

Digital Discovery

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: Y. Yun, D. Park, J. Choi, D. S. Choi and Y. Jung, *Digital Discovery*, 2025, DOI: 10.1039/D5DD00579E.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

ARTICLE

Enhancing High-Dimensional Neural Network Potentials Accuracy in OLED Systems via Element-Relabeling

Yonghwan Yun^a, Dongmin Park^a, Junyoung Choi^a, Dong Shin Choi^b, Yousung Jung^{*ac}Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Accurate atomistic simulations are essential for understanding organic light-emitting diode (OLED) materials in complex condensed-phase environments. However, *ab initio* methods are computationally demanding, whereas classical force fields often lack sufficient accuracy to represent complex molecular behavior. In this work, we present a high-dimensional neural network potential (HDNNP) framework specifically optimized for complex OLED systems. Detailed atomic-level analysis reveals that local environments, measured by intramolecular neighbor density, serve as a key factor of force prediction heterogeneity. To reduce force prediction errors arising from intrinsic molecular complexity, we introduce a conceptually simple yet effective strategy: relabeling atoms exhibiting large force errors as distinct pseudo-elements. By providing more environment-specific local descriptors, this targeted refinement improves force and total-energy prediction accuracy. Such accuracy gains lead to higher fidelity in the reproduction of key simulation observables. To demonstrate the transferability of the element-relabeling strategy, we applied it to chemically diverse benchmark systems, specifically aspirin and double-walled carbon nanotube. Our approach substantially reduces the accuracy gap with state-of-the-art graph-based models while preserving the computational efficiency of conventional HDNNPs, enabling large scale simulations, exemplified by accurate glass transition temperature prediction.

Introduction

Organic light-emitting diode (OLED) technology has emerged as a prominent innovation in the modern display and lighting industries.¹ Over the past few decades, atomistic simulations have significantly advanced the understanding of key aspects of OLED materials, including molecular configurations and charge carrier mobility.^{2, 3} These simulations play a crucial role in bridging microscopic properties with macroscopic behavior, thus providing an efficient pathway for the design and optimization of OLED devices.⁴ The accuracy and relevance of insights derived from these simulations fundamentally depend on the precision and reliability of the interatomic potentials employed.^{5, 6} Accordingly, ongoing efforts to refine these potentials are essential to enhance the predictive capability of simulation-based approaches.

Classical molecular dynamics (MD) has long served as a foundational tool for investigating molecular behavior.⁷ However, the use of general organic force fields, such as the all-atom optimized potential for liquid simulations (OPLS-AA),^{8, 9} in the study of OLED materials can lead to deficiencies in accurately capturing the specific interactions and electronic properties essential to OLED functionality. These limitations necessitate an extensive parameterization and

validation process, which must be benchmarked against experimental or high-level computational data.^{2, 9-11} Even with rigorous parameterization, accuracy may still be constrained by the intrinsic functional forms of the force field (e.g., bonds, angles, and torsions). On the other hand, *ab initio* molecular dynamics (AIMD) simulations¹², based on first principles, offer higher accuracy in describing the intricate behavior of molecular systems.¹³ Nevertheless, the considerable computational cost associated with AIMD often restricts its applicability to smaller systems and shorter timescales. Predicting certain properties, such as the glass transition temperature, remains prohibitive using *ab initio* methods.^{2, 14}

Machine learning interatomic potentials (MLIPs) offer a promising hybrid approach that integrates the accuracy of *ab initio* methods with the computational efficiency of classical MD.^{15, 16} By learning the potential energy surface (PES) with quantum chemical accuracy, these potentials enable energy evaluations at speeds that are orders of magnitude faster than conventional methods.¹⁷ MLIPs can be classified based on how they encode atomic environments and propagate that information to approximate the PES. Local MLIPs—such as the Behler–Parrinello high-dimensional neural-network potential (HDNNP)¹⁸, Gaussian Approximation Potential (GAP)¹⁹, the spectral neighbor analysis potential (SNAP)²⁰, atomic-cluster expansion (ACE)²¹ and moment-tensor potentials (MTP)²²—encode atomic neighborhoods via compact, symmetry-invariant basis expansions. Their inherent computational efficiency enables large-scale and long-timescale molecular dynamics simulations to be performed on standard computing architectures.²³ In contrast to local-descriptor models, graph-based MLIPs represent molecular structures as atomic graphs, where node embeddings are iteratively refined via

^a Department of Chemical and Biological Engineering, and Institute of Chemical Processes, Seoul National University, 1 Gwa-nak-ro, Gwanak-gu, Seoul 08826, Republic of Korea.

^b Virtual Design Engineering Division, LG Display, 245 LG-ro, Wollongong-myeon, Paju-si, Gyeonggi-do 10845, Republic of Korea

^c Institute of Engineering Research, Seoul National University, 1 Gwa-nak-ro, Gwanak-gu, Seoul 08826, Republic of Korea.

† Electronic supplementary information (ESI) available. See DOI: 10.1039/x0xx00000x



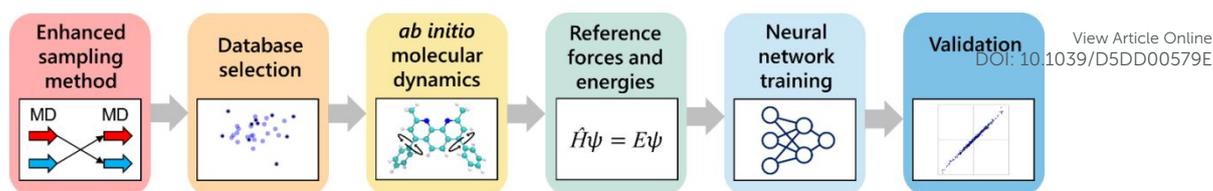


Fig. 1 Schematic workflow for constructing an HDNNP for OLED materials, comprising sampling, data selection, reference calculations, model training, and validation.

message-passing neural networks (MPNNs).²⁴ Widely adopted examples include SchNet²⁵, DimeNet²⁶, GemNet²⁷, SpookyNet²⁸, M3GNet²⁹ and the E(3)-equivariant family comprising NequIP³⁰, MACE³¹ and Equiformer³². Iterative message passing significantly improves predictive accuracy by capturing higher-order many-body correlations within the interaction cutoff. However, repeated neighbor communications introduce substantial computational overhead. Recent equivariant architectures that reduce the depth of message passing (e.g., two-layer higher-order MACE^{31,34,35}) or eliminate it entirely (Allegro³³) have achieved substantial computational speedups while maintaining state-of-the-art accuracy. Despite these advances, comprehensive benchmarks indicate that, on standard CPUs, graph-based MLIPs generally remain slower than local-descriptor models, and comparable throughput typically requires specialized GPU acceleration—thereby underscoring the fundamental trade-off between accuracy and efficiency.³⁶⁻³⁸

Although local-descriptor MLIPs have successfully modeled small gas-phase molecular systems,³⁹⁻⁴¹ their applicability to large π -conjugated molecules, especially in condensed-phase environments, has not been thoroughly explored. In this work, we demonstrate that the predictive accuracy of local descriptor-based HDNNPs can be substantially improved for structurally complex molecular systems, such as condensed-phase OLED molecules, without compromising their intrinsic computational efficiency. Atomic-level analyses reveal a strong correlation between force prediction accuracy and the local density of intra-molecular neighboring atoms, indicating systematically higher force errors for atoms situated in densely coordinated environments. Such structural heterogeneity requires advanced approaches capable of accurately capturing subtle atomic-scale variations. To address this, atoms with large prediction errors were reassigned to distinct pseudo-element labels, allowing descriptors to better capture their unique chemical environments. This strategy overcomes a fundamental limitation of conventional local-descriptor methods, which typically assign a uniform descriptor to all atoms of a given element, irrespective of significant variations in their local chemical environments. With improved atomic force and total-energy predictions, the proposed approach reliably reproduces key structural and thermodynamic properties. Additionally, we demonstrate the generality and broad applicability of the element-relabeling strategy by validating its performance across chemically and structurally diverse molecular systems. The element-relabeled HDNNP substantially narrows the accuracy gap with state-of-the-art graph-based models, while its simpler architectural design facilitates superior parallelization and enhanced computational scalability. Such computational efficiency is particularly advantageous for accurately predicting properties that necessitate extensive sampling over nanosecond timescales and large molecular ensembles, exemplified by calculations of the glass transition temperature.

