

<https://doi.org/10.1038/s41524-024-01240-7>

Principal component analysis enables the design of deep learning potential precisely capturing LLZO phase transitions

Check for updates

Yiwei You¹, Dexin Zhang¹, Fulun Wu¹, Xinrui Cao^{1,2}, Yang Sun¹, Zi-Zhong Zhu^{1,2} & Shunqing Wu¹ ✉

The development of accurate and efficient interatomic potentials using machine learning has emerged as an important approach in materials simulations and discovery. However, the systematic construction of diverse, converged training sets remains challenging. We develop a deep learning-based interatomic potential for the $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ (LLZO) system. Our interatomic potential is trained using a diverse dataset obtained from databases and first-principles simulations. We propose using the coverage of the training and test sets as the convergence criteria for the training iterations, where the coverage is calculated by principal component analysis. This results in an accurate LLZO interatomic potential that can describe the structure and dynamical properties of LLZO systems meanwhile greatly reducing computational costs compared to density functional theory calculations. The interatomic potential accurately describes radial distribution functions and thermal expansion coefficient consistent with experiments. It also predicts the tetragonal-to-cubic phase transition behaviors of LLZO systems. Our work provides an efficient training strategy to develop accurate deep-learning interatomic potential for complex solid-state electrolyte materials, providing a promising simulation tool to accelerate solid-state battery design and applications.

The performance of traditional lithium-ion batteries is approaching its limit¹. The demand for energy density has increased due to the popularity of electric vehicles and the development of 5 G technology^{2–4}. To increase energy density, one strategy is to use a lithium metal anode^{5,6}. However, the safety of lithium metal is poor due to its reactivity with the electrolyte⁷. Therefore, the trend in battery development is shifting towards all-solid-state lithium metal batteries, which are composed of a chemically stable solid-state electrolyte instead of a liquid electrolyte. Among many solid-state electrolytes, $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ (LLZO) has gained extensive attention due to its excellent thermal stability, high Li^+ conductivity at room temperature, and wide electrochemical window^{8–11}. LLZO exists in both tetragonal and cubic phases, where the cubic phase exhibits higher Li-ion conductivity but lacks stability at room temperature^{12,13}.

Compared to batteries with liquid electrolytes, solid-state lithium batteries using LLZO as the solid electrolyte exhibit poor contact at the interface^{14,15}, phase transition issues^{16–19}, structural disorder, and chemical segregation²⁰, all of which contribute to the increased impedance at the

interface. Moreover, the propagation of Li dendrites along grain boundaries (GB) can cause short circuits during cycling²¹.

Currently, the main challenges faced by LLZO involve the microscale interfacial mechanisms of phase transformation, ionic transport, dendrite growth, and interface structure evolution that are difficult to directly observe in experiments^{22,23}. Molecular dynamics (MD) simulations can provide critical insights into these mechanisms by accessing the atomic-level processes, thus accelerating the design and optimization of LLZO-based solid electrolytes. However, utilizing Density Functional Theory (DFT) for large system models is impractical and computationally expensive. On the other hand, studying LLZO often produces results that are challenging to validate.

The deep interatomic potential (DP) was initially proposed by Behler and Parrinello in 2007 as the neural network potential (NNP)²⁴. Over the years, various interatomic potentials have been developed, including Gaussian approximation potential (GAP)²⁵, moment tensor potential (MTP), gradient-domain machine learning (GDML)²⁶, and deep potential for molecular dynamics (DeePMD)²⁷. DeePMD-kit²⁸ utilizes deep neural networks (DNNs) to train interatomic potential functions using ab initio

¹Department of Physics, OSED, Key Laboratory of Low Dimensional Condensed Matter Physics (Department of Education of Fujian Province), Xiamen University, Xiamen 361005, China. ²Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, Xiamen University, Xiamen 361005, China.

✉ e-mail: wsq@xmu.edu.cn

data, encompassing total potential energy, forces on individual atoms, and virials for a set of atomic configurations. These DNN interatomic potentials can accurately reproduce potentials and forces in the training dataset and largely improve efficiency for molecular dynamics simulations^{29–32}. The computational cost of MD simulations with DNN interatomic potentials scales linearly with the system size. However, a key factor in developing an accurate DP using machine learning is constructing an appropriate training set. Generating a massive number of random atomic configurations is computationally wasteful, as most contain redundant data and fail to adequately sample relevant configurations. An efficient approach must balance training set compactness against coverage of key atomic environments and interactions. This is particularly challenging for multi-component solid-state Li battery materials, which possess complex interface chemistry and a significantly large configuration space. It is vital to devise a reasonable method that produces a sufficiently comprehensive training set to accurately model the diverse atomic environments and complex interface phenomena in these materials.

In our study, we developed a potential function training process and obtained the training set for the Li–La–Zr–O quaternary system after several iterations. We extracted the local structural feature matrices of the training set and test set through principal component analysis (PCA). By comparing the coverage of the two feature matrices, we evaluated the comprehensiveness of the training set. By training a DP model, we obtained an interatomic potential that demonstrated effectiveness in predicting energy, force, and kinetic processes. The DP method proves to be reliable in predicting these properties for various chemical compositions, including crystalline and amorphous materials, as well as randomly generated compounds. This versatility enables us to perform cost-effective molecular dynamics simulations to investigate the complex structure of large-scale interfaces.

Results

The training and iteration procedure is as follows: first, we utilized data calculated by the crystal database and DFT as the initial training set. As shown in Fig. 1, we use the trained potential function to perform molecular dynamics simulations and then use PCA to calculate the coverage of local structural features in the dynamics trajectory to determine its convergence. If it does not converge, the molecular dynamics trajectories of the process are used for DFT calculations to compare their energy accuracy. Through error verification, structures with energy errors greater than 1% are used as supplementary training data sets.

During the process, we calculated the feature matrices of the training set and test set using PCA. We defined the coverage rate as the percentage of configurations in the test dataset that have similar representations in the training dataset. The coverage rate can qualitatively demonstrate the rationality and effectiveness of the training set.

Construction of training set

To develop an interatomic potential and a deep learning model capable of accurately describing the dynamics of complex LLZO systems, we created a comprehensive training set. The initial training set consists of three components.

