npj | computational materials

Article

Published in partnership with the Shanghai Institute of Ceramics of the Chinese Academy of Sciences

https://doi.org/10.1038/s41524-025-01740-0

DPmoire: a tool for constructing accurate machine learning force fields in moiré systems

Jiaxuan Liu[1,2], Zhong Fang[1,2,3], Hongming Weng[1,2,3] ✉ & Quansheng Wu[1,2] ✉

In moiré systems, the impact of lattice relaxation on electronic band structures is significant, yet the computational demands of first-principles relaxation are prohibitively high due to the large number of atoms involved. To address this challenge, We introduce a robust methodology for the construction of machine learning potentials specifically tailored for moiré structures and present an open-source software package ***DPmoire*** designed to facilitate this process. Utilizing this package, we have developed machine learning force fields (MLFFs) for MX_2 (M = Mo, W; X = S, Se, Te) materials. Our approach not only streamlines the computational process but also ensures accurate replication of the detailed electronic and structural properties typically observed in density functional theory (DFT) relaxations. The MLFFs were rigorously validated against standard DFT results, confirming their efficacy in capturing the complex interplay of atomic interactions within these layered materials.

In recent years, two-dimensional twisted moiré structures have captured significant interest due to the diverse physical phenomena they exhibit. By varying the interlayer twist angle, researchers can tune the band structure of these materials, enabling the experimental observation of novel phenomena. For instance, in twisted graphene, when the twist angle reaches the so-called "magic angle", the valence band flattens, prompting electrons to transition from a weakly correlated to a strongly correlated state. This shift gives rise to a host of intriguing behaviors, including unconventional superconductivity, Mott insulating states, and the quantum anomalous Hall effect[1–12]. Similar phenomena have also been observed in moiré bilayers of transition metal dichalcogenides (TMDs)[13–19].

In twisted structures, the moiré potential narrows the bandwidth as the periodicity of the structure increases. For instance, the bandwidth of bilayer twisted graphene at a twist angle of 1.08° is only a few meV[4,20], while the bandwidth of bilayer twisted $MoTe_2$ at 3.89° is just over 10 meV[14]. Such narrow bands are highly susceptible to the effects of lattice relaxation, which significantly influences their electronic properties. Theoretical calculations reveal that the electronic band structures of rigid twisted graphene differ markedly from those of relaxed systems[20]. Additionally, experimental studies using scanning tunneling microscopy have also documented the relaxation patterns in TMDs resulting from lattice reconstruction[21,22].

To accurately model the electronic properties of moiré structures, density functional theory (DFT) is often employed, particularly for structures with large twist angles, where it is considered essential for reliable structural relaxation[14,15,19]. However, despite its high level of accuracy, the computational complexity of DFT scales cubically with the number of atoms. The atoms in moiré structures increase dramatically as the twist angle decreases (Table 1), rendering the DFT calculation impractical for smaller-angle structures due to the sheer number of atoms involved.

To address this computational challenge, researchers have developed parameterized continuum models that are better suited for structures with small twist angles[23–30]. While these models provide a computationally feasible alternative, they typically do not reach the accuracy levels of DFT relaxation. For materials such as graphene[31–33] and transition metal dichalcogenides (TMDs)[34], empirical force fields have been effectively utilized for structural relaxation[12,35–38]. However, in other systems, robust and extensively validated empirical potentials remain scarce, limiting the scope of studies that can be conducted.

Machine learning force fields (MLFF) offer a promising solution to the computational challenges posed by moiré structures[39–50]. Recent advancements in universal MLFFs have shown great promise in terms of versatility, efficiency, and accuracy for materials discovery and high-throughput calculations[51–56]. Universal MLFFs typically can achieve an energy error of several tens of millielectron volts (meV) per atom. For example, the mean absolute energy errors of CHGNET[52] and ALIGNN-FF[54] are 33 meV per atom and 86 meV per atom, respectively. However, in the context of moiré systems, the energy scales of electronic bands are often on the order of meV, a range comparable to the accuracy limits of these universal MLFFs. This indicates that while universal MLFFs provide broad applicability, their precision may be insufficient for structural relaxation tasks in moiré systems, necessitating the development of MLFFs specifically tailored to individual material systems. Algorithms such as NequIP[41] and Allegro[40] can

[1]Beijing National Laboratory for Condensed Matter Physics and Institute of Physics, Chinese Academy of Sciences, Beijing, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]Songshan Lake Materials Laboratory, Dongguan, Guangdong, China. ✉e-mail: hmweng@iphy.ac.cn; quansheng.wu@iphy.ac.cn

achieve errors of approximately a fraction of a meV per atom when trained on specific materials, which is accurate enough for moiré systems.

Previous efforts have successfully constructed MLFFs for twisted structures, achieving encouraging outcomes. Some studies have developed MLFFs for large twist angles and then applied these models to smaller angles[57,58], while others have trained MLFFs on non-twisted structures before using them to relax twisted configurations[14,59]. Additionally, a few approaches have combined initial training on non-twisted structures with subsequent transfer learning on large twist-angle structures to efficiently relax twisted configurations[60,61]. This multifaceted strategy highlights the adaptability of MLFFs in addressing the specific challenges posed by the diverse configurations encountered in moiré systems.

While these innovative approaches have shown promise, their validation has often been limited to specific materials, and a comprehensive tool for constructing MLFFs tailored to twisted structures is still lacking. Moiré systems offer a unique platform for exploring novel phenomena such as strong correlations and topological states, with numerous experimental and theoretical advances highlighting their potential. Given the rapid development in this field, there is a pressing need for a universal tool that can conveniently and efficiently construct MLFFs for such complex systems. To bridge this gap, we propose a new methodology and introduce an open-source software, *DPmoire*, designed specifically for moiré systems. *DPmoire* leverages non-twisted structures to construct training datasets, facilitating the automated generation of MLFFs tailored to the unique challenges of moiré systems. This tool aims to streamline the MLFF construction process, enabling researchers to more effectively study and model the intricate behaviors exhibited by twisted materials.