Methods

Workflow for OLED HDNNP Development

The construction of an HDNNP for OLED materials involves several sequential steps. Fig. 1 provides a schematic overview of the workflow, which includes the following key processes:

- 1. Enhanced Sampling Method:** To efficiently explore a diverse ensemble of condensed-phase configurations, replica exchange molecular dynamics (REMD) simulations are performed using the OPLS-AA classical force field.^{8,9}
- 2. Database Selection:** A representative subset of structures is selected using the farthest point sampling (FPS) algorithm³⁸ to ensure configurational diversity while minimizing redundancy.
- 3. *Ab initio* Molecular Dynamics:** Short AIMD simulations (1ps) based on density functional theory (DFT) are employed to augment the dataset and mitigate limitations inherent to configurations generated via classical MD force fields.
- 4. Reference Forces and Energies:** For each configuration obtained from AIMD simulations, the reference forces and energies for each selected configuration are recomputed using a more rigorous computational framework to ensure high-fidelity training data.
- 5. Neural Network Training:** The neural network is trained to reproduce the reference forces and energies using a composite loss function that balances force and energy contributions.
- 6. Validation:** Finally, model accuracy and generalization are evaluated against independent test sets and reference calculations to validate the predictive performance of the constructed HDNNP. To further assess possible train/test leakage arising from short AIMD trajectories, we also examined a trajectory-wise split in which paired endpoint configurations from the same AIMD trajectory were always assigned to the same subset.



Training Data Generation for OLED HDNNPs via REMD and AIMD

The construction of a robust dataset is a fundamental step in developing any machine learning model. However, collecting training data for HDNNPs tailored to OLED molecules in the condensed phase poses unique challenges. OLED molecules are typically large, conjugated systems with many degrees of freedom, enabling them to adopt a wide range of conformations. In the amorphous condensed phase characteristic of OLED devices, this diversity is further amplified by intermolecular interactions. As a result, the configuration space that needs to be explored expands substantially, making it increasingly difficult to generate a training dataset that captures all relevant molecular conformations.

OLED material configurations are typically simulated using molecular dynamics (MD), where atomistic interactions are represented through classical force fields. However, achieving high accuracy with these force fields demands extensive parameterization.^{2, 9-11} Even with careful parameterization, the attainable accuracy is still constrained by the inherent limitations of predefined functional forms such as bonds, angles, and torsions. In contrast, AIMD computes forces and energies on-the-fly without relying on predefined potentials, making it well suited for systems with complex, dynamic interactions. While this approach captures true physical behavior more accurately, it is computationally intensive, with each *ab initio* calculation generating a single, often highly correlated data point that limits sampling efficiency and introduces redundancy.

In this study, we employ methods from different theoretical levels to construct a comprehensive dataset. First, we utilize REMD using the classical OPLS-AA force field (with only atomic partial charges reparameterized⁴⁴) to broadly sample the configuration space of OLED materials. Also known as parallel tempering, REMD enhances phase-space sampling by enabling exchanges between replicas simulated at different temperatures.^{45, 46} In this work, REMD simulations were conducted with LAMMPS (2 Aug 2023), and the corresponding initial molecular packings were generated with Packmol (v20.14.2). For details on the REMD methodology, refer to the Computational Details section of the Supplementary information (SI).

To assess the difference in sampling regions explored by regular MD and REMD, we projected the resulting structures into a low-dimensional space. The low-dimensional map provides a visually interpretable representation, simplifying the identification of a diverse set of structures.^{41, 47-49} This approach provides an efficient way for conducting comparative analysis. Each structure is encoded by a vector of atom-centered symmetry functions (ACSF), and a two-dimensional projection of the data was obtained through principal component analysis (PCA) implemented using the ASAP⁴⁹ code. As shown in Fig. 2a, REMD samples a significantly broader conformational space than regular MD. Moreover, the dihedral angle distributions of a representative OLED molecule, Bathocuproine (BCP), further confirm this enhanced diversity (Fig. S1). Fig. 2b illustrates that REMD explores a broader range of structural variations across diverse thermodynamic states. To minimize structural redundancy, we employ the FPS algorithm⁴³ to select a representative subset of 1,000 snapshots from a pool of 3,500 initial configurations.

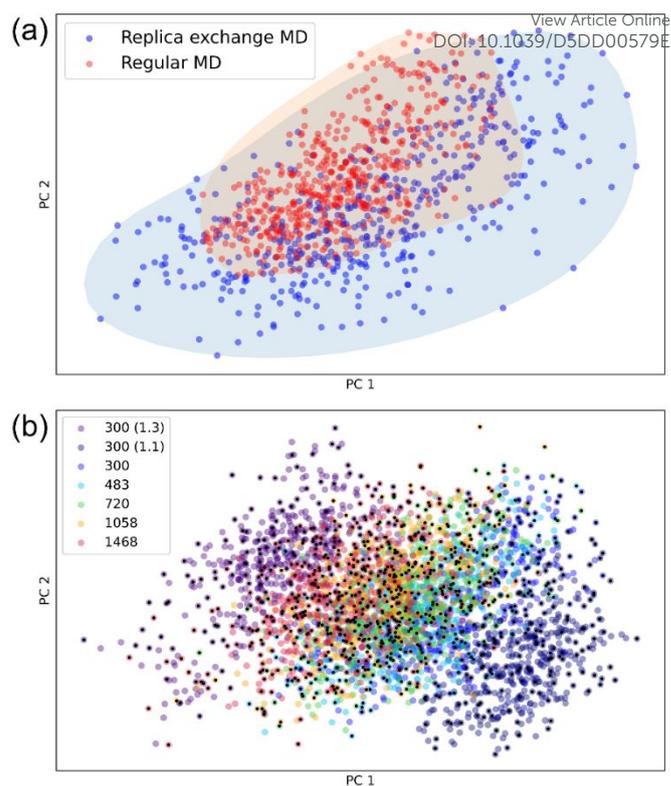


Fig. 2 PCA projections of MD simulation structures, where each point represents the averaged ACSF descriptor of a training set structure. (a) Configuration space comparison between REMD and regular MD simulations at 300 K. (b) REMD simulation structures obtained at temperatures ranging from 300 to 1468 K and densities of 1.1, 1.2, and 1.3 g·cm⁻³. Parenthetical values indicate densities at 300 K, with 1.2 g·cm⁻³ used as default when unspecified. Black dots indicate configurations selected using the FPS algorithm.

Fig. S2 presents the number of snapshots selected by FPS under various temperature and density conditions.

