_{0.5}O_3$, (h) $K_{0.5}Na_{0.5}NbO_3$, (i) 0.36PIN–0.36PMN–0.28PT.

The universal interatomic potential developed here can serve as a readily deployable tool to reveal the local structures and lattice dynamics of perovskite oxides with atomic spatial resolution. Using three 50-ps equilibrium DPMD trajectories for 0.29PIN–0.45PMN–0.26PT, each with different *B*-site cation arrangements and modeled with a 40 000-atom supercell, we analyze the distributions of local displacements of cations with respect to the center of their surrounding oxygen cages. For describing inversion-symmetry breaking in ferroelectric materials, polarization is often the natural order parameter. Given a configuration from MD simulations, the polarization for simple perovskite systems such as $PbTiO_3$ can be estimated using the following formula

$$\mathbf{P}^m(t) = \frac{1}{V_{uc}}\left[\frac{1}{8}\mathbf{Z}^*_{Pb}\sum_{i=1}^{8}\mathbf{r}^m_{Pb,i}(t) + \mathbf{Z}^*_{Ti}\mathbf{r}^m_{Ti}(t) + \frac{1}{2}\mathbf{Z}^*_{O}\sum_{i=1}^{6}\mathbf{r}^m_{O,i}(t)\right],$$

where $\mathbf{P}^m(t)$ is the polarization of unit cell m at time t, V_{uc} is the volume of the unit cell, $\mathbf{Z}^*_{Pb}$, $\mathbf{Z}^*_{Ti}$, and $\mathbf{Z}^*_{O}$ are the Born effective charges of Pb, Ti, and O atoms, $\mathbf{r}^m_{Pb,i}(t)$, $\mathbf{r}^m_{Ti,i}(t)$, and $\mathbf{r}^m_{O,i}(t)$ are the instantaneous atomic positions. The polarization is thus the dipole moment per unit volume. However, PIN-PMN-PT is considered a charge-frustrated system where the *B* sites host cations of different charges: Mg^{2+}, In^{3+}, Ti^{4+}, and Nb^{5+}. It becomes problematic to define a local electric dipole for a unit cell that is not locally charge neutral, (e.g., a local $PbInO_3$ unit cell in PIN-PMN-PT solid solutions). Therefore, we use local atomic displacements of cations to measure the degree of inversion-symmetry breaking at the unit cell level.

In Figs. 4(a)–4(c), the probability density distributions of cation displacements (d_x, d_y, d_z) along the Cartesian axes are shown as a function of temperature. At a low temperature of 200 K, the distributions of d_x and d_y are mostly Gaussian-like and centered around zero for all cations. In contrast, the peak position for d_z of Pb deviates the most away from zero, followed by Ti, Nb, In, and Mg in sequence. This deviation in the d_z distribution, compared to d_x and d_y, aligns with the global tetragonal symmetry indicated by the lattice constants. Importantly, these distributions highlight a distinct difference in the local structures between relaxor ferroelectric PIN–PMN–PT and prototypical ferroelectrics like $PbTiO_3$. In PIN–PMN–PT, the distributions of cation displacements are more dispersive and isotropic, spanning both positive and negative values along all three Cartesian directions. Conversely, both Pb and Ti cations in ferroelectric $PbTiO_3$ displace along the same polar direction (e.g., [001]) [44]. At an elevated temperature of 300 K, the peak positions of the d_z distributions for all cations shift toward values lower than those at 200 K. When temperature rises above T_C, distributions become Gaussian, all centering at zero.