Table 1 | Number of atoms in moiré cell of twisted bilayer TMDs

Twist angle (°)	21.79	13.17	9.43	7.34	6.01	5.09
Number of atoms	42	114	222	366	546	762
Twist angle (°)	4.41	3.89	3.48	3.15	2.88	2.65
Number of atoms	1014	1302	1626	1986	2382	2814

Results

MLFF for moiré systems

To develop an MLFF for moiré superlattice structures, we initially constructed 2×2 supercells of non-twisted bilayers and introduced in-plane shifts to generate various stacking configurations. Subsequently, structural relaxations were performed for each configuration, ensuring that the x and y coordinates of a reference atom from each layer remained fixed to prevent structural drift toward energetically favorable stackings. The lattice constants were also held constant throughout the simulations. The relaxation data were compiled into a training dataset.

Following the relaxation phase, Molecular Dynamics (MD) simulations were conducted under the aforementioned constraints to augment the training data pool. For these simulations, we employed the VASP MLFF module to explore a wide range of atomic configurations. VASP MLFF module is an on-the-fly MLFF algorithm, which will be described in detail in section "Machine learning force fields". Then, we selectively incorporated data solely from DFT calculation steps. Given the potential instability when initiating MD simulations with VASP MLFF from an untrained state, we initially established a baseline MLFF using single-layer structures before proceeding with the full simulations. To ensure the MLFF's applicability to moiré systems and to mitigate overfitting to non-twisted structures, we constructed the test set using large-angle moiré patterns. These were subjected to ab initio relaxations, with the resultant data serving as the test set.

Finally, the compilation of the aforementioned datasets facilitated the training of a robust and accurate MLFF. While we utilized the Allegro framework for MLFF training in this study, other MLFF algorithms, such as

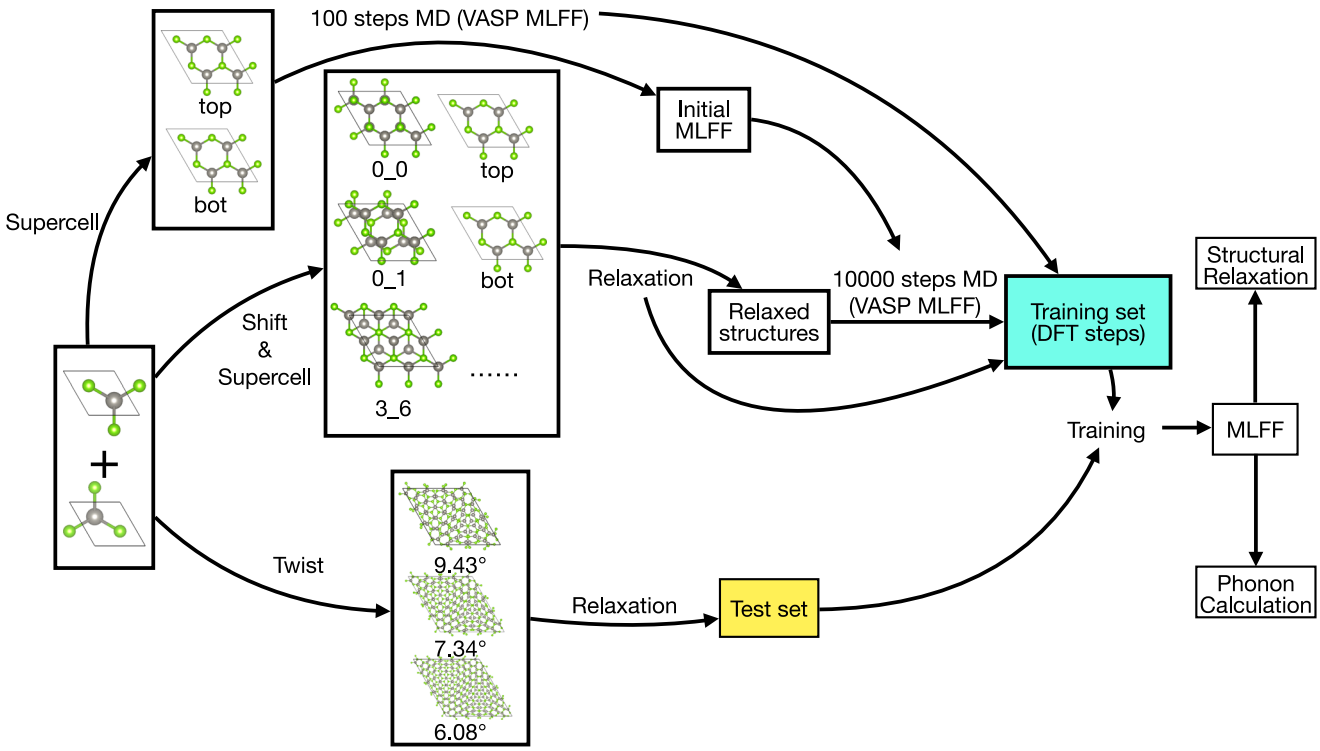

Fig. 1 | Schematic overview of the process for constructing the MLFF. Initially, an MLFF is generated for monolayer structures to stabilize subsequent molecular dynamics (MD) simulations for bilayer systems. We then create non-twisted bilayer structures with various stacking configurations, relax these structures, and run MD simulations using the VASP MLFF module to construct the training dataset. The coordinates (x and y) of a selected atom from each layer are maintained constant during relaxation to preserve the integrity of the stacking order. Subsequently, the twisted structures are relaxed using density functional theory (DFT) to generate the test dataset. The MLFF is ultimately trained on these collected datasets, ensuring it can accurately predict the physical behaviors of moiré systems.