Due to the high degrees of freedom in large OLED molecules, traditional force field parameter fitting becomes increasingly challenging. To overcome potential deficiencies in the dataset obtained from classical MD, we performed short AIMD simulations (1 ps) for each selected snapshot CP2K (v9.1). These AIMD runs capture a broad range of energy distributions during equilibration (Fig. S3) and yield diverse geometric structures, which together enrich the structural diversity of the dataset (Fig. S4). The use of structures obtained from REMD enables the parallel execution of AIMD simulations, with simulation temperatures naturally determined by the REMD protocol. However, due to the high computational cost of performing AIMD on all 1,000 selected structures, we adopted a cost-effective strategy: after a 1 ps AIMD run, the reference energies and forces were recalculated for the initial and final snapshots of each trajectory using a more rigorous computational setup. **Because both endpoint structures from each short AIMD trajectory were incorporated into the reference dataset, we additionally examined whether a configuration-level random split could introduce train/test correlation. To address this possibility, we performed an additional trajectory-wise split in which the paired initial and final configurations from each AIMD trajectory were always assigned to the same subset. The resulting force and energy**



prediction errors were comparable to those obtained from repeated random splits (Fig. S7), indicating that the reported model performance is not driven by inadvertent separation of paired AIMD endpoint structures across the training and test sets. The use of approximately 1 ps-scale structure saving is also consistent with prior dataset-generation practice; for example, Williams et al. reported sampling structures every 1 ps for reference-dataset construction and noted that longer MD trajectories help reduce temporal correlation between sampled structures.⁵⁰ Further details regarding the AIMD simulations, recalculation procedure, and split validation are provided in the Computational Details section of the SI.

Behler-Parrinello Symmetry Functions

We employed the HDNNP framework, originally introduced by Behler and Parrinello.¹⁸ This section provides a concise overview of the functional forms and parameterization of the two primary families of ACSFs employed in this study. Further details regarding the construction and theoretical background of ACSFs are available in previous literature.^{51, 52} Within this framework, the total energy of a system is expressed as the sum of atomic contributions:

$$E = \sum_{j=1}^N E_j(\mathbf{G}_j)$$

where N is the number of atoms and \mathbf{G}_j is the descriptor that characterizes the local environment of atom j . ACSFs, which are rotationally, translationally, and permutationally invariant, serve as local structural descriptors for the input layer of the HDNNP. In this study, both radial (\mathbf{G}_i^2) and angular (\mathbf{G}_i^3) ACSFs were employed, defined as follows:

$$\mathbf{G}_i^2 = \sum_{j \neq i} e^{-\eta(R_{ij}-R_s)} f_c(R_{ij})$$

$$\mathbf{G}_i^3 = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos \theta_{ijk})^\zeta e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk})$$

where R_{ij} is the distance between atoms i and j , and θ_{ijk} is the angle formed by atoms i , j , and k with i as the vertex. The shape of the radial symmetry functions is controlled by the parameters η and R_s , whereas the angular symmetry functions depend on η , ζ , and λ . A cutoff function $f_c(R_{ij})$, defined with a cutoff radius R_c is given by:

$$f_c(R_{ij}) = \begin{cases} \tanh^3 \left(1 - \frac{R_{ij}}{R_c} \right), & (R_{ij} \leq R_c) \\ 0, & (R_{ij} > R_c) \end{cases}$$

An initial pool of 138 ACSFs was generated per element. A subset of ACSFs was selected to optimally represent the diverse atomic environments encountered in the training dataset. Specifically, 113, 111, and 79 functions were selected for hydrogen, carbon, and nitrogen atoms, respectively. Further details on symmetry function selection are provided in the Computational Details section of the SI. A cutoff radius of 12 Bohr (6.35 Å) was applied to define the local atomic environment.

Results and discussion

View Article Online

DOI: 10.1039/D5DD00579E

Force Prediction Error Evaluation at the Atomic Level in HDNNPs

Neural network parameters were optimized using the n2p2 package (v2.2.0)⁵³ to accurately reproduce reference forces and energies. Further details regarding the neural network training protocol are presented in the Computational Details section of the SI. Throughout this study, an ensemble of four HDNNP models with distinct random seeds was employed to ensure robustness and reduce variance in predictions for the OLED system.

Due to their larger size compared to typical organic molecules, OLED molecules frequently contain atoms of the same element occupying distinctly different local atomic environments. The structural

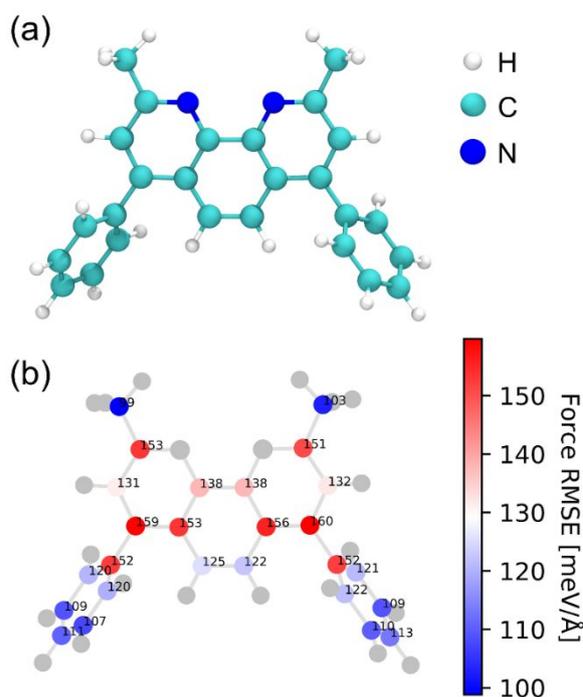


Fig. 3 (a) Chemical structure of the BCP molecule. (b) Spatial distribution of per-atom force RMSEs for carbon atoms. Nitrogen and hydrogen atoms are shown in gray for clarity.

heterogeneity inherent to OLED systems presents challenges for HPNNPs based on local descriptors. Accordingly, force predictions were used to assess atomic-level accuracy rather than global accuracy. Analysis of force errors at the individual-atom level provides detailed insight into the capability of the HPNNP to capture subtle variations in forces across diverse atomic environments. This type of analysis helps pinpoint where the HPNNP systematically underperforms, thereby enabling targeted refinements. Figure 3 shows the distribution of per-atom force root mean square errors (RMSEs) within the carbon framework of a BCP molecule. The per-atom force RMSE distribution spans a broad range (99–160 meV Å⁻¹), reflecting pronounced site-to-site heterogeneity. Our analysis reveals that carbon atoms located in sterically congested regions, such as the central rings of the polyaromatic core, exhibit markedly higher force RMSEs. Conversely, terminal or peripheral atoms situated within sparsely populated environments exhibit comparatively modest errors. In a recent study⁵⁴, specific atomic environments were found to exhibit



significantly elevated force errors. Notably, this heterogeneity was consistently observed across multiple MLIP frameworks, irrespective of whether the underlying models employed kernel-based methods or neural networks.