Snapshots of instantaneous cation displacements in configurations sampled from MD are displayed in Figs. 4(d)–4(e). We find that the magnitude of local Pb displacements exhibits less temperature sensitivity compared to the *B*-site cations. Even at 400 K which is above the theoretical T_C, many Pb atoms remain locally displaced; their displacements grow more chaotic, manifesting diminished long-range order. The displacements of *B*-site cations, however, display more vulnerability: some are already nonpolar at 200 K, and an
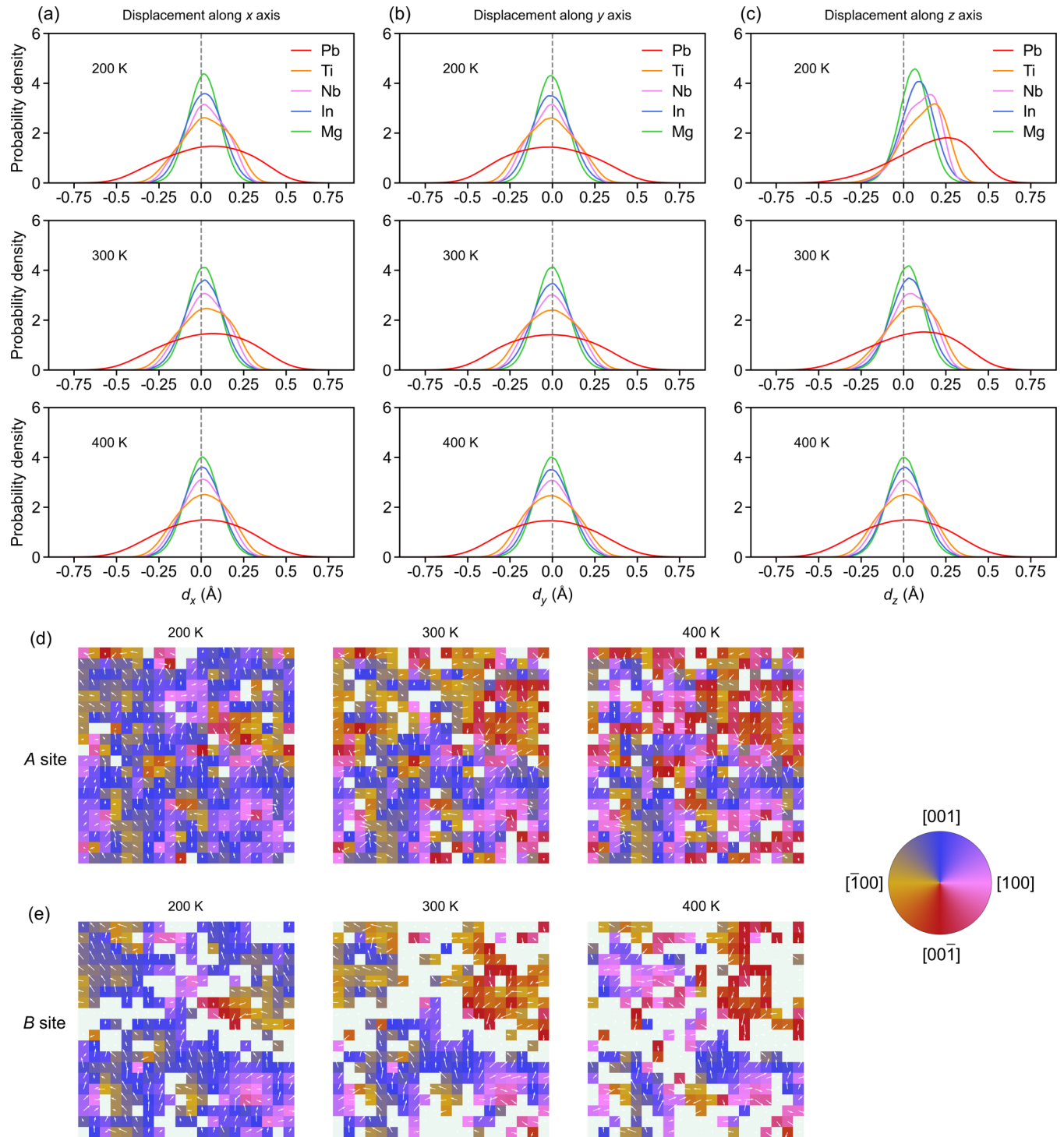
FIG. 4. Probability density distributions of cation displacements along the (a) *x* axis, (b) *y* axis, (c) *z* axis as temperature increases in 0.29PIN–0.45PMN–0.26PT modeled with a 40 000-atom supercell. Snapshots of (d) *A*-site and (e) *B*-site cation displacements at different temperatures obtained from MD simulations. Each white arrow indicates the local atomic displacement within a unit cell, with its background color representing the direction. Note that if the magnitude of the local displacement is less than 0.1 Å, the background is colored white.

increasing percentage becomes nonpolar as temperature increases. Notably, we do not observe any polar nanoregions (PNRs) embedded in a nonpolar matrix, as the majority of Pb atoms are displaced, in agreement with previous MD simulations of PMN–PT [46]. The markedly different behaviors between Pb and *B*-site cations could be important for

understanding the piezoelectric response of lead-based relaxor ferroelectrics. These findings underscore the quantitative nature of the DPA-1 model.