The first part includes element and multiplex compounds of Li, La, Zr, and O, as well as elemental materials from the Materials Project (MP)³³ database. Additionally, we included structures derived from scaling the lattice constants for all chemical compositions and space groups. This part provided a diverse array of structures that served as fundamental building blocks for LLZO.

The second part includes structures obtained from first-principles molecular dynamics simulations to further expand the variety of LLZO structures in the training set. We performed simulations of LLZO crystals at different temperatures (400 K, 800 K, 1200 K, and 1600 K) and obtained amorphous structures by melting LLZO at 3000 K and cooling it to 300 K. We included structures at various temperatures (3000 K, 2000 K, 1000 K, 500 K, and 300 K) to capture the kinetic information of both crystalline and amorphous LLZO. This part of the training set provides valuable insights into kinetic properties such as energy, force, and dynamic processes.

The third part of the training set introduced a two-body potential to address the issue of atoms being too close or too far apart during molecular dynamics simulations. By incorporating this data, we effectively constrained the interatomic distances in the dynamic processes.

Potential iteration

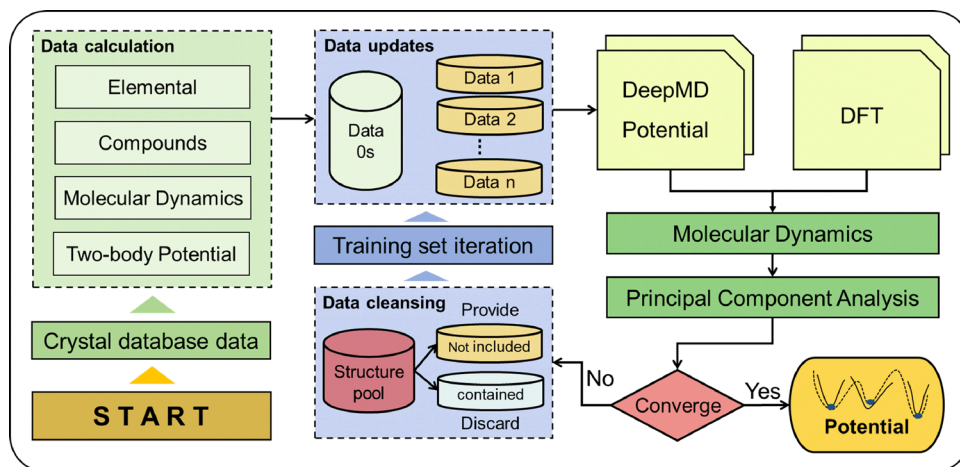
Figure 2a displays the error test results for crystal, amorphous, and slab structures using the interatomic potential solely obtained from the Data 0 training set. It can be observed that the error is within a few millielectronvolts for LLZO crystals but slightly larger for amorphous and surface structures. To address this, we constructed additional structures containing more surface and amorphous information. Specifically, we heated and melted a $3 \times 3 \times 3$ supercell structure (5184 atoms), subsequently cooling it down to obtain an amorphous structure.

We selected ten structures in the MD trajectories at three temperatures: 1000 K, 2000 K, and 3000 K. Because the amorphous only has short and medium-range orders, the small pieces should contain sufficient structure motifs that can represent the overall amorphous phase. To evaluate their DFT properties, we divided them into 64 smaller blocks and placed them in an empty cavity of $40 \text{ \AA} \times 40 \text{ \AA} \times 40 \text{ \AA}$ for calculation. This avoids unnecessary interactions from periodicity. The energy and force calculations to evaluate the errors between DP and DFT results, as shown in Fig. 2d. The parameters for the test calculations were consistent with those used to generate the initial training set.

We iterate this operation for structures at temperatures of 1000 K, 2000 K, and 3000 K and compared the errors of these structures, as shown in

Fig. 1 | Interatomic potential training flow chart.

The training set composition of the interatomic potential, error verification, and iterative process of the potential function.