Local Structure–Error Correlation in OLED Molecules

The present analysis examines the correlation between force RMSE and the number of atoms within a defined cutoff sphere (Figure 4a). This correlation analysis aims to uncover the underlying factors responsible for the observed variations in force RMSE distributions. At a 3 Å cutoff radius, RMSE values increase monotonically with the number of neighboring atoms for both carbon and hydrogen atoms (Figure 4b), suggesting a strong positive correlation between local atomic density and force prediction error. This observation reveals that local structural complexity, as measured by neighbor density, is a key factor in force prediction errors, with highly coordinated atomic sites exhibiting higher RMSE values. To quantify this relationship more systematically, Pearson correlation coefficients were computed as a function of cutoff distance, with specific focus on carbon atoms (Figure 4c). In this analysis, atoms within each cutoff sphere were classified into two distinct groups: intramolecular atoms (belonging to the same molecule) and total atoms (including both intra- and intermolecular neighbors). This classification enables a more detailed assessment of how each neighbor group contributes to force prediction accuracy. As shown in Figure 4c, both correlation curves reach their maximum near 3 Å, indicating that short-range neighborhoods are most informative for predicting force errors. Notably, intramolecular neighbor density exhibits a consistently stronger correlation with force errors compared to the total atomic density within the same cutoff radius. This stronger correlation observed for intramolecular neighbors likely results from inherently stronger covalent interactions compared to weaker intermolecular forces. These findings imply that, even in the condensed phase, force-prediction heterogeneity is primarily determined by the intrinsic complexity of molecular structure. These observations further suggest that simple structural descriptors, such as intramolecular coordination number or local neighbor density, may serve as practical predictive criteria for identifying candidate atoms for relabeling. In this study, however, we retained the diagnostic error-based procedure, as it directly identifies the atomic environments associated with the largest residual errors in the baseline model and thus provides the most conservative basis for atom selection. Figure 4d presents the relative proportions of intra- and intermolecular neighbors as a function of cutoff distance, revealing the expected decay of intramolecular influence at longer distances.

Element Relabelling for Capturing Structural Complexity in HDNNP

Addressing the structural-complexity-driven heterogeneity in force-prediction errors requires descriptors that accurately capture subtle variations within local atomic environments. However, modifying the descriptor architecture on a per-atom basis is infeasible, as standard neural network imposes a fixed functional form and uniform parameterization across all atoms of a given element. Although each descriptor value changes with geometry, the use of a fixed functional form may inadequately represent the diversity of local atomic environments. A straightforward yet effective approach is to

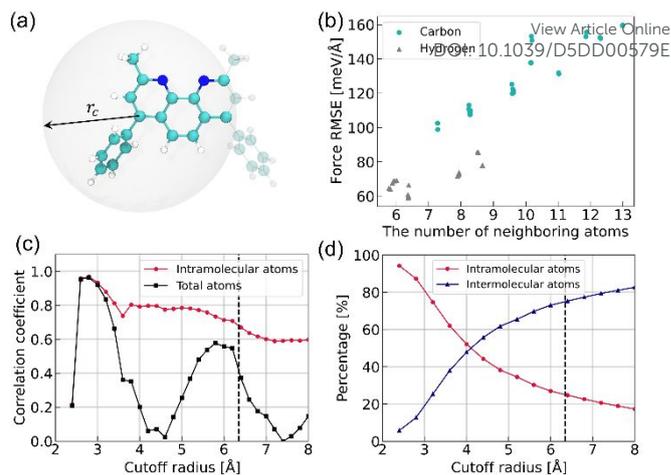


Fig. 4 (a) Schematic illustration of the BCP molecule highlighting local atomic environments defined by cutoff spheres. (b) Correlation between the number of neighboring atoms within a 3 Å cutoff sphere (x-axis) and the force RMSE (y-axis) for carbon and hydrogen atoms. (c) Pearson correlation coefficient between the number of atoms within the cutoff sphere and the force RMSE for carbon atoms as a function of cutoff distance. "Total atoms" refers to the combined intra- and intermolecular contributions. (d) Percentage of intramolecular and intermolecular atoms among total atoms within the cutoff sphere for carbon atoms, as a function of cutoff distance.

reclassify as a pseudo-element "X," thereby enabling the design of descriptors that capture environment-specific features. As illustrated in Figure 5, relabeling specific carbon atoms as pseudo-element X requires modifications to their symmetry function vectors and corresponding atomic neural network parameters. Specifically, introducing pseudo-element X requires defining an additional atomic neural network and constructing symmetry functions to explicitly encode interactions between X and hydrogen, carbon, and nitrogen atoms. In practice, this approach involves simply relabeling selected atoms in the dataset by changing their elemental labels in the position and force datasets, and incorporating the newly defined symmetry functions. This method was applied to the ten carbon atoms within the BCP molecule that exhibited the highest force RMSE values. These atoms are characterized by a high density of intramolecular neighbors and a lack of covalent bonds to hydrogen (Figure 3b), representing structurally challenging environments for force prediction. Here, the pseudo-element X is not intended to represent a new chemical species, but rather a label for a distinct local-environment class among atoms that remain chemically carbon. This relabeling therefore provides a separate descriptor-to-network mapping for these crowded carbon environments, rather than forcing all carbon atoms into a single shared representation. Initially, standard sets of symmetry functions were



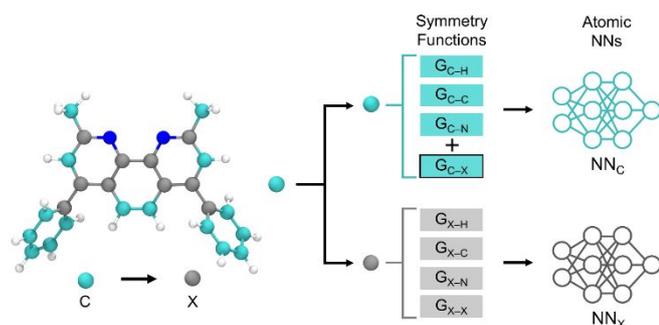


Fig. 5 Schematic representation of the modifications in symmetry function vectors and the corresponding atomic neural networks (NNs) following the relabeling of carbon atoms as the pseudo-element "X". For clarity, only two representative components of the radial symmetry function vector are shown in the schematic; angular symmetry function components (typically three per triplet interaction) are omitted for simplicity.

utilized for each relevant elemental interaction. To mitigate the combinatorial increase in feature dimensionality resulting from additional element types, symmetry functions exhibiting negligible variability across the dataset were eliminated. Specifically, symmetry functions with dataset ranges (maximum–minimum values) below 0.01 were discarded. Further methodological details are provided in the Computational Details section of the SI.

Data-Driven Element-Relabeling Improves HDNNP Accuracy

Figure 6 shows that element relabeling markedly improves atomic force predictions: the mean force RMSE for carbon reduced from 129.5 to 98.0 $\text{meV}\cdot\text{\AA}^{-1}$, while its standard deviation reduced from 33.5 to 23.4 $\text{meV}\cdot\text{\AA}^{-1}$. The reduction in RMSE variability reflects decreased heterogeneity in force predictions, indicating that the element-relabeled HDNNP attains more uniform accuracy across chemically diverse atomic environments. Such improved uniformity is essential for ensuring energy conservation and maintaining stability in long-timescale MD simulations. These gains are transferable, lowering the mean force RMSE for hydrogen (69.3 \rightarrow 53.7 $\text{meV}\cdot\text{\AA}^{-1}$) and nitrogen (137.4 \rightarrow 97.0 $\text{meV}\cdot\text{\AA}^{-1}$) even though only carbon atoms were relabeled. This transferability arises because relabeling carbon atoms as the pseudo-element X modifies not only the descriptors of the relabeled sites themselves, but also those of neighboring atoms through the introduction of X-containing interaction channels in the ACSF representation. Consequently, hydrogen and nitrogen atoms located near relabeled carbon sites are described by more environment-specific descriptors, and this improved local representation is further reflected in the force predictions because atomic forces are obtained as derivatives of the total HDNNP energy. Because the HDNNP represents the total potential energy as an explicit sum of atom-centered contributions, improvements in per-atom force predictions naturally propagate to the global scale. Accordingly, the total energy RMSE decreased from 419.5 to 293.6 meV, confirming that the relabeled model yields more accurate and internally consistent energetics. This targeted, data-driven relabeling strategy, guided by atomic-level force-error analysis, proved more effective than conventional heuristic approaches, such as differentiation based on hybridization states (e.g., sp^2 vs. sp^3), typically employed in force-field parameterizations (Figure S5).