We briefly comment on the main limitations and challenges of machine-learning force fields like DPA-1, which center around interpretability and computational

efficiency. While the large number of learnable parameters of DPA-1 facilitates the representation of complex interactions, it simultaneously hinders our understanding of these very interactions. In essence, each of these parameters adjusts itself during the learning process to minimize prediction errors, but the collective behavior of such a high number of parameters can be enigmatic. This contrasts with traditional force fields which, through their explicit functional forms, often allow for clearer physical interpretation. Although the DPA-1 model demonstrates a clear speed advantage over first-principles methods like DFT, it still lags behind conventional force fields. For instance, the bond-valence model [44] typically runs two orders of magnitude faster than the DPA-1 model. Solutions to the aforementioned challenges exist. Principal component analysis can be performed to analyze parameters of the DPA-1 model to explore the information within the latent space. For example, in a recent work of DPA-1 [22], the visualizations of the element-related parameters learned by the model coincide with the periodic table, partially demonstrating the model's interpretability. As for computational efficiency, model compression [47] is set to alleviate this problem by using a small number of polynomial coefficients to fit the learned parameters in the model. This approach can significantly reduce the computational cost while maintaining accuracy.

The current state of DPA-1 for perovskite oxides can be viewed as a pre-trained model, which is already sufficiently accurate for many perovskite oxides as discussed above. Drawing parallels from the development history of ChatGPT [48] in natural language processing, such a pre-trained model can be fine-tuned for specific downstream tasks with minimal extra effort using few-shot or zero-shot learning techniques [49]. As the community continues to amass high-quality electronic structure data, it is conceivable that this pre-trained model will progressively expand its coverage, encompassing an ever broader range of chemical and structural spaces.

In summary, using the ABO_3-type perovskite oxide as a model system, our results demonstrate that a deep neural network, augmented with an attention scheme, can deliver the requisite representability and transferability to establish a universal interatomic potential for solid materials spanning diverse element types and compositions, at least for materials systems with a fixed crystal structure. Even with the complex nature of target systems involving 15 elements, the final training database remains impressively concise, containing no more than 20 000 configurations; the associated computational cost is manageable, requiring fewer than 200 000 CPU hours. The DPA-1-based force field, UniPero, developed in this work already exhibits satisfactory accuracy, proving suitable for large-scale MD simulations of distinct ferroelectric perovskites and reproducing the correct phase transition sequences driven by temperature. By publicly releasing the training database and the automated testing workflow, we anticipate our initiative will facilitate a systematic improvement and extension of a generalized force field with an even larger scope (e.g., free surfaces and defects). This, in turn, hopefully could foster the routine and efficient use of molecular dynamics for simulating solid materials, ultimately providing valuable insights for computer-aided materials discovery.

---

[1] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, J. Comput. Chem. **4**, 187 (1983).

[2] D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz Jr, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, The Amber biomolecular simulation programs, J. Comput. Chem. **26**, 1668 (2005).

[3] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, and W. F. van Gunsteren, The GROMOS biomolecular simulation program package, J. Phys. Chem. A **103**, 3596 (1999).

[4] A. C. Pan, D. W. Borhani, R. O. Dror, and D. E. Shaw, Molecular determinants of drug–receptor binding kinetics, Drug Discov. Today **18**, 667 (2013).

[5] M. Karplus and J. A. McCammon, Molecular dynamics simulations of biomolecules, Nat. Struct. Biol. **9**, 646 (2002).

[6] D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan *et al.*, Atomic-level characterization of the structural dynamics of proteins, Science **330**, 341 (2010).

[7] J. Åqvist, V. B. Luzhkov, and B. O. Brandsdal, Ligand binding affinities from MD simulations, Acc. Chem. Res. **35**, 358 (2002).

[8] S. Sakkiah, S. Thangapandian, and K. W. Lee, Pharmacophore modeling, molecular docking, and molecular dynamics simulation approaches for identifying new lead compounds for inhibiting aldose reductase 2, J. Mol. Model. **18**, 3267 (2012).