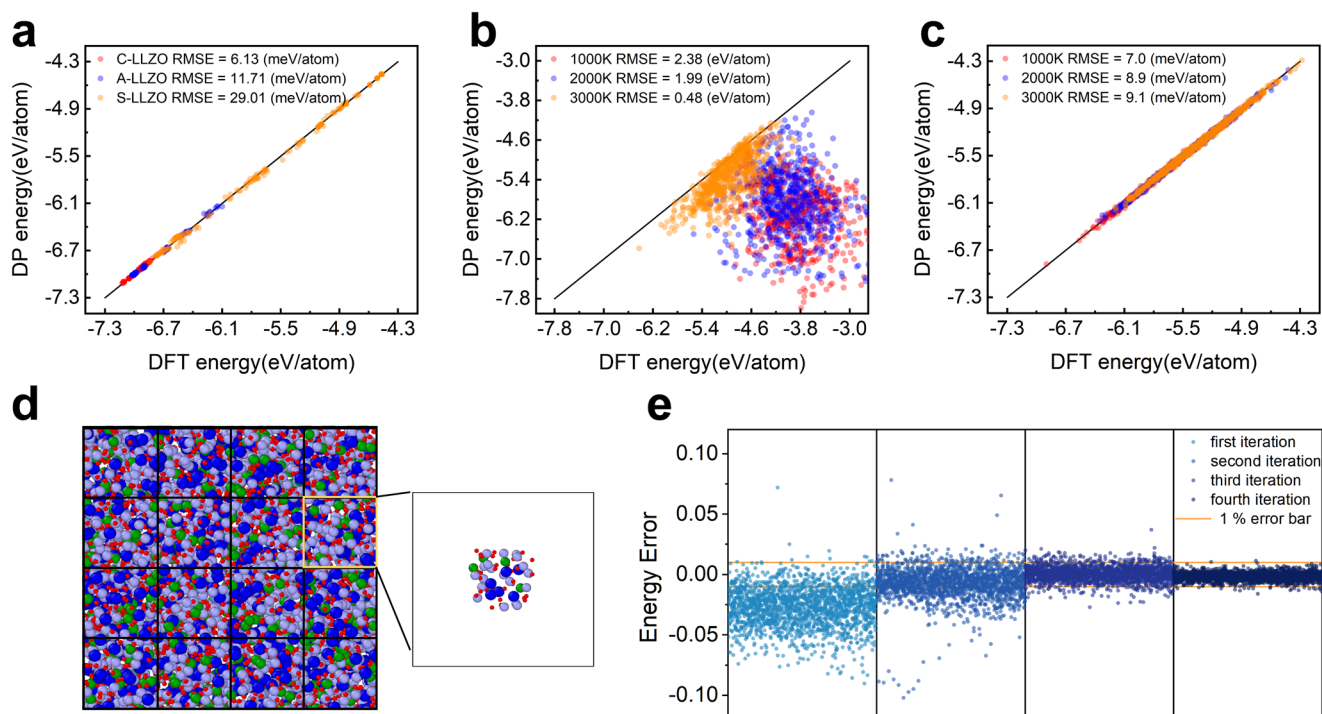


Fig. 2 | Error verification of the iterative process. **a** Testing of test sets with primary potential. **b** Testing of 1/64 structure with primary potential. **c** Testing of 1/64 structure with final potential. **d** Schematic diagram of 1/64 structures. **e** Error of the fourth iteration.

Fig. 2b. At the first iteration, the energy error is very significant, reaching the electron-volts level. We use structures with DFT and DP energy errors greater than 1% as supplementary training datasets. After four iterations, the energy error was greatly reduced to the millielectronvolt level, and the corresponding structure is shown in Fig. 2c. The error percentage for each iteration process is presented in Fig. 2e, and the final error is within 1%. This demonstrates the effectiveness of the present training strategy.

Figure 3a presents the coverage of the training set in the test set, both non-iteration and after the fourth iteration. It also provides the average coverage rate after each iteration. Initially, the coverage rate of the structure (test set) and training set generated by the potential function, without iteration, is only 75.34%. After four iterations, the coverage rate significantly improved to 99.51%. This improvement indicates that the existing training set adequately covers the generated structure, confirming the convergence of our training process.

Figure 3b displays the change in each iteration's root-mean-squared errors (RMSE). The iteration process demonstrates effective performance, with the error change stabilizing in the final iteration. Additionally, Fig. 3c illustrates the relationship between the coverage rate and the error percentage, demonstrating that utilizing the coverage rate is a reliable method to assess the sufficiency of the training set.

Reliability and validation of DP

To assess the accuracy of the DNN interatomic potential, we conducted a comparison between the energies and forces predicted by the DP model and DFT. Figure 3d–g provides a visual representation of this comparison, demonstrating the accuracy of DP in predicting energy and force per atom for crystal LLZO, amorphous LLZO, and slab LLZO in comparison to DFT. These plots reveal a strong correlation between the DP-predicted values and the results obtained from DFT. The RMSE for forces is below 200 meV/Å, and for energy, it is below 4 meV/atom. The wide energy distribution observed in the figure suggests a complex configuration space when exploring the potential energy surface (PES). In the second part of the Supporting Information, we provide details of tests conducted on the error of amorphous structures with varying ratios of Li, La, Zr, and O.

Moreover, we performed a comparison of the radial distribution function (RDF) for the cubic phase and amorphous LLZO during dynamic processes. Figure 4 showcases the RDFs of 1000 K cubic LLZO and 3000 K amorphous LLZO, acquired through molecular dynamic simulations of ab initio molecular dynamics (AIMD) and deep potential molecular dynamics (DPMD). The interatomic potential accurately predicts the RDF of both crystal and amorphous systems, further affirming the reliability and accuracy of our interatomic potential in dynamic processes. Additional RDFs for the system can be found in Supplementary Fig. 2 and Supplementary Fig. 3.

Furthermore, we believe that the application of two-body potentials is very effective for machine learning-based interatomic potential generation. By comparing the two-body potentials and forces between DFT and DP, we identified problematic atomic interactions requiring enhanced training. The two-body analysis also helped avoid unphysical artifacts from closely spaced atoms. More details are presented in the third part of the Supporting Information,

Application of DP in LLZO phase transition

The tetragonal phase of LLZO has ionic conductivity lower than the cubic phase by two to three orders of magnitude. However, the tetragonal phase is more stable at room temperature. Therefore, understanding the phase transition can facilitate the fabrication of the high ionic conductivity cubic phase.

One of the important characteristics of LLZO is its transition from the tetragonal phase to the cubic phase. In our study, we investigated the phase transition of LLZO from the tetragonal phase to the cubic phase using npt ensemble molecular dynamics simulations of a $3 \times 3 \times 3$ supercell of t-LLZO with 5148 atoms. The simulations, depicted in Fig. 5a, demonstrate that the interatomic potential effectively captures the thermal phase transitions of LLZO. We observed the tetragonal-to-cubic phase transition occurring around 900 K, as evidenced by the lattice constant, which is consistent with the experimental value of 923 K³⁴.

To further validate the phase transition, we examined the X-ray diffraction (XRD) pattern variation with temperature, as shown in Fig. 5d. It can be observed that as the temperature increases, the characteristic peaks of the tetragonal phase gradually disappear at 300 K and are completely