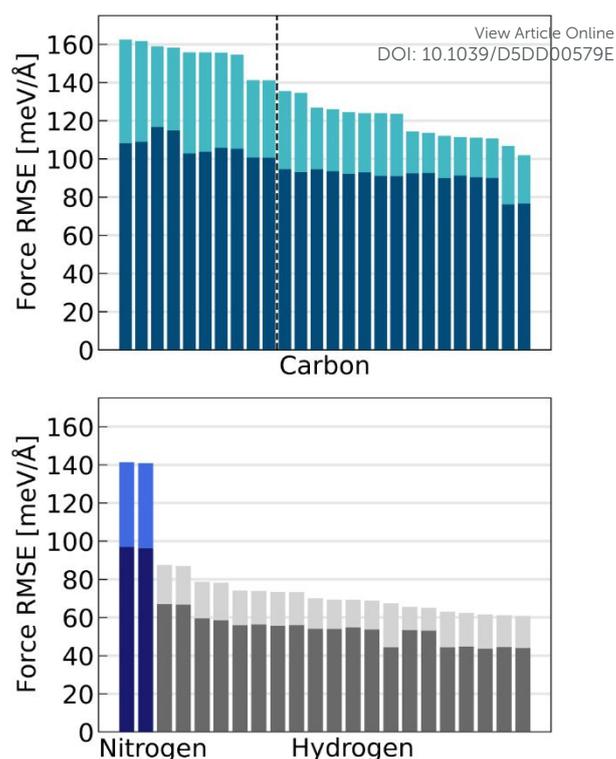


Fig. 6 Individual atomic force RMSE variations before and after carbon relabeling with pseudo element 'X'. Initial RMSE values, arranged by magnitude, are depicted using bar colors corresponding to the standard atomic colors for carbon, nitrogen, and hydrogen, while dark-colored bars represent the RMSE values after relabeling. In the upper panel, a dashed line separates the atoms; those to the left of the line correspond to sites where carbon was relabeled as 'X'.

These results underscore the advantages of employing data-driven, error-based strategies to systematically enhance the predictive accuracy of HDNNPs.

Validation Beyond Force and Energy: RDF and PMF Comparisons

While improvements in atomic force and total-energy predictions are beneficial, such refinements alone may not ensure overall fidelity in MD simulations. Thus, rigorous validation of MLIPs should extend beyond force and energy accuracy to include their ability to accurately reproduce physically relevant simulation observables.^{56, 57} Such comprehensive validation requires assessing the accuracy of simulation observables predicted by HDNNPs with respect to AIMD reference data. To this end, we conducted a comparative analysis of structural and thermodynamic properties derived from 60-ps trajectories generated by AIMD, baseline HDNNP, and element-relabeled HDNNP. Specifically, we employed a quantitative evaluation framework that measures deviations from AIMD reference trajectories using percentage-based accuracy metrics.⁵⁵ Structural accuracy was evaluated by computing radial distribution functions (RDFs) with respect to the OLED molecule's center of mass, while molecular flexibility was characterized through potentials of mean force (PMFs) as a function of the dihedral angle ϕ . The results of these validation analyses are summarized in Figure 7. Our analysis demonstrates that the element-relabeling strategy enhances the



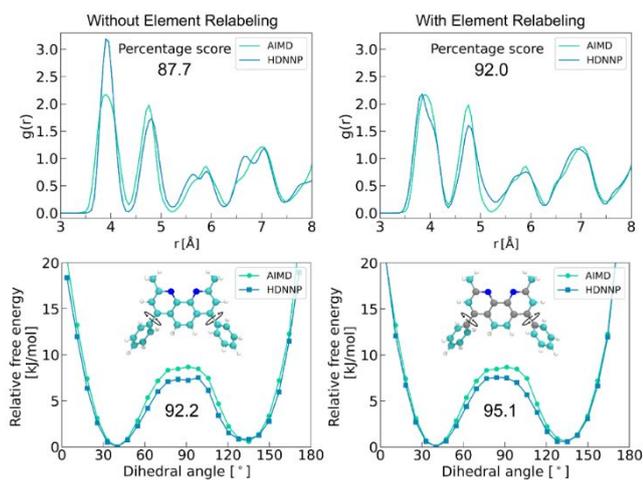


Fig. 7 RDFs and PMFs obtained from MD trajectories using the baseline HDNNP (left) and the element-relabeled HDNNP (right). For quantitative assessment, these properties were converted to numerical scores using established difference metrics (see ref. 47 for details). The resulting evaluation scores are shown in the figure for direct comparison.

reproduction of key simulation-derived properties, such as RDFs and PMFs, in addition to improving the accuracy of both atomic force and total energy predictions.

Application to Other Systems

To evaluate the transferability and robustness of the element-relabeling strategy, we applied it to two chemically and structurally distinct systems: aspirin (MD17 dataset⁵⁸) and double-walled carbon nanotube (MD22 dataset⁵⁴). This extension demonstrates that our approach is not limited to specific molecular structures but can be generalized across diverse chemical environments and molecular systems. The accuracy improvements for these two systems are presented in Figure 8. To benchmark our approach against graph-based models, all prediction errors were recalculated in terms of mean absolute error (MAE) for direct comparison with EGraFFBench results⁵⁷. For aspirin, the baseline HDNNP yielded energy and force MAEs of 12.31 meV and 31.45 meV·Å⁻¹, respectively. Analysis of the force error distribution identified the three carbon atoms with high MAEs, which were subsequently relabeled. Application of the element-relabeling strategy lowered the MAEs, 12.31 → 6.61 meV for energy and 31.45 → 19.79 meV·Å⁻¹ for force. When compared with the median error values of eight equivariant GNNs in EGraFFBench (7.4 meV and 13.95 meV·Å⁻¹, respectively⁵⁷), the relabeled HDNNP achieved comparable accuracy, indicating that element relabeling markedly enhances the performance of conventional HDNNPs and narrows the performance gap relative to state-of-the-art GNNs (Table 1). Similarly, for the double-walled carbon nanotube, energy and force MAEs were reduced from 146.0 meV and 88.93 meV·Å⁻¹ to 113.5 meV and 61.76 meV·Å⁻¹, respectively, although a different strategy was used. We first confirmed that a force-error-based selection also reduces the error in the nanotube system, indicating that it provides a consistent and effective starting criterion across the systems examined here. However, for the highly symmetric DWCNT system, relabeling the carbon atoms in the outer shell proved even more effective, as it