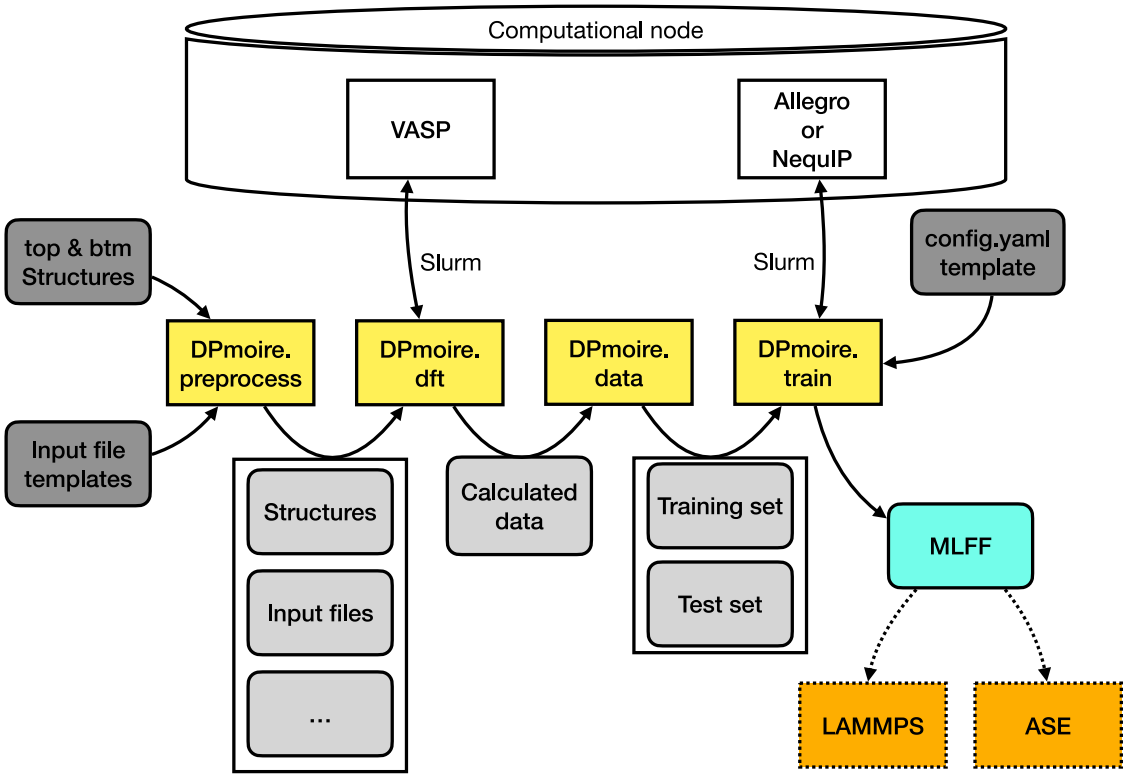

Fig. 2 | Overview of the DPmoire package workflow. Initially, the **preprocess** module utilizes the provided structural files for each layer along with an input template to generate the necessary input files for subsequent VASP DFT calculations. The **dft** module then orchestrates these calculations using the Slurm management system. Upon completion, the **data** module collects the results and compiles them into datasets. Subsequently, the **train** module begins training a machine learning force field using these datasets, adhering to the parameters specified in the MLFF configuration template file. Once trained, the MLFF can be integrated with software packages such as LAMMPS[63] or ASE[62] to facilitate structural relaxation.

Table 2 | Settings in training MLFFs

Settings	l_{max}	n_{layer}	learning rate	optimizer	lr scheduler	schduler patience	schduler factor
Values	2	2	0.001	Adam	ReduceLROnPlateau	50	0.5

l_{max} denotes the maximum order of irreducible representation in E(3) equivariant networks, and n_{layer} denotes the number of equivariant layers in allegro MLFF.

DeepMD[42], could also be effectively trained on these datasets to potentially enhance predictive accuracy and transferability across similar complex structures (Fig. 1).

Eventually, the procedure described above was implemented in DPmoire. As shown in Fig. 2, DPmoire is structured into four functional modules: DPmoire.preprocess, DPmoire.dft, DPmoire.data, and DPmoire.train. Firstly, as provided the unit cell structures of each layer, DPmoire.preprocess module will automatically combine two layers and generate shifted structures of a 2×2 supercell. The twisted structure for the building test set will also be prepared. The preprocess module will take care of the input files for VASP according to the provided templates. After that, the DPmoire.dft module will submit VASP calculation jobs through slurm system. When all the calculation is done, the DFT-calculated data (Energy, Force, and Stress) will be collected by DPmoire.data module. Then, DPmoire.data will merge the data into training set and test set. DPmoire.train module will modify the system-dependent settings in the configuration file according to given template for training Allegro or NequIP MLFF, and submit the training job. After the training is done, the trained MLFF can be used in ASE[62] or LAMMPS[63] to perform structural relaxation.

Performance of generated MLFF

The accuracy of MLFF is critically dependent on the precision of underlying density functional theory (DFT) calculations. Particularly in layered materials, the van der Waals (vdW) interactions play a crucial role in determining the DFT-calculated interlayer distances, making their inclusion indispensable. Over the years, a plethora of vdW correction methodologies have been developed[64–80]. Despite these developments, the predicted interlayer distances using different vdW corrections can vary by a few tenths of an Ångstrom.

Given this variation, it is crucial to identify the most appropriate vdW correction for each material prior to the training of MLFFs. To this end, we evaluated the lattice constants obtained under various vdW corrections, comparing them against experimental measurements to ascertain the optimal vdW correction for each material. The details of this comparative analysis and the optimal vdW corrections are documented in Section I of Supplementary Information, providing a rigorous foundation for the subsequent MLFF training. These tailored corrections are crucial for enhancing the accuracy of DFT calculations, thereby improving the robustness of the developed MLFFs for different TMD materials.

Then, the MLFF is constructed utilizing the previously determined optimal vdW corrections for both AA and AB stacking configurations of MX_2 (M = Mo, W; X = S, Se, Te) materials, as thoroughly discussed in Section I of Supplementary Information. Settings used to train the MLFFs are shown in Table 2. We specifically examined AA WSe_2 and AA MoS_2 as representative examples. The efficacy of the MLFF is demonstrated through a comparison of predicted and DFT-calculated forces within the test set, as illustrated in Fig. 3. The comparison shows a strong alignment between the