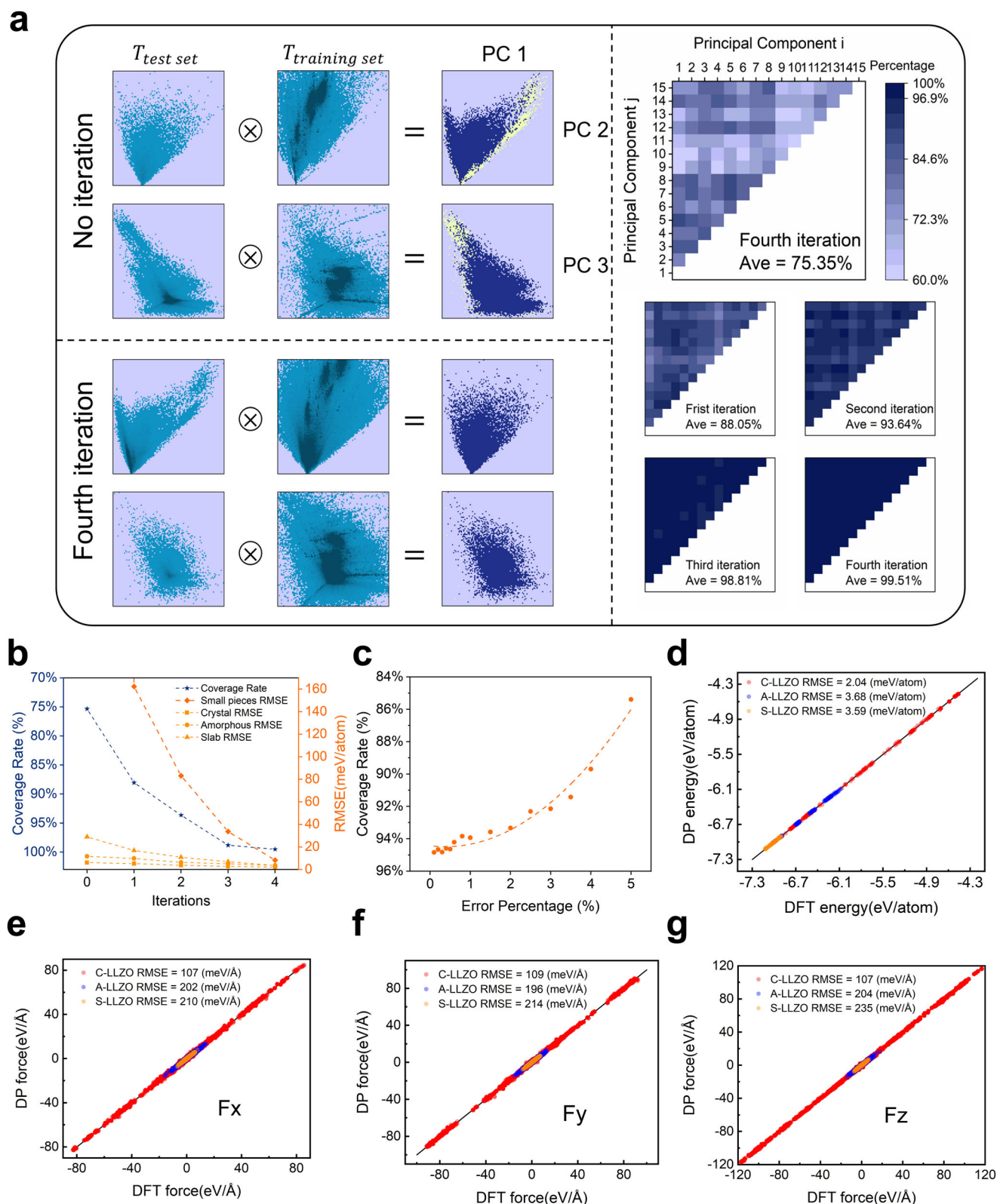


Fig. 3 | Changes in iterative process coverage. **a** Principal component analysis of the test set and training set before and after iteration, and the change of coverage. **b** Changes in RMSE of each category for each iteration. **c** The relationship between error and coverage. **d–g** The RMSE of the final interatomic potential for energy and force in each direction.

replaced by the characteristic peaks of the cubic phase at nearly 900 K. Additionally, we predicted the thermal expansion coefficient of c-LLZO between 1000 K and 1500 K to be $5.48 \times 10^{-5} \text{ K}^{-1}$, which is similar to the experimentally³⁴ measured value of $1.30 \times 10^{-5} \text{ K}^{-1}$. We observed a

significant volume mutation between 1900 K and 2100 K, resulting in random lattice constants a , b , and c . At the same time, the results of structural RDF (Fig. 5e) can explain that LLZO melts to form liquid amorphous LLZO at this time.

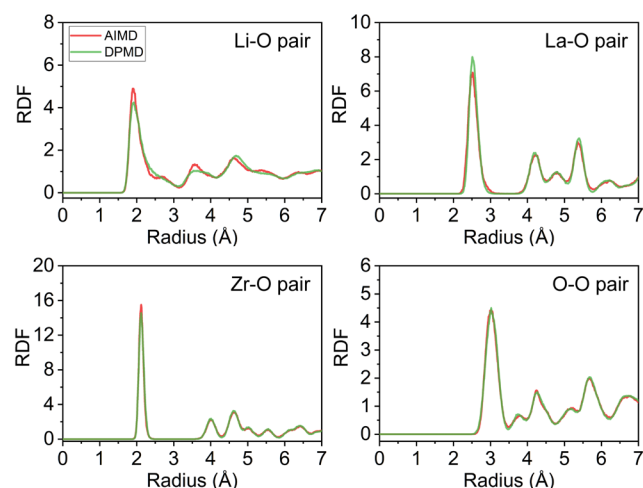
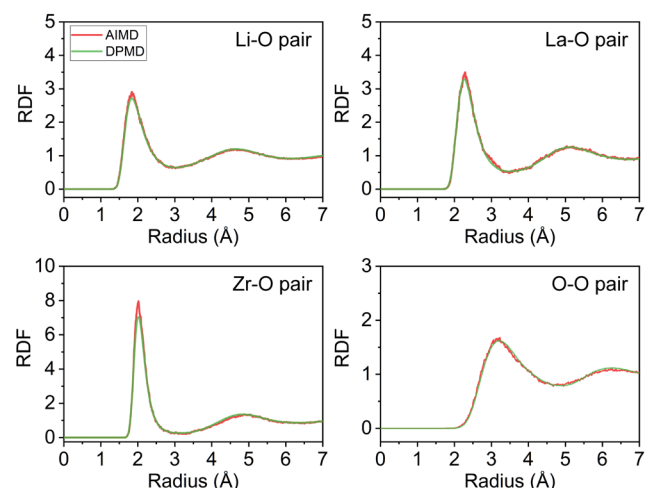
a Cubic LLZO at 1200K**b Amorphous LLZO at 3000K**

Fig. 4 | Comparison of RDFs of atomic pairs derived from AIMD and DPMD. a Cubic LLZO at 1200 K; **b** Amorphous LLZO at 3000 K.

Furthermore, Fig. 5b, c presents PCA comparisons of the crystal structure and amorphous structure with the training set, respectively. Both the crystalline structure and the amorphous structure are adequately covered in the training, enabling an accurate description of the tetragonal-to-cubic phase transition and subsequent melting.