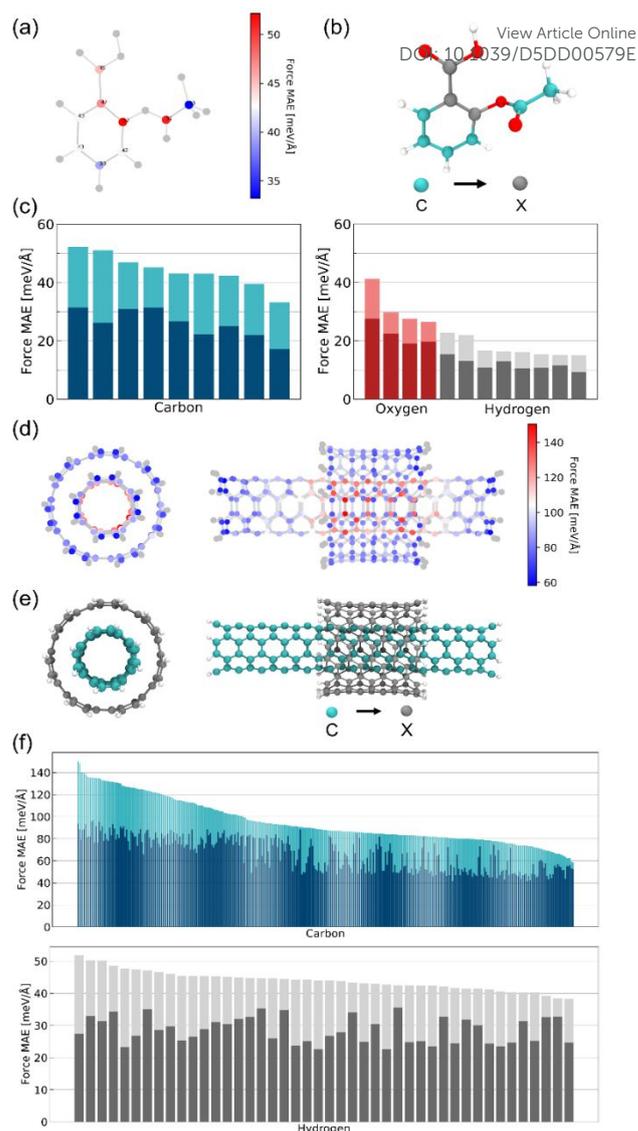


Fig. 8 (a) Spatial distribution of per-atom force MAEs for carbon atoms in aspirin. Oxygen and hydrogen atoms are shown in gray for clarity. (b) Schematic illustration depicting the relabeling of three carbon atoms with high MAEs as the pseudo-element 'X'. (c) Individual atomic force MAE before and after relabeling selected carbon atoms as pseudo-element 'X'. Initial MAE values, arranged by magnitude, are depicted using bar colors corresponding to the standard atomic colors, while dark-colored bars represent the MAE values after relabeling. (d) Spatial distribution of per-atom force MAEs for carbon atoms in a double-walled nanotube with hydrogen atoms depicted in gray for clarity. (e) Schematic illustration depicting the relabeling of carbon atoms in the outer nanotube as the pseudo-element 'X'. (f) Individual atomic force MAE before and after relabeling outer-shell carbon atoms as pseudo-element 'X'. Initial MAE values, arranged by magnitude, are depicted using bar colors corresponding to the standard atomic, while dark-colored bars represent the MAE values after relabeling.

provides a cleaner structural partition of distinct local environments (see Fig. S8). The consistent accuracy improvements observed across chemically and structurally diverse systems indicate that the element-relabeling strategy offers a generalizable framework for systematically improving the performance of local-descriptor-based HDNNPs.



Table 1 Energy (E) and force (F) MAEs, given in meV and meV·Å⁻¹, respectively, for models trained on the aspirin system. Bold and italicized values indicate the best and second-best results, respectively, for both energy and force. DOI: 10.1039/D5DD00579E

NequiP		Allegro		BOTNet		MACE		Equiformer		TorchMDNet		PaiNN		DimeNET++		Baseline HDNNP		Relabeled HDNNP	
E	F	E	F	E	F	E	F	E	F	E	F	E	F	E	F	E	F	E	F
6.84	13.8 9	5.00	9.17	7.99	14.0 6	8.53	14.0 1	6.15	15.2 9	5.33	8.97	41.4 9	12.4 1	133. 24	22.0 7	12.3 1	31.4 5	6.61	19.7 9

Local Descriptors vs. Graph-based Models: Trade-offs in Accuracy and Efficiency

Graph-based MLIPs achieve remarkable accuracy by explicitly propagating information across neighboring atomic environments.²⁴ However, this often entails computational overhead from message passing^{26,27} and from equivariant tensor-product contractions.^{28,60} Message-passing architectures employing angular/triplet interactions (e.g., DimeNet²⁶, GemNet²⁷) scale nearly quadratically with the number of neighbors, since each central atom forms $k(k-1)/2$ triplets, yielding an overall $O(Nk^2)$ complexity (N : total atoms; k : average neighbors within cutoff r_c) that limits their applicability to large condensed-phase systems. Although Allegro reduces some overhead by enforcing strict cutoffs and eliminating explicit message passing, it remains substantially slower than descriptor-based HDNNPs on equivalent CPU hardware.^{33,36}

In contrast, the local descriptor-based HDNNP is inherently more computationally efficient. By leveraging ACSFs, the framework enables efficient parallelization and exhibits near-linear scaling with system size, facilitating simulations at the nanosecond timescale or at the billion-atom scale.^{61,62} The present element-relabeling strategy should be situated within this local-descriptor paradigm rather than among hybrid or equivariant architectures. Specifically, it does not introduce message passing or equivariant tensor operations, but instead increases representational flexibility at the descriptor level by assigning distinct pseudo-element labels to atoms associated with systematically underrepresented local environments. In this sense, the method may be viewed as a descriptor-level refinement of the HDNNP framework that improves the resolution of local structural heterogeneity while retaining the practical efficiency advantages of descriptor-based models. We further mitigated ACSF combinatorial growth by eliminating symmetry functions with negligible variability and retaining informative descriptors. Our approach reduces the accuracy gap with state-of-the-art graph-based models while maintaining high computational efficiency, which is particularly advantageous for simulations over extended timescales and large system sizes. To make this point more explicit, we summarize the descriptor growth and the effect of pruning in Table S3 of the Supporting Information; in the present case, the total number of symmetry functions increases from 414 to 928 upon introducing X, but pruning reduces this to 619, indicating that a substantial fraction of the formally introduced channels is not informative in practice. At the same time, this mitigation does not remove the intrinsic combinatorial growth of ACSF-based descriptors.

Leveraging the Efficiency of Local Descriptors for Large-Scale Molecular Simulations

To evaluate the practicality and scalability of the element-relabeling HDNNP in OLED applications, we applied it to predict the glass transition temperature (T_g). Predicting T_g of amorphous OLED materials represents a demanding task, requiring nanosecond-scale sampling and extensive molecular ensembles. Previous studies have primarily relied on classical MD simulations to meet these computational demands.^{2,44} The present approach aims to deliver DFT-level accuracy by leveraging the computational efficiency of HDNNPs. The initial dataset consisted exclusively of high-density configurations (1.1, 1.2, and 1.3 g·cm⁻³). To broaden configurational diversity required for T_g prediction, additional low-density (0.7 and 0.9 g·cm⁻³) and high-temperature (825 and 1202 K) structures were incorporated. Each density-temperature combination contributed 100 structural snapshots, resulting in a total of 400 configurations. Glass transition temperatures were predicted using the computational protocol established by Lin et al.⁴² The simulation protocol began with classical equilibration of 128 BCP molecules (6,144 atoms), followed by a 2.5 ns NPT equilibration at 800 K and 1 atm performed using the element-relabeled HDNNP. The equilibrated system was subsequently subjected to linear cooling from 800 K to 0 K at a rate of 50 K·ns⁻¹, during which density and temperature were continuously recorded. To ensure robust T_g estimation, linear regressions were performed on density-temperature data across three temperature intervals (200, 250, and 300 K). Despite pronounced density fluctuations attributable to the small system size (128 molecules), high coefficients of determination ($R^2 > 0.9$) confirmed the robustness of the linear regressions within the selected temperature intervals (Figure S6). Although the estimated glass transition temperatures varied slightly depending on the regression interval (200, 250, and 300 K), the overall trends remained consistent. The resulting T_g values were 116.6 ± 12.7 °C for HDNNP-MD and 152.3 ± 7.2 °C for classical MD, reported as the mean and standard deviation of estimates obtained from three distinct temperature intervals (200, 250, and 300 K). The T_g predicted by HDNNP-MD showed closer agreement with the experimental value (89 °C),⁶³ indicating improved accuracy over classical force fields. Throughout the cooling process, HDNNP-MD trajectories consistently exhibited higher bulk densities by 0.02 to 0.25 g·cm⁻³ than those obtained from classical MD simulations. The higher densities observed in HDNNP-MD trajectories suggest tighter molecular packing relative to classical MD. One possible contributor is differences in the treatment of short-range exchange (Pauli) repulsion: HDNNP inherits this from the underlying electronic-structure data, whereas many classical force fields rely on Lennard-Jones-type repulsion, whose short-range behavior can be inadequate for some systems.⁶⁴ These findings demonstrate that the element-



relabeling HDNNP can be effectively utilized for predicting properties derived from large-scale MD simulations, with improved accuracy over classical potentials.