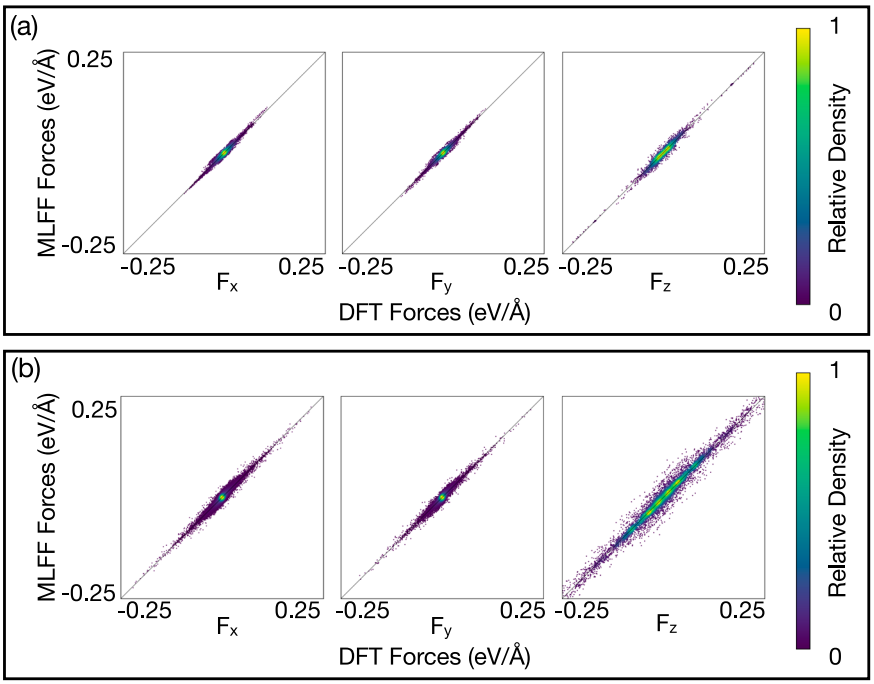

Fig. 3 | Force error of generated MLFF. a MLFF-predicted versus DFT-calculated forces for AA WSe_2 in test set of 7.34° twist. **b** Similar comparison for AA MoS_2 in test set including 9.34°, 7.34°, and 6.08° twists.

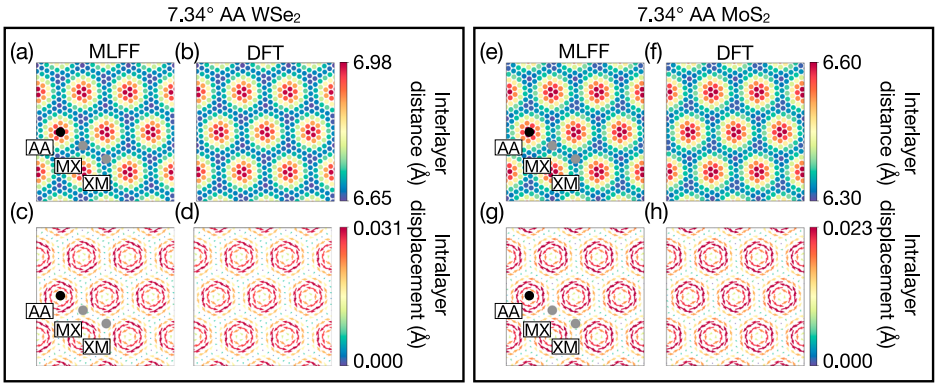

Fig. 4 | Relaxation pattern of 7.34° AA WSe_2 and MoS_2. a, **c** correspond to the interlayer distance and intralayer displacement in MLFF-relaxed WSe_2, respectively. **b**, **d** correspond to the interlayer distance and intralayer displacement in DFT-relaxed WSe_2. **e**, **g** correspond to the interlayer distance and intralayer displacement in MLFF-relaxed MoS_2. **f**, **h** correspond to the interlayer distance and intralayer displacement in DFT-relaxed MoS_2.

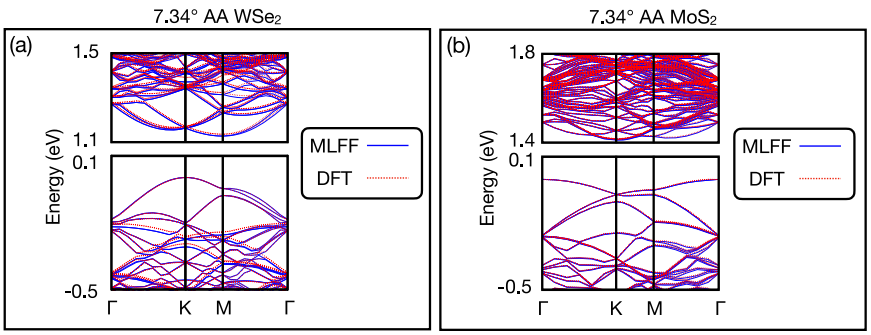

Fig. 5 | Band structure comparison of 7.34° AA WSe_2 and MoS_2. a Comparison of electronic band structure between MLFF-relaxed structure and DFT-relaxed structure in 7.34° AA WSe_2. **b** Comparison of electronic band structure between MLFF-relaxed structure and DFT-relaxed structure in 7.34° AA MoS_2.

MLFF predictions and the DFT calculations, with root mean square errors of 0.007 eV/Å and 0.014 eV/Å for WSe_2 and MoS_2, respectively, underscoring the accuracy of the MLFF in capturing the essential physical interactions in these materials.

We further evaluated the performance of the trained MLFFs by relaxing a structure with a 7.34° twist angle, followed by a comparison relaxation using DFT. As depicted in Fig. 4, the relaxation outcomes from the MLFF are nearly indistinguishable from those obtained via DFT, with no significant deviations observed. The maximum differences in atomic positions were found to be 0.039 Å in WSe_2 and 0.003 Å in MoS_2. In the relaxed structures, regions characterized by MX and XM stacking exhibited lower interlayer distances compared to the AA regions. Throughout the relaxation process, atoms near the AA regions tend to rotate counterclockwise, which intensifies the local twist effect. Conversely, atoms in proximity to the MX and XM regions rotate clockwise. This differential rotation behavior strategically