Discussion

Because of the limitation of DFT, an issue arises: the interatomic potential trained on small systems can accurately describe small-scale models, but its transferability and accuracy for large-scale simulations are unclear. To demonstrate the accuracy of large-scale simulations, we adopted Principal Component Analysis to extract local structural feature matrices from the large-scale simulation results, which allows us to assess the reliability of the results. This method can evaluate the similarity between any structure and its training set, thus providing strong validation for the accuracy of large-scale simulations.

In this manuscript, we trained the DNN interatomic potentials for LLZO systems through DeepMD. Compared to DFT, the DP model can accurately predict the energies, forces, and molecular dynamics properties of the LLZO system at a greatly reduced cost. The training set for this study comprised three main components: databases, first-principles simulations, and two-body potentials. Through iterative training and validation, we achieved convergence and demonstrated the accuracy of the potential. In the iterative process, we employed PCA and coverage analysis to judge the convergence of training and indicate the reliability of the results.

We utilize this potential function to investigate the phase transition behavior of LLZO, specifically the transition from the tetragonal phase to the cubic phase. Molecular dynamics simulations captured the transition temperature and thermal expansion coefficient in good agreement with experimental values. The potential accurately predicted the radial distribution function for both crystal and amorphous systems, further validating its performance in dynamic processes.

Overall, by constructing a diverse training set and validating convergence, we develop a generalizable approach for training high-precision machine learning potentials. Our study verifies the efficacy of the deep learning-based interatomic potential in capturing the dynamics and phase transitions of LLZO systems. The DP model provides an accurate and efficient tool to investigate the microscale interface phenomena in solid-state Li batteries, which is challenging to explore experimentally. The accuracy, transferability, and convergence of the interatomic potential make it a valuable tool that enables extensive simulations to provide atomic-level insights into complex processes governing the performance of LLZO-based solid-state batteries.

Methods

First-principles calculations

Our first-principles calculations were performed by using the projector augmented wave (PAW)^{35,36} method within the density functional theory (DFT) as implemented in the Vienna Ab initio Simulation Package (VASP)^{37,38}. The electronic exchange–correlation function was treated within the spin-polarized generalized gradient approximation (GGA) parameterized by Perdew–Burke–Ernzerhof (PBE)³⁹. The convergence criteria were set to be 10^{-4} eV for the energy of the unit cell in the electronic minimization and 0.1 eV/Å for the force on each atom in relaxation, respectively. Electronic occupancies were decided using Gaussian smearing and an energy width of 0.1 eV in relaxation. The cutoff energy was set to 500 eV. Brillouin zones were sampled to accommodate different cell sizes by using the VASP k-spacing parameter $sk = 0.25$.

DP training details

The DeePMD-Kit²⁷ package utilizes the smoothing method during the training process of the DP model. For this work, we used the descriptor “se_e2_a”, which is short for the Deep Potential Smooth Edition (DeepPot-SE) constructed from all information (both angular and radial) of atomic configurations. The cutoff radius of adjacent atoms in the model was set to 6.0 Å, and the inverse distance was gradually smoothed from 0.5 Å to 6 Å. The filtering neural network was composed of three hidden layers [10, 20, 40], while the fitting network consisted of [120, 120, 120]. The neural network was initialized with random parameters, and the total number of training steps was 6000000. The Adam stochastic gradient descent method was used for training the model⁴⁰, which caused the learning rate to exponentially decrease relative to the starting value of 0.001. The decay step and decay rate were set to 2000 and 0.996, respectively.

The loss function L was defined as follows:

$$L(p_e, p_f, p_\xi) = \frac{p_e}{N} \Delta E^2 + \frac{p_f}{3N} \sum_i |\Delta F_i|^2 + \frac{p_\xi}{9N} \|\Delta \Xi\|^2 \quad (1)$$

where ΔE and ΔF_i represent the mean square errors in energies and forces, respectively. The energy prefactor p_e decreased from 0.02 to 1, while the force prefactor p_f decreased from 1000 to 1. It is worth noting that the training process did not include viral data.

Principal component analysis

Principal component analysis (PCA)⁴¹ is a widely adopted algorithm for reducing data dimensionality. Its primary objective is to transform