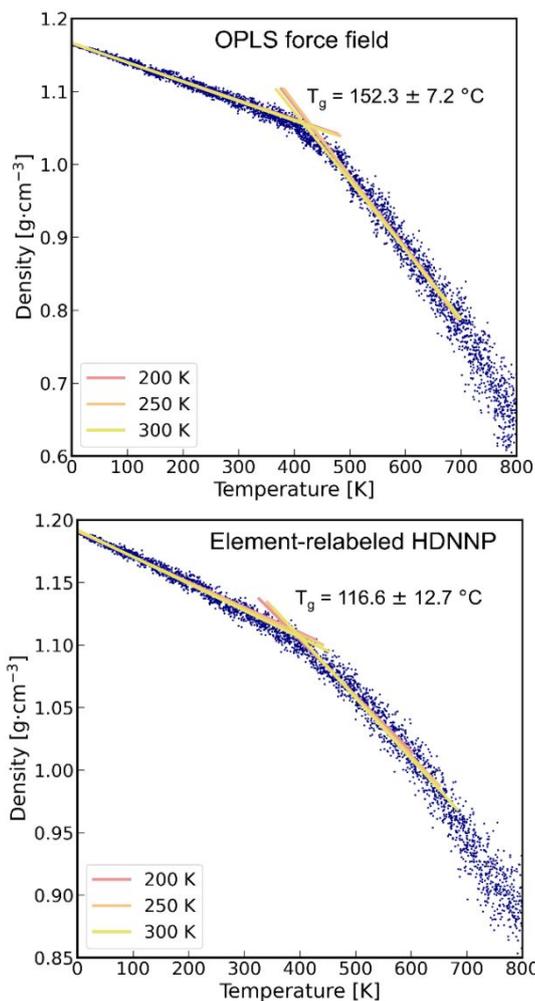


Fig. 9 Density–temperature plots obtained from the OPLS-AA force field and the element-relabeling HDNNP, shown on the top and bottom panels, respectively. Blue data points indicate instantaneous values of bulk density obtained during linearly controlled cooling simulations. Colored solid lines denote linear least-squares fits applied within distinct temperature windows (200 K, 250 K, and 300 K). The intersection of linear regressions in the high- and low-temperature regimes defines the glass transition temperature for each simulation method.

Conclusions

In this study, we presented a HDNNP framework specifically designed to address the computational challenges associated with structurally complex OLED systems. Atomic force analysis reveals that force errors are strongly correlated with intramolecular neighbor density, increasing consistently with higher coordination numbers. These insights motivated us to implement an element-relabeling strategy, relabeling high-error atoms as pseudo-elements to provide environment-specific descriptors. The accuracy of the HDNNP was improved using an element-relabeling strategy designed to overcome the limitations

inherent to single-descriptor approaches when applied to diverse local environments. Relative to conventional approaches, the developed HDNNP framework demonstrated superior accuracy in predicting atomic forces and total-energy and reproducing essential structural and thermodynamic characteristics of OLED materials. Application across chemically and structurally diverse systems confirmed the generality of the proposed framework. This strategy preserves the computational efficiency of HDNNPs while achieving accuracy comparable to graph-based MLIPs, enabling large-scale materials simulations such as glass transition temperature prediction. This represents a significant advancement in the modeling of OLED materials. Nonetheless, one important limitation should be noted. Although descriptor pruning substantially reduces the number of formally introduced symmetry functions, the present approach does not eliminate the intrinsic combinatorial growth of ACSF-based descriptors. As additional relabeled classes and elemental species are introduced, the resulting increase in descriptor dimensionality, memory demand, and training cost will become increasingly significant. The method is therefore best regarded as a targeted and practically useful refinement strategy when only a limited number of distinct environment classes require separate treatment.

Author contributions

Y. Y. and Y. J. conceptualized the project and acquired funding. Y. Y. curated the dataset, developed the neural network potential models, conducted molecular dynamics simulations, and analysed the results. Y. Y. also drafted the initial manuscript and generated the figures. D. P., J. C., and D. C. reviewed and edited the manuscript. Y. J. provided critical revisions and supervised the project.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data and code relevant to this study are deposited in Zenodo [DOI:10.5281/zenodo.19078768]. The deposit includes the raw HDNNP reference datasets in the n2p2 input.data format; the explicit train/test split files corresponding to each trained model; the deposited input.nn files specifying the random_seed values used for the internal n2p2 split; the trained model artifacts (scaling.data and trained weight files) for the OLED reference and relabeled committees and for the aspirin and DWCNT reference/relabeling models; the FPS selection script; additional dataset-preparation scripts; and representative CP2K input files (aimd.inp and endpoint_recalculation.inp). Additional methodological details and supplementary figures are provided in the ESI. The external benchmark datasets used in this study are publicly available. The aspirin benchmark was taken from the revised MD17 (rMD17) dataset on Figshare (DOI: 10.6084/m9.figshare.12672038.v3), and the double-walled



carbon nanotube benchmark was obtained from the MD22 dataset portal hosted by sGDML (accessed March 10, 2026).

Acknowledgements

This work was supported by the New Faculty Startup Fund from Seoul National University and LG Display (C2021004543).