[9] J. D. Durrant and J. A. McCammon, Molecular dynamics simulations and drug discovery, BMC Biol. **9**, 71 (2011).

[10] J. A. Lemkul and D. R. Bevan, The role of molecular simulations in the development of inhibitors of amyloid β-peptide aggregation for the treatment of Alzheimer's disease, ACS Chem. Neurosci. **3**, 845 (2012).

[11] T. Ohira, T. Inamuro, and T. Adachi, Molecular dynamics simulation of amorphous silicon with tersoff potential, Sol. Energy Mater. Sol. Cells **34**, 565 (1994).

[12] M. Sepliarsky, A. Asthagiri, S. R. Phillpot, M. G. Stachiotti, and R. L. Migoni, Atomic-level simulation of ferroelectricity in oxide materials, Curr. Opin. Solid State Mater. Sci. **9**, 107 (2005).

[13] Y.-H. Shin, V. R. Cooper, I. Grinberg, and A. M. Rappe, Development of a bond-valence molecular-dynamics model for complex oxides, Phys. Rev. B **71**, 054104 (2005).

[14] M. Sepliarsky and R. E. Cohen, First-principles based atomistic modeling of phase stability in PMN–xPT, J. Phys.: Condens. Matter **23**, 435902 (2011).

[15] R. B. Wexler, Y. Qi, and A. M. Rappe, Sr-induced dipole scatter in $Ba_xSr_{1-x}TiO_3$: Insights from a transferable-bond valence-based interatomic potential, Phys. Rev. B **100**, 174109 (2019).

[16] H. W. Sheng, M. J. Kramer, A. Cadien, T. Fujita, and M. W. Chen, Highly optimized embedded-atom-method potentials for fourteen fcc metals, Phys. Rev. B **83**, 134118 (2011).

[17] M. J. Noordhoek, T. Liang, Z. Lu, T.-R. Shan, S. B. Sinnott, and S. R. Phillpot, Charge-optimized many-body (COMB) potential for zirconium, J. Nucl. Mater. **441**, 274 (2013).

[18] R. Ganser, S. Bongarz, A. von Mach, L. Azevedo Antunes, and A. Kersch, Piezo- and pyroelectricity in zirconia: A study with machine-learned force fields, Phys. Rev. Appl. **18**, 054066 (2022).

[19] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, ReaxFF: A reactive force field for hydrocarbons, J. Phys. Chem. A **105**, 9396 (2001).

[20] T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, and A. C. T. van Duin, The ReaxFF reactive force-field: Development, applications and future directions, npj Comput. Mater. **2**, 15011 (2016).

[21] L. Zhang, J. Han, H. Wang, R. Car, and W. E, Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics, Phys. Rev. Lett. **120**, 143001 (2018).

[22] D. Zhang, H. Bi, F.-Z. Dai, W. Jiang, L. Zhang, and H. Wang, DPA-1: Pretraining of attention-based deep potential model for molecular simulation, arXiv:2208.08236.

[23] J. Wu, J. Yang, L. Ma, L. Zhang, and S. Liu, Modular development of deep potential for complex solid solutions, Phys. Rev. B **107**, 144102 (2023).

[24] S. Zhang, J. Luo, W. Hackenberger, and T. R. Shrout, Characterization of $Pb(In_{1/2}Nb_{1/2})O_3$–$Pb(Mg_{1/3}Nb_{2/3})O_3$–$PbTiO_3$ ferroelectric crystal with enhanced phase transition temperatures, J. Appl. Phys. **104**, 064106 (2008).

[25] X. Liu, S. Zhang, J. Luo, T. R. Shrout, and W. Cao, Complete set of material constants of $Pb(In_{1/2}Nb_{1/2})O_3$–$Pb(Mg_{1/3}Nb_{2/3})O_3$–$PbTiO_3$ single crystal with morphotropic phase boundary composition, J. Appl. Phys. **106**, 074112 (2009).

[26] D. Yi, N. Lu, X. Chen, S. Shen, and P. Yu, Engineering magnetism at functional oxides interfaces: Manganites and beyond, J. Phys.: Condens. Matter **29**, 443004 (2017).

[27] R. Ramesh and D. G. Schlom, Creating emergent phenomena in oxide superlattices, Nat. Rev. Mater. **4**, 257 (2019).