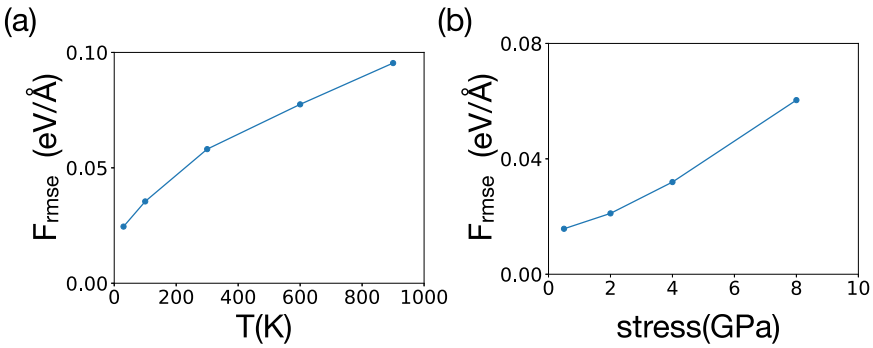

Fig. 6 | RMSE of forces of 7.34° AB MoS_2 in different temperatures and stresses. a RMSE under different temperatures. For each data point, we sampled 10 structures evenly from 1-ps MLFF-MD simulation to calculate the error. **b** RMSE under different stresses. For each data point, we sampled 10 structures evenly from the relaxation trajectories. Stresses in the z-direction are applied by imposing forces on the top and bottom sulfur atoms.

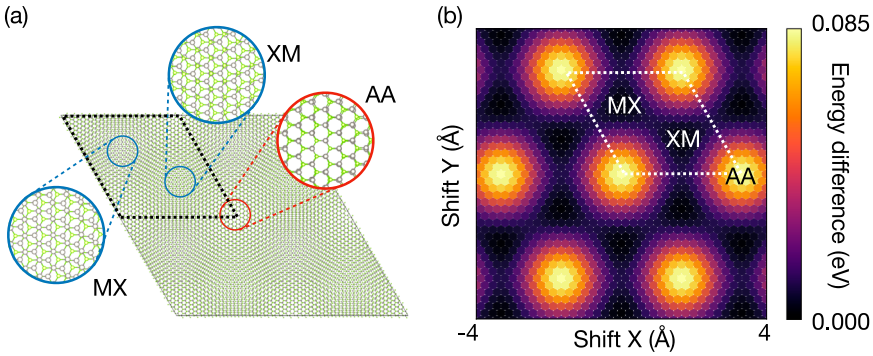

Fig. 7 | Characteristic of moiré structures. a Moiré crystal structure of WSe_2 with a 2.13° AA stacking twist, resembling the atomic layout of non-twisted bilayer WSe_2. **b** Energy profile of non-twisted bilayer WSe_2 based on relative in-plane shifts between layers, where X and Y axes represent shift vectors, and color indicates unit cell energy. Energy at MX and XM stackings is zeroed. Interlayer distance is 6.8 Å.

maximizes the area of MX and XM regions while minimizing the AA region. These findings align well with previous theoretical studies[23].

To further investigate how different relaxation approaches affect the computed band structures, we also performed band structure calculations on both MLFF-relaxed and DFT-relaxed structures for AA WSe_2 and AA MoS_2, as shown in Fig. 5. The band structures of the two methods are nearly identical, with only minor differences, demonstrating that the MLFF is sufficiently accurate to capture the essential physical phenomena in moiré structures without the need for additional DFT relaxation. As detailed in Section II of Supplementary Information, MLFFs for other materials also exhibited robust performance. For MoS_2, WS_2, AB $MoTe_2$, and WTe_2, the structures relaxed by MLFF and DFT methods were nearly identical, and their corresponding band structures closely matched. However, for materials like $MoSe_2$ and AA-stacked $MoTe_2$, slight variations in interlayer distances led to minor differences in their band structures. We further analyzed the 5.09° twist angle in AA and AB stacked $MoSe_2$ (Supplementary Fig. 13), where the discrepancies between DFT-relaxed and MLFF-relaxed structures were reduced, suggesting that the observed suboptimal performance in these materials may be due to the larger twist angles. In large-angle structures, the lattice mismatch between layers is not negligible, and such atomic configurations rarely appear in the training dataset.

Moreover, we evaluated the transferability of our MLFF with 7.34° AB MoS_2 as an example. The root mean squared errors (RMSE) under different temperatures and stresses are shown in Fig. 6. For the temperature tests, we conducted 1-picosecond MLFF-MD simulations at each temperature and sampled 10 structures evenly from the trajectories to compute the force errors against DFT references. For the stress tests, we performed structural relaxations starting from a rigid structure and similarly sampled 10 structures from each relaxation trajectory. Stresses in the z-direction are applied by imposing forces on the top and bottom sulfur atoms. The results demonstrate excellent transferability of the MLFF across varying temperatures and stresses. This indicates that our MLFF is not only suitable for structural relaxation but also robust for MD simulations under diverse conditions.

Discussion

In this work, we introduced a universal methodology and developed an open-source tool, DPmoire, for constructing MLFF tailored to moiré structures. Utilizing the VASP MLFF module, DPmoire effectively generates training sets and constructs validation sets based on large-twist-angle configurations. We successfully trained accurate MLFFs for MX_2 (M = Mo, W; X = S, Se, Te) systems, which precisely replicate both the relaxation patterns and electronic band structures observed in DFT relaxations, but at a significantly reduced computational cost.

This innovative tool enables the effective relaxation of moiré systems across a broader range of smaller angles and varied materials. Additionally, it facilitates phonon calculations within these complex systems. We anticipate that DPmoire will significantly enhance the understanding of physical phenomena influenced by relaxation effects and spur the discovery of novel moiré materials.

Moreover, we found that for moiré systems, carefully constructing the training set can significantly improve the accuracy of MLFF. We believe that for other systems, designing the training set according to the common characteristic might also be a promising approach to get an accurate model.

Methods

Moiré structures

Moiré twisted materials could be constructed by either applying a twist angle between layers of two layered materials or stacking two materials with a slight lattice constant mismatch. Generally, the smaller the twist angle, the larger the resulting moiré supercell. Different regions of a moiré structure exhibit various stacking arrangements. Taking twisted AA WSe_2 as an example (Fig. 7), in the AA region, the W/Se atoms in the top layer are aligned with the corresponding W/Se atoms in the bottom layer. In the MX region, the W atoms in the top layer align with the Se atoms in the bottom layer, while in the XM region, the Se atoms in the top layer align with the W atoms in the bottom layer. In non-twisted structures, various stacking configurations correspond to different energy states, as illustrated in Fig.