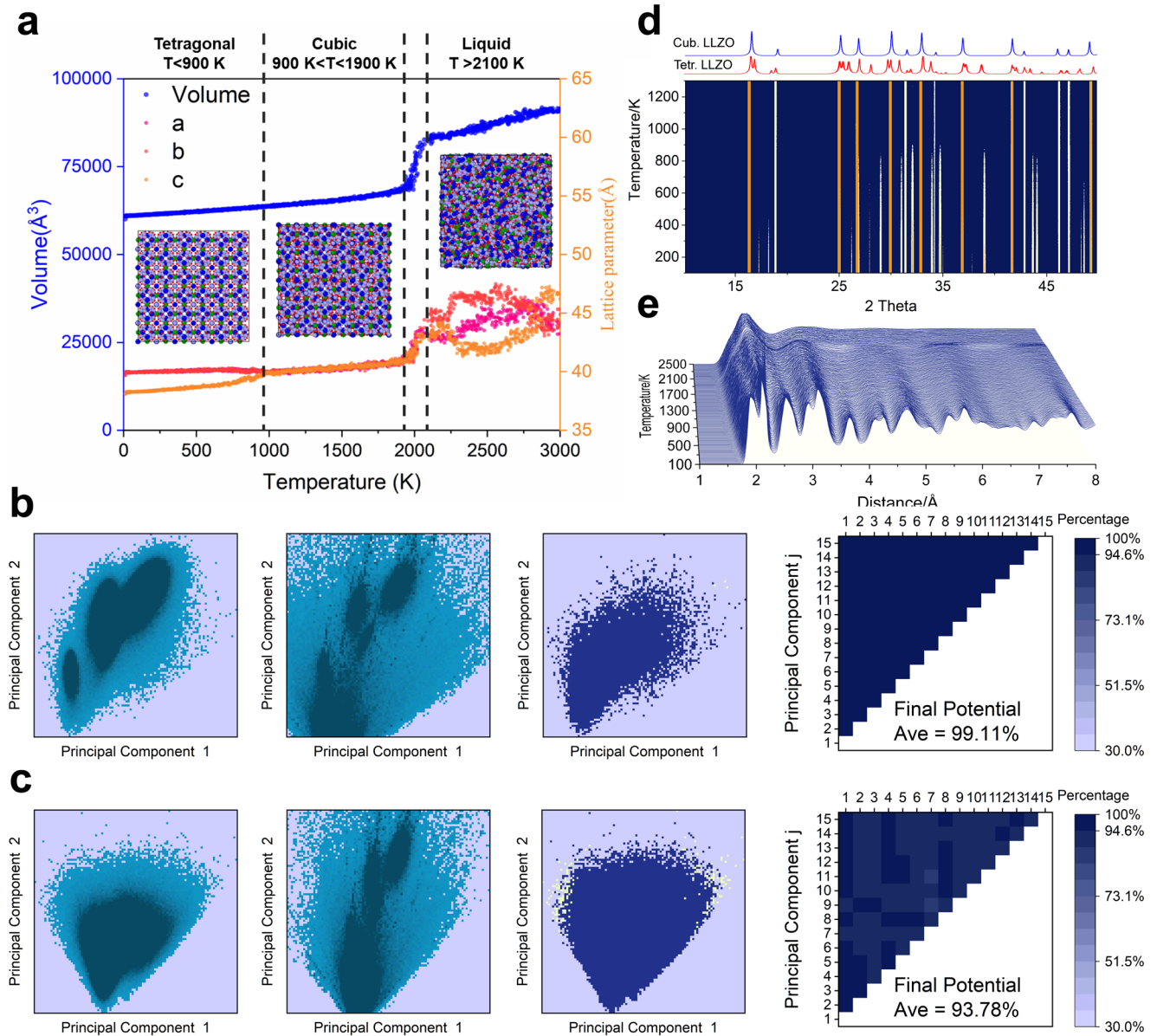


Fig. 5 | Coverage analysis and XRD comparison of the phase change process. **a** The lattice parameters and volume of LLZO as a function of temperature. **b** PCA comparison and coverage of the crystal structure with the training set. **c** PCA

comparison and coverage of the amorphous structure with the training set. **d** The relationship between XRD and temperature transformation. **e** The relationship between RDF and temperature transformation.

p -dimensional features into m -dimensions, where these m -dimensions are orthogonal features known as principal components. These principal components are reconstructed from the original p -dimensional features, providing a representation of the data in m -dimensions.

To analyze the local structures using descriptors, we obtain a matrix $x_{n \times p}$ that describes numerous local structures. We then standardize this matrix to obtain the standardized matrix $X_{n \times p}$. Next, we calculate the covariance matrix $R_{p \times p}$ based on $X_{n \times p}$. For a detailed computation process, please refer to Fig. 6 and the first section of the Supporting Information.

$$x_{n \times p} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}, X = \begin{bmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{np} \end{bmatrix}, R_{p \times p} = \begin{bmatrix} r_{11} & \cdots & r_{1p} \\ \vdots & \ddots & \vdots \\ r_{p1} & \cdots & r_{pp} \end{bmatrix} \quad (2)$$

As a result, we obtain the eigenvectors (a_1, a_2, \dots, a_p) and eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_p$.

$$a_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{p1} \end{bmatrix}, a_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{p2} \end{bmatrix}, \dots, a_p = \begin{bmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{pp} \end{bmatrix} \quad (3)$$

By multiplying the eigenvectors (a_1, a_2, \dots, a_p) with the standardized matrix $X_{n \times p}$, we obtain our target matrix $T_{n \times p}$.