[28] L. Zhang, D.-Y. Lin, H. Wang, R. Car, and W. E, Active learning of uniformly accurate interatomic potentials for materials simulation, Phys. Rev. Mater. **3**, 023804 (2019).

[29] Y. Zhang, H. Wang, W. Chen, J. Zeng, L. Zhang, H. Wang, and W. E, DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models, Comput. Phys. Commun. **253**, 107206 (2020).

[30] J. Wu, Y. Zhang, L. Zhang, and S. Liu, Deep learning of accurate force field of ferroelectric HfO_2, Phys. Rev. B **103**, 024108 (2021).

[31] J. Wu, L. Bai, J. Huang, L. Ma, J. Liu, and S. Liu, Accurate force field of two-dimensional ferroelectrics from deep learning, Phys. Rev. B **104**, 174107 (2021).

[32] M. Chen, G. -C. Guo, and L. He, Systematically improvable optimized atomic basis sets for *ab initio* calculations, J. Phys.: Condens. Matter **22**, 445501 (2010).

[33] P. Li, X. Liu, M. Chen, P. Lin, X. Ren, L. Lin, C. Yang, and L. He, Large-scale *ab initio* simulations based on systematically improvable atomic basis, Comput. Mater. Sci. **112**, 503 (2016).

[34] M. Schlipf and F. Gygi, Optimization algorithm for the generation of ONCV pseudopotentials, Comput. Phys. Commun. **196**, 36 (2015).

[35] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Restoring the density-gradient expansion for exchange in solids and surfaces, Phys. Rev. Lett. **100**, 136406 (2008).

[36] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, J. Comput. Phys. **117**, 1 (1995).

[37] H. Wang, L. Zhang, J. Han, and W. E, DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics, Comput. Phys. Commun. **228**, 178 (2018).

[38] https://nb.bohrium.dp.tech/detail/6516485150.

[39] B. Jaffe, R. S. Roth, and S. Marzullo, Piezoelectric properties of lead zirconate-lead titanate solid-solution ceramics, J. Appl. Phys. **25**, 809 (1954).

[40] B. Jaffe, W. R. Cook, and H. Jaffe, *Piezoelectric Ceramics* (Academic Press, Cambridge, MA, 1971).

[41] T. Riste, E. Samuelsen, K. Otnes, and J. Feder, Critical behaviour of $SrTiO_3$ near the 105 K phase transition, Solid State Commun. **9**, 1455 (1971).

[42] Q. Li, S. Danilkin, G. Deng, Z. Li, R. L. Withers, Z. Xu, and Y. Liu, Soft phonon modes and diffuse scattering in $Pb(In_{1/2}Nb_{1/2})O_3$–$Pb(Mg_{1/3}Nb_{2/3})O_3$-$PbTiO_3$ relaxor ferroelectrics, J. Materiomics **4**, 345 (2018).

[43] D. Wang, M. Cao, and S. Zhang, Phase diagram and properties of $Pb(In_{1/2}Nb_{1/2})O_3$–$Pb(Mg_{1/3}Nb_{2/3})O_3$–$PbTiO_3$ polycrystalline ceramics, J. Eur. Ceram. Soc. **32**, 433 (2012).

[44] S. Liu, I. Grinberg, H. Takenaka, and A. M. Rappe, Reinterpretation of the bond-valence model with bond-order formalism: An improved bond-valence-based interatomic potential for $PbTiO_3$, Phys. Rev. B **88**, 104102 (2013).

[45] Y. Qi, S. Liu, I. Grinberg, and A. M. Rappe, Atomistic description for temperature-driven phase transitions in $BaTiO_3$, Phys. Rev. B **94**, 134308 (2016).

[46] H. Takenaka, I. Grinberg, S. Liu, and A. M. Rappe, Slush-like polar structures in single-crystal relaxors, Naure **546**, 391 (2017).

[47] D. Lu, W. Jiang, Y. Chen, L. Zhang, W. Jia, H. Wang, and M. Chen, DP compress: A model compression scheme for generating efficient deep potential models, arXiv:2107.02103.

[48] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray *et al.*, Training language models to follow instructions with human feedback, Adv. Neural Inf. Process. Syst. **35**, 27730 (2022).

[49] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, Generalizing from a few examples: A survey on few-shot learning, ACM Comput. Surv. **53**, 1 (2020).