7(b). When the interlayer twist angle is minimal, the lattice vectors of both layers closely match, making the local atomic configurations in the moiré structure similar to those in non-twisted structures. By modeling the potential energy surfaces of these non-twisted configurations, we can effectively reconstruct the potential energy landscape of twisted structures, thereby advancing our understanding of their unique properties.

Machine learning force fields

MLFF[39–50] refers to machine learning algorithms for predicting the energy and forces of crystal structures. Typically, to train an MLFF, it needs a dataset consisting of a set of crystal structures along with their corresponding energies and forces. Once training is complete, the MLFF can rapidly predict the energies and forces of similar structures. The computational cost of MLFF prediction scales linearly with the number of atoms, making the cost of relaxation manageable even for very large structures.

However, constructing a comprehensive dataset can be a time-consuming endeavor. Directly using ab-initio MD simulations to build datasets is a relatively inefficient approach, as structures that are close in time within an MD trajectory are very similar. This similarity results in a redundancy that offers little added value to the training dataset, posing a challenge for efficient MLFF deployment.

On-the-fly MLFF approaches like DP-GEN[81] and the MLFF module of the Vienna Ab initio Simulation Package (VASP)[39,82] provide effective solutions for managing computational costs in MD simulations. This article focuses on the MLFF module within VASP. This module automates the process of data collection, MLFF training, and its immediate application to accelerate MD simulations within a continuous loop. The MLFF module operates based on Bayesian linear regression, which allows it to directly estimate the error in its predictions without needing to compare them against ab-initio results. During an MD simulation, if the module estimates a small error, it applies the MLFF-predicted results directly. Conversely, if a large error is estimated, it discards these results and performs a density functional theory (DFT) step to obtain accurate data. This ab-initio data is then added to the training dataset for refining the MLFF. The error criteria also evolve with the averaged predicted error. The initial error criteria in our calculations are set to 2 meV. This iterative process repeats throughout the MD simulation, allowing for extensive sampling from MD trajectories, which could involve tens of thousands of steps, while only requiring DFT calculations for several hundred steps. In our test cases, around two hundred DFT steps were computed during the 10,000-step molecular dynamics simulation. As a result, a high-quality dataset can be constructed with minimal computational expense, optimizing both resources and time.

The MLFF algorithm in VASP is designed to be relatively lightweight, which significantly reduces the training time required during the simulation loop. However, this streamlined approach means that the accuracy of the VASP MLFF may not rival that of more complex neural network-based MLFF algorithms. Consequently, we first utilize VASP MLFF and collect only the DFT data generated in this iteration, subsequently employing a more accurate neural network-based MLFF to fit the collected DFT dataset.

One such advanced approach is NequIP, a machine learning force field based on an E(3)-equivariant graph neural network[41]. This method ensures covariance among the inputs, outputs, and hidden layers, leading to enhanced data efficiency and model accuracy. Another notable E(3)-equivariant algorithm is Allegro, which is particularly well-suited for large structures and optimized for parallel computing[40]. While this article primarily focuses on the application of Allegro, the dataset generated using our approach is versatile and can be employed to train other MLFF models as well. This flexibility facilitates the exploration and application of various advanced MLFF techniques in computational material science.

Data availability

MLFFs for TMDs can be accessed in the Electronic Laboratory for Material Science https://in.iphy.ac.cn/eln/link.html#/125/F3p0.

Code availability

The code for DPmoire is publicly available on GitHub at this link https://github.com/JiaxuanLiu-Arsko/DPmoire.

Received: 7 January 2025; Accepted: 17 July 2025;
Published online: 01 August 2025

References

1. Cao, Y. et al. Unconventional superconductivity in magic-angle graphene superlattices. *Nature* **556**, 43–50 (2018).
2. Cao, Y. et al. Correlated insulator behaviour at half-filling in magic-angle graphene superlattices. *Nature* **556**, 80–84 (2018).
3. Xie, Y. et al. Fractional Chern insulators in magic-angle twisted bilayer graphene. *Nature* **600**, 439–443 (2021).
4. Bistritzer, R. & MacDonald, A. H. Moiré bands in twisted double-layer graphene. *Proc. Natl. Acad. Sci.* **108**, 12233–12237 (2011).
5. Liu, J., Ma, Z., Gao, J. & Dai, X. Quantum valley Hall effect, orbital magnetism, and anomalous Hall effect in twisted multilayer graphene systems. *Phys. Rev. X* **9**, 031021 (2019).
6. Liu, J. & Dai, X. Theories for the correlated insulating states and quantum anomalous Hall effect phenomena in twisted bilayer graphene. *Phys. Rev. B* **103**, 035427 (2021).
7. Lu, X. et al. Superconductors, orbital magnets and correlated states in magic-angle bilayer graphene. *Nature* **574**, 653–657 (2019).
8. Kerelsky, A. et al. Maximized electron interactions at the magic angle in twisted bilayer graphene. *Nature* **572**, 95–100 (2019).
9. Jiang, Y. et al. Charge order and broken rotational symmetry in magic-angle twisted bilayer graphene. *Nature* **573**, 91–95 (2019).
10. Serlin, M. et al. Intrinsic quantized anomalous Hall effect in a moiré heterostructure. *Science* **367**, 900–903 (2020).
11. Xie, Y. et al. Spectroscopic signatures of many-body correlations in magic-angle twisted bilayer graphene. *Nature* **572**, 101–105 (2019).
12. Shen, C. et al. Correlated states in twisted double bilayer graphene. *Nat. Phys.* **16**, 520–525 (2020).
13. Cai, J. et al. Signatures of fractional quantum anomalous Hall states in twisted $MoTe_2$. *Nature* **622**, 63–68 (2023).
14. Jia, Y. et al. Moiré fractional Chern insulators. I. First-principles calculations and continuum models of twisted bilayer $MoTe_2$. *Phys. Rev. B* **109**, 205121 (2024).
15. Wang, C. et al. Fractional Chern insulator in twisted bilayer $MoTe_2$. *Phys. Rev. Lett.* **132**, 036501 (2023).
16. Zeng, Y. et al. Thermodynamic evidence of fractional Chern insulator in moiré $MoTe_2$. *Nature* **622**, 69–73 (2023).
17. Li, T. et al. Quantum anomalous Hall effect from intertwined moiré bands. *Nature* **600**, 641–646 (2021).
18. Wang, L. et al. Correlated electronic phases in twisted bilayer transition metal dichalcogenides. *Nat. Mater.* **19**, 861–866 (2020).
19. Devakul, T., Crépel, V., Zhang, Y. & Fu, L. Magic in twisted transition metal dichalcogenide bilayers. *Nat. Commun.* **12**, 6730 (2021).
20. Leconte, N., Javvaji, S., An, J., Samudrala, A. & Jung, J. Relaxation effects in twisted bilayer graphene: a multiscale approach. *Phys. Rev. B* **106**, 115410 (2022).
21. Shabani, S. et al. Deep moir\'e potentials in twisted transition metal dichalcogenide bilayers. *Nat. Phys.* **17**, 720–725 (2021).
22. Tilak, N., Li, G., Taniguchi, T., Watanabe, K. & Andrei, E. Y. Moiré potential, lattice relaxation, and layer polarization in marginally twisted MoS_2 bilayers. *Nano Lett.* **23**, 73–81 (2023).
23. Carr, S. et al. Relaxation and domain formation in incommensurate two-dimensional heterostructures. *Phys. Rev. B* **98**, 224102 (2018).
24. Nakatsuji, N., Kawakami, T. & Koshino, M. Multiscale lattice relaxation in general twisted trilayer graphenes. *Phys. Rev. X* **13**, 041007 (2023).
25. Jung, J., DaSilva, A. M., MacDonald, A. H. & Adam, S. Origin of band gaps in graphene on hexagonal boron nitride. *Nat. Commun.* **6**, 6308 (2015).