$$T_{n \times p} = X_{np} \times (a_1, a_2, \dots, a_p) \quad (4)$$

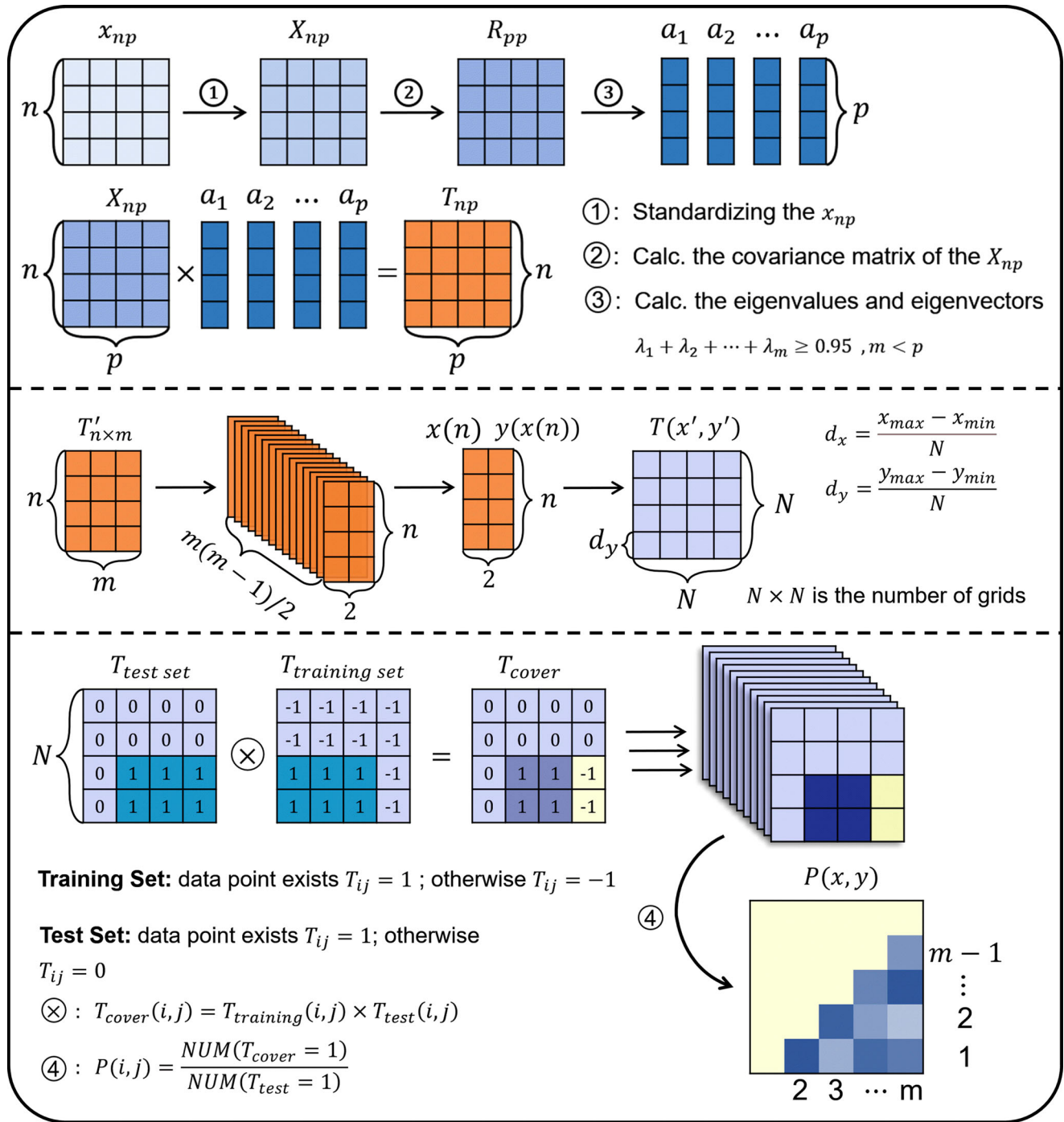


Fig. 6 | Schematic diagram of coverage calculation. PCA is the data dimensionality reduction process and coverage calculation method.

In the matrix $T_{n \times p}$, n still represents the number of local structures, and p remains the dimensionality.

Additionally, λ_m represents the contribution rate of the m th dimension. Typically, a contribution rate exceeding 95% is considered reliable. When the cumulative contribution rate $\lambda_1 + \lambda_2 + \dots + \lambda_m \geq 0.95 (m < p)$ is reached, the dimensionality is reduced from p -dimensions to m -dimensions. As a result, we obtain the transformed matrix $T'_{n \times m}$.

Calculation of coverage rate

In order to compare the coverage of the feature matrices of the training set and the test set, we need to check whether the values in each dimension match. However, the dimensionality is too high to be understood through visualization. So, we convert it into a projection on a two-dimensional plane

for understanding. This process can be understood by comparing the degree of coincidence of points in three-dimensional space. We only need to confirm that these points coincide in xy , xz , and yz , and you can determine that this point coincides in three-dimensional space. The same concept can be generalized to any dimension.

$$(m - 1) + (m - 2) + \dots + 1 = \frac{m(m - 1)}{2} \tag{5}$$

Consequently, the matrix $T'_{n \times m}$ can be decomposed into $\frac{m(m-1)}{2} n \times 2$ -dimensional matrices, which can then be projected onto a two-dimensional plane. By dividing the two-dimensional plane into equal partitions, we can

analyze the distribution of these points on the two-dimensional grid to determine if they are equal.

More importantly, each dimension can represent a certain feature of the local structure. Even if these points do not completely overlap in the high-dimensional space, as long as these structural features are consistent, the description of the test set can also be accurate.

We project the training set and test set onto a two-dimensional grid. If a point from the test set falls within a grid cell, we set $T_{ij} = 1$; otherwise, it is set $T_{ij} = -1$. For the training set, we set $T_{ij} = 1$ if a point exists within a grid cell, otherwise, it is set $T_{ij} = 0$. Consequently, we obtain two $N \times N$ matrices, $T_{\text{test set}}$ and $T_{\text{train set}}$. Additionally, we define $T_{\text{cover}}(i, j) = T_{\text{training}}(i, j) \times T_{\text{test}}(i, j)$, and by computing it, we obtain the distribution on the T_{cover} grid. A value of 1 in the T_{cover} matrix indicates that the local structure from the test set can be found in the training set, while a value of 0 means the structure does not exist in the test set, and -1 indicates that the structure from the test set is not found in the training set. By utilizing equation 9, we can determine the ratio of structures in the training set to those in the test set. Ideally, if all structures from the test set can be found in the training set, the coverage rate would be 100%.

$$P(i, j) = \frac{\text{NUM}(T_{\text{cover}} = 1)}{\text{NUM}(T_{\text{test}} = 1)}, i \in N, j \in N \quad (6)$$

Data availability

The training set is provided at <https://doi.org/10.5281/zenodo.10556106>. Example input files for DPMD calculations performed in this work are provided as Supplementary Information.

Code availability

The machine learning training potential function adopts the open-source DeepPMD-kit code (<https://github.com/deepmodeling/deepmd-kit>). All molecular dynamics simulations were performed with the open-source LAMMPS code (<https://github.com/lammps/lammps>).

Received: 16 September 2023; Accepted: 6 March 2024;