26. Jung, J., Raoux, A., Qiao, Z. & MacDonald, A. H. Ab initio theory of moiré superlattice bands in layered two-dimensional materials. *Phys. Rev. B* **89**, 205414 (2014).
27. Koshino, M. & Nam, N. N. T. Effective continuum model for relaxed twisted bilayer graphene and moiré electron-phonon interaction. *Phys. Rev. B* **101**, 195425 (2020).
28. Nam, N. N. T. & Koshino, M. Lattice relaxation and energy band modulation in twisted bilayer graphene. *Phys. Rev. B* **96**, 075311 (2017).
29. Koshino, M. et al. Maximally localized Wannier orbitals and the extended Hubbard model for twisted bilayer graphene. *Phys. Rev. X* **8**, 031087 (2018).
30. Miao, W., Li, C., Han, X., Pan, D. & Dai, X. Truncated atomic plane wave method for subband structure calculations of moiré systems. *Phys. Rev. B* **107**, 125112 (2023).
31. Stuart, S. J., Tutein, A. B. & Harrison, J. A. A reactive potential for hydrocarbons with intermolecular interactions. *J. Chem. Phys.* **112**, 6472–6486 (2000).
32. Ouyang, W., Mandelli, D., Urbakh, M. & Hod, O. Nanoserpents: graphene nanoribbon motion on two-dimensional hexagonal materials. *Nano Lett.* **18**, 6009–6016 (2018).
33. Kolmogorov, A. N. & Crespi, V. H. Registry-dependent interlayer potential for graphitic systems. *Phys. Rev. B* **71**, 235415 (2005).
34. Naik, M. H., Maity, I., Maiti, P. K. & Jain, M. Kolmogorov-Crespi potential for multilayer transition-metal dichalcogenides: capturing structural transformations in moiré superlattices. *J. Phys. Chem. C.* **123**, 9770–9778 (2019).
35. Nielsen, C. E. M., Da Cruz, M., Torche, A. & Bester, G. Accurate force-field methodology capturing atomic reconstructions in transition metal dichalcogenide moiré system. *Phys. Rev. B* **108**, 045402 (2023).
36. Long, M. et al. An atomistic approach for the structural and electronic properties of twisted bilayer graphene-boron nitride heterostructures. *npj Comput. Mater.* **8**, 73 (2022).
37. Herzog-Arbeitman, J. et al. Moiré fractional Chern insulators. II. First-principles calculations and continuum models of rhombohedral graphene superlattices. *Phys. Rev. B* **109**, 205122 (2024).
38. Haddadi, F., Wu, Q., Kruchkov, A. J. & Yazyev, O. V. Moiré flat bands in twisted double bilayer graphene. *Nano Lett.* **20**, 2410–2415 (2020).
39. Jinnouchi, R., Karsai, F. & Kresse, G. On-the-fly machine learning force field generation: application to melting points. *Phys. Rev. B* **100**, 014105 (2019).
40. Musaelian, A. et al. Learning local equivariant representations for large-scale atomistic dynamics. *Nat. Commun.* **14**, 579 (2023).
41. Batzner, S. et al. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453 (2022).
42. Wang, H., Zhang, L., Han, J. & E, W. DeePMD-kit: a deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.* **228**, 178–184 (2018).
43. Zhang, L., Han, J., Wang, H., Car, R. & E, W. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett.* **120**, 143001 (2018).
44. Zeng, J. et al. DeePMD-kit v2: a software package for deep potential models. *J. Chem. Phys.* **159**, 054801 (2023).
45. Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
46. Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).
47. Schütt, K. T., Sauceda, H. E., Kindermans, P.-J., Tkatchenko, A. & Müller, K.-R. SchNet—a deep learning architecture for molecules and materials. *J. Chem. Phys.* **148**, 241722 (2018).
48. Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B* **99**, 014104 (2019).
49. Park, C. W. et al. Accurate and scalable graph neural network force field and molecular dynamics with direct force architecture. *npj Comput. Mater.* **7**, 73 (2021).
50. Xie, Y., Vandermause, J., Sun, L., Cepellotti, A. & Kozinsky, B. Bayesian force fields from active learning for simulation of inter-dimensional transformation of stanene. *npj Comput. Mater.* **7**, 40 (2021).
51. Chen, C. & Ong, S. P. A universal graph deep learning interatomic potential for the periodic table. *Nat. Comput. Sci.* **2**, 718–728 (2022).
52. Deng, B. et al. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nat. Mach. Intell.* **5**, 1031–1041 (2023).
53. Batatia, I. et al. A foundation model for atomistic materials chemistry. https://doi.org/10.48550/arXiv.2401.00096 (2024).
54. Choudhary, K. et al. Unified graph neural network force-field for the periodic table: solid state applications. *Digit. Discov.* **2**, 346–355 (2023).
55. Chen, C., Ye, W., Zuo, Y., Zheng, C. & Ong, S. P. Graph networks as a universal machine learning framework for molecules and crystals. *Chem. Mater.* **31**, 3564–3572 (2019).
56. Xie, F., Lu, T., Meng, S. & Liu, M. GPTFF: a high-accuracy out-of-the-box universal AI force field for arbitrary inorganic materials. *Sci. Bull.* **69**, 3525–3532 (2024).
57. Liu, X., Peng, R., Sun, Z. & Liu, J. Moiré phonons in magic-angle twisted bilayer graphene. *Nano Lett.* **22**, 7791–7797 (2022).
58. Zhang, X.-W. et al. Polarization-driven band topology evolution in twisted $MoTe_2$ and WSe_2. *Nat. Commun.* **15**, 4223 (2024).
59. Zhang, Y. et al. Universal moiré-model-building method without fitting: application to twisted $MoTe_2$ and WSe_2 https://doi.org/10.48550/arXiv.2411.08108 [cond-mat] (2024).
60. Xu, C., Mao, N., Zeng, T. & Zhang, Y. Multiple Chern bands in twisted $MoTe_2$ and possible non-Abelian states. *Phys. Rev. Lett.* **134**, 066601. https://doi.org/10.1103/PhysRevLett.134.066601 (2024).
61. Mao, N. et al. Transfer learning relaxation, electronic structure and continuum model for twisted bilayer $MoTe_2$. *Commun. Phys.* **7**, 262 (2024).
62. Larsen, A. H. et al. The atomic simulation environment—a Python library for working with atoms. *J. Phys. Condens. Matter* **29**, 273002 (2017).
63. Thompson, A. P. et al. LAMMPS—a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comp. Phys. Comm.* **271**, 108171 (2022).
64. Dion, M., Rydberg, H., Schröder, E., Langreth, D. C. & Lundqvist, B. I. Van der Waals density functional for general geometries. *Phys. Rev. Lett.* **92**, 246401 (2004).
65. Klimeš, J., Bowler, D. R. & Michaelides, A. Van der Waals density functionals applied to solids. *Phys. Rev. B* **83**, 195131 (2011).
66. Klimeš, J., Bowler, D. R. & Michaelides, A. Chemical accuracy for the van der Waals density functional. *J. Phys. Condens. Matter* **22**, 022201 (2010).
67. Berland, K. & Hyldgaard, P. Exchange functional that tests the robustness of the plasmon description of the van der Waals density functional. *Phys. Rev. B* **89**, 035412 (2014).
68. Sabatini, R., Gorni, T. & De Gironcoli, S. Nonlocal van der Waals density functional made simple and efficient. *Phys. Rev. B* **87**, 041108 (2013).
69. Peng, H., Yang, Z.-H., Perdew, J. P. & Sun, J. Versatile van der Waals density functional based on a meta-generalized gradient approximation. *Phys. Rev. X* **6**, 041005 (2016).
70. Ning, J. et al. Workhorse minimally empirical dispersion-corrected density functional with tests for weakly bound systems: r^2 SCAN + rVV10. *Phys. Rev. B* **106**, 075422 (2022).
71. Hamada, I. van der Waals density functional made accurate. *Phys. Rev. B* **89**, 121103 (2014).