Published online: 18 March 2024

References

1. Janek, J. & Zeier, W. G. Challenges in speeding up solid-state battery development. *Nat. Energy* **8**, 230–240 (2023).
2. Yang, Z. et al. Electrochemical energy storage for green grid. *Chem. Rev.* **111**, 3577–3613 (2011).
3. Hesse, H., Schimpe, M., Kucevic, D. & Jossen, A. Lithium-ion battery storage for the grid—a review of stationary battery storage system design tailored for applications in modern power grids. *Energies* **10**, 2107 (2017).
4. Diouf, B. & Pode, R. Potential of lithium-ion batteries in renewable energy. *Renew. Energy* **76**, 375–380 (2015).
5. Wang, P. et al. Electro-chemo-mechanical issues at the interfaces in solid-state lithium metal batteries. *Adv. Funct. Mater.* **29**, 950 (2019).
6. Yang, L. et al. Lithium deposition on graphite anode during long-term cycles and the effect on capacity loss. *RSC Adv.* **4**, 26335–26341 (2014).
7. Chen, R. et al. Approaching practically accessible solid-state batteries: stability issues related to solid electrolytes and interfaces. *Chem. Rev.* **120**, 6820–6877 (2020).
8. Thangadurai, V. & Weppner, W. $\text{Li}_6\text{AlLa}_2\text{Ta}_2\text{O}_{12}$ (A = Sr, Ba): novel garnet-like oxides for fast lithium ion conduction. *Adv. Funct. Mater.* **15**, 107–112 (2005).
9. Liu, Q. et al. Challenges and perspectives of garnet solid electrolytes for all solid-state lithium batteries. *J. Power Sources* **389**, 120–134 (2018).
10. Ramakumar, S. et al. Lithium garnets: synthesis, structure, Li^+ conductivity, Li^+ dynamics and applications. *Prog. Mater. Sci.* **88**, 325–411 (2017).
11. Hiebl, C. et al. Proton bulk diffusion in cubic $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ garnets as probed by single X-ray diffraction. *J. Phys. Chem. C* **123**, 1094–1098 (2018).
12. Samson, A. J., Hofstetter, K., Bag, S. & Thangadurai, V. A bird’s-eye view of Li-stuffed garnet-type $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ ceramic electrolytes for advanced all-solid-state Li batteries. *Energy Environ. Sci.* **12**, 2957–2975 (2019).
13. Meier, K., Laino, T. & Curioni, A. Solid-state electrolytes: revealing the mechanisms of li-ion conduction in tetragonal and cubic LLZO by first-principles calculations. *J. Phys. Chem. C* **118**, 6668–6679 (2014).
14. Narayanan, S., Hitz, G. T., Wachsman, E. D. & Thangadurai, V. Effect of excess Li on the structural and electrical properties of garnet-type $\text{Li}_6\text{La}_3\text{Ta}_{1.5}\text{Y}_{0.5}\text{O}_{12}$. *J. Electrochem. Soc.* **162**, A1772–A1777 (2015).
15. Ohta, S., Kobayashi, T., Seki, J. & Asaoka, T. Electrochemical performance of an all-solid-state lithium ion battery with garnet-type oxide electrolyte. *J. Power Sources* **202**, 332–335 (2012).
16. Banerjee, A. et al. Interfaces and interphases in all-solid-state batteries with inorganic solid electrolytes. *Chem. Rev.* **120**, 6878–6933 (2020).
17. Park, K. et al. Electrochemical nature of the cathode interface for a solid-state lithium-ion battery: interface between LiCoO_2 and garnet- $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$. *Chem. Mater.* **28**, 8051–8059 (2016).
18. Ma, C. et al. Interfacial stability of Li metal-solid electrolyte elucidated via in situ electron microscopy. *Nano Lett.* **16**, 7030–7036 (2016).
19. Rettenwander, D. et al. Interface instability of Fe-stabilized $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ versus Li metal. *J. Phys. Chem. C. Nanomater Interfaces* **122**, 3780–3785 (2018).
20. Li, Y., Cao, Y. & Guo, X. Influence of lithium oxide additives on densification and ionic conductivity of garnet-type $\text{Li}_{6.75}\text{La}_3\text{Zr}_{1.75}\text{Ta}_{0.25}\text{O}_{12}$ solid electrolytes. *Solid State Ion.* **253**, 76–80 (2013).
21. Cheng, E. J., Sharafi, A. & Sakamoto, J. Intergranular Li metal propagation through polycrystalline $\text{Li}_{6.25}\text{Al}_{0.25}\text{La}_3\text{Zr}_2\text{O}_{12}$ ceramic electrolyte. *Electrochim. Acta* **223**, 85–91 (2017).
22. Sharafi, A. et al. Controlling and correlating the effect of grain size with the mechanical and electrochemical properties of $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$ solid-state electrolyte. *J. Mater. Chem. A* **5**, 21491–21504 (2017).
23. Liu, X. et al. Local electronic structure variation resulting in Li ‘filament’ formation within solid electrolytes. *Nat. Mater.* **20**, 1485–1490 (2021).
24. Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
25. Bartok, A. P., Payne, M. C., Kondor, R. & Csanyi, G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).
26. Chmiela, S. et al. Machine learning of accurate energy-conserving molecular force fields. *Sci. Adv.* **3**, e1603015 (2017).
27. Zhang, L. et al. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett.* **120**, 143001 (2018).
28. Wang, H., Zhang, L., Han, J. & E, W. DeepPMD-kit: a deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.* **228**, 178–184 (2018).
29. Chen, W. K. et al. Deep learning for nonadiabatic excited-state dynamics. *J. Phys. Chem. Lett.* **9**, 6702–6708 (2018).
30. Zhang, L., Wang, H., Car, R. & Weinan, E. Phase diagram of a deep potential water model. *Phys. Rev. Lett.* **126**, 236001 (2021).
31. Calegari Andrade, M. F. et al. Free energy of proton transfer at the water- TiO_2 interface from ab initio deep potential molecular dynamics. *Chem. Sci.* **11**, 2335–2341 (2020).
32. Tang, L. et al. Development of interatomic potential for Al-Tb alloys using a deep neural network learning method. *Phys. Chem. Chem. Phys.* **22**, 18467–18479 (2020).
33. Jain, A. et al. Commentary: the materials project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).

34. Larraz, G., Orera, A. & Sanjuán, M. Cubic phases of garnet-type $\text{Li}_7\text{La}_3\text{Zr}_2\text{O}_{12}$: the role of hydration. *J. Mater. Chem. A* **1**, 11419–11428 (2013).
35. Blochl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
36. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
37. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B Condens Matter* **47**, 558–561 (1993).
38. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
39. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
40. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.1412.6980> (2014).
41. Abdi, H. & Williams, L. J. Principal component analysis. *Wiley Interdiscip. Rev.* **2**, 433–459 (2010).

Acknowledgements

This research was supported by the National Natural Science Foundation of China (11874307). Shaorong Fang and Tianfu Wu from the Information and Network Center of Xiamen University are acknowledged for their help with Graphics Processing Unit (GPU) computing.

Author contributions

Y.Y. performed all calculations of LLZO (DFT, DeePMD-kit, and DPMD) and provided part of the training set. D.Z. provided part of the training. F.W. provided the processing procedure for PCA. S.W. directed the entire project. All authors contributed to analyzing the data and writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-024-01240-7>.

Correspondence and requests for materials should be addressed to Shunqing Wu.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024