72. Lee, K., Murray, É. D., Kong, L., Lundqvist, B. I. & Langreth, D. C. Higher-accuracy van der Waals density functional. *Phys. Rev. B* **82**, 081101 (2010).
73. Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **27**, 1787–1799 (2006).
74. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).
75. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465 (2011).
76. Tkatchenko, A. & Scheffler, M. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **102**, 073005 (2009).
77. Gould, T., Lebégue, S., Ángyán, J. G. & Bučko, T. A fractionally ionic approach to polarizability and van der Waals many-body dispersion calculations. *J. Chem. Theory Comput.* **12**, 5920–5930 (2016).
78. Gould, T. & Bučko, T. C_6 coefficients and dipole polarizabilities for all atoms and many ions in rows 1-6 of the periodic table. *J. Chem. Theory Comput.* **12**, 3603–3613 (2016).
79. Kim, H., Choi, J.-M. & Goddard, W. A. Universal correction of density functional theory to include London dispersion (up to Lr, element 103). *J. Phys. Chem. Lett.* **3**, 360–363 (2012).
80. Steinmann, S. N. & Corminboeuf, C. Comprehensive benchmarking of a density-dependent dispersion correction. *J. Chem. Theory Comput.* **7**, 3567–3577 (2011).
81. Zhang, Y. et al. DP-GEN: a concurrent learning platform for the generation of reliable deep learning based potential energy models. *Comput. Phys. Commun.* **253**, 107206 (2020).
82. 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).

Acknowledgements

J.L. thanks Hanqi Pi for helpful discussions. This work was supported by the Science Center of the National Natural Science Foundation of China (Grant no. 12188101), the National Key R&D Program of China (Grant no. 2023YFA1607400, 2024YFA1408400, 2022YFA1403800), the National Natural Science Foundation of China (Grant nos. 12274436, 11925408, 11921004), and H.W. acknowledges support from the New Cornerstone Science Foundation through the XPLORER PRIZE. The AI-driven experiments, simulations and model training were performed on the robotic AI-Scientist platform of the Chinese Academy of Science.

Author contributions

Q.W. and H.W. supervised and conceptualized the project. J.L. conducted the calculations, developed the code, and drafted the initial manuscript. J.L., Z.F., H.W., and Q.W. participated in discussions and contributed to the refinement of 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/s41524-025-01740-0.

Correspondence and requests for materials should be addressed to Hongming Weng or Quansheng Wu.