Notes and references

- B. Geffroy, P. Le Roy, and C. Prat, *Polymer International* 2006, **55** (6), 572-582.
- A. Mondal, L. Paterson, J. Cho, K. Lin, B. van der Zee, G. Wetzelaer, A. Stankevych, A. Vakhnin, J.-J. Kim and A. Kadashchuk, *Chemical Physics Reviews* 2021, **2** (3).
- L. Paterson, F. May, and D. Andrienko, *Journal of Applied Physics* 2020, **128** (16).
- Y. Qiu, L. Wang, L. Chen, Y. Tian, Z. Zhou and J. Wu, *Chem. Sci.*, 2025, **16**, 17450–17460.
- F. Vitalini, A. Mey, F. Noé and B. G. Keller, *The Journal of Chemical Physics* 2015, **142** (8).
- T. Lewis-Atwell, P. Townsend and M. N. Grayson, *Tetrahedron*, 2021, **79**, 131865.
- S. Hollingsworth and R. O. Dror, *Neuron*, 2018, **99** (6), 1129-1143.
- W. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *Journal of the American Chemical Society*, 1996, **118** (45), 11225-11236.
- G. Kaminski, R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen, *The Journal of Physical Chemistry B*, 2001, **105** (28), 6474-6487.
- M. Moral, W. Son, J. Sancho-Garcia, Y. Olivier and L. Muccioli, *Journal of Chemical Theory and Computation*, 2015, **11** (7), 3383-3392.
- C. Tonnelé, M. Stroet, B. Caron, A. Clulow, R. C. Nagiri, A. K. Malde, P. L. Burn, I. R. Gentle, A. E. Mark and B. J. Powell, *Angewandte Chemie International Edition*, 2017, **56** (29), 8402-8406.
- D. Marx and J. Hutter, Cambridge University Press, 2009, 1–520.
- Y. Fu, L. Bernasconi, and P. Liu, *Journal of the American Chemical Society*, 2021, **143** (3), 1577-1589.
- Y. Qiu, X. Zhang, Y. Tian and Z. Zhou, *Chinese Journal of Structural Chemistry*, 2023, **42**, 100118.
- P. Friederich, F. Häse, J. Proppe, and A. Aspuru-Guzik, *Nature Materials*, 2021, **20** (6), 750-761.
- O. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko and K.-R. Müller, *Chemical Reviews*, 2021, **121** (16), 10142-10186.
- V. Deringer, M. A. Caro, G. Csányi, *Advanced Materials*, 2019, **31** (46), 1902765.
- J. Behler and M. Parrinello, *Physical Review Letters*, 2007, **98** (14), 146401.
- A. Bartók, M. C. Payne, R. Kondor and G. Csányi, *Physical Review Letters*, 2010, **104** (13), 136403.
- A. Thompson, L. P. Swiler, C. R. Trott, S. M. Foiles and G. J. Tucker, *Journal of Computational Physics*, 2015, **285**, 316-330.
- R. Drautz, *Physical Review B*, 2019, **99**, 014104.
- A. Shapeev, *Multiscale Modeling & Simulation*, 2016, **14** (3), 1153-1173.
- A. P. Thompson, H. Aktulga, R. Berger, D. Bolintineanu, W. Brown, P. Crozier, P. Veld, A. Kohlmeyer, S. Moore, T. Nguyen, R. Shan, M. Stevens, J. Tranchida, C. Trott and S. Plimpton, *Computer Physics Communications*, 2021, **271**, 108171. DOI: 10.1039/D5DD000579E
- J. Gilmer, S. Schoenholz, P. Riley, O. Vinyals and G. Dahl, *Proceedings of the 34th International Conference on Machine Learning*, 2017, **70**, 1263-1272.
- K. Schütt, P. Kindermans, H. Sauceda Felix, S. Chmiela, A. Tkatchenko, Müller, K.-R. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *Advances in neural information processing systems* 2017, 30.
- J. Klicpera, J. Groß, J. and S. Günnemann, *International Conference on Learning Representations*, 2020, 1-13.
- J. Gasteiger, F. Becker, and S. Günnemann, *Advances in Neural Information Processing Systems*, 2021, **34**, 6790-6802.
- O. Unke, S. Chmiela, M. Gastegger, K. T. Schütt, H. E. Sauceda and K.-R. Müller, *Nature Communications*, 2021, **12** (1), 7273.
- C. Chen and S. Ong, *Nature Computational Science*, 2022, **2** (11), 718-728.
- S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt and B. E. Kozinsky, *Nature Communications*, 2022, **13** (1), 2453.
- I. Batatia, D. Kovacs, G. Simm, C. Ortner and G. Csányi, *Advances in Neural Information processing systems*, 2022, **35**, 11423-11436.
- Y. Liao and T. Smidt, *International Conference on Learning Representations*, 2023.
- A. Musaelian, S. Batzner, A. Johansson, L. Sun, C. Owen, M. Kornbluth and B. Kozinsky, *Nature Communications*, 2023, **14** (1), 579.
- D. P. Kovacs, I. Batatia, E. Arany and G. Csyani, *The Journal of Chemical Physics*, 2023, **159**, 044118.
- D. Kovacs, J. Moore, N. Browning, I. Batatia, J. Horton, Y. Pu, V. Kapil, W. Witt, I.-B. Magdau, D. Cole and G. Csyani, *Journal of the American Chemical Society*, 2025, **147** (21), 17598-17611.
- N. Leimeroth, L. Erhard, K. Albe and J. Rohrer, *Modelling and Simulation in Materials Science and Engineering*, 2025, **33**, 065012.
- G. Wang, C. Wang, X. Zhang, Z. Li, J. Zhou, and Z. Sun, *iScience*, 2024, **27**, 109673.
- W. Stark, C. Oord, I. Batatia, Y. Zhang, B. Jiang, G. Csyani and R. Maurer, *Machine Learning: Science and Technology*, 2024, **5**, 030501.
- J. Smith, O. Isayev and A. Roitberg, *Chemical Science*, 2017, **8** (4), 3192-3203.
- M. Gastegger, J. Behler, and P. Marquetand, *Chemical Science*, 2017, **8** (10), 6924-6935.
- S. Manzhos and T. Carrington Jr, *Chemical Reviews*, 2020, **121** (16), 10187-10217.
- M. Gastegger, L. Schwiedrzik, M. Bittermann, F. Berzsenyi, and P. Marquetand, *The Journal of Chemical Physics* 2018, **148** (24).
- A. Bartók, S. De, C. Poelking, N. Bernstein, J. R. Kermode, G. Csányi and M. Ceriotti, *Science Advances*, 2017, **3** (12), e1701816.
- K. Lin, L. Paterson, F. May, and D. Andrienko, *npj Computational Materials*, 2021, **7** (1), 179.
- D. Earl, M. W. Deem, *Physical Chemistry Chemical Physics*, 2005, **7** (23), 3910-3916.
- A. Patriksson and D. van der Spoel, *Physical Chemistry Chemical Physics*, 2008, **10** (15), 2073-2077.
- O. Isayev, D. Fourches, E. Muratov, C. Oses, K. Rasch, A. Tropsha and S. Curtarolo, *Chemistry of Materials*, 2015, **27** (3), 735-743.
- M. Ceriotti, *The Journal of Chemical Physics*, 2019, **150**, 150901.



- 49 B. Cheng, R. Griffiths, S. Wengert, C. Kunkel, T. Stenczel, B. Zhu, V. Deringer, N. Bernstein, J. T. Margraf, K. Reuter, *Accounts of Chemical Research*, 2020, **53** (9), 1981-1991.
- 50 C. D. Williams, J. Kalayan, N. A. Burton and R. A. Bryce, *Chem. Sci.*, 2024, **15**, 12780–12795.
- 51 J. Behler, *Chemical Reviews*, 2021, **121** (16), 10037-10072.
- 52 J. Behler, *The Journal of Chemical Physics* 2011, **134** (7), 074106.
- 53 A. Singraber, T. Morawietz, J. Behler, and C. Dellago *Journal of Chemical Theory and Computation*, 2019, **15** (5), 3075-3092.
- 54 I. Poltavsky, A. Charkin-Gorbulin, M. Puleva, G. Fonseca, I. Batatia, N. Browning, S. Chmiela, M. Cui, J. T. Frank, and S. Heinen, *Chemical Science*, 2025, **16**, 3738-3754.
- 55 C. Schran, F. Thiemann, P. Rowe, E. A. Müller, O. Marsalek and A. Michaelides, *Proceedings of the National Academy of Sciences*, 2021, **118** (38), e2110077118.
- 56 X. Fu, Z. Wu, W. Wang, T. Xie, S. Keten, R. Gomez-Bombarelli, and T. Jaakkola, *arXiv preprint arXiv:2210.07237*, 2022.
- 57 V. Bihani, S. Mannan, U. Pratiush, T. Du, Z. Chen, S. Miret, M. Micoulaut, M. Smedskjaer, S. Ranu and N. A. Krishnan *Digital Discovery*, 2024, **3** (4), 759-768.
- 58 S. Chmiela, A. Tkatchenko, H. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, *Science Advances*, 2017, **3** (5), e1603015.
- 59 S. Chmiela, V. Vassilev-Galindo, O. Unke, A. Kabylda, H. E. Sauceda, A. Tkatchenko and K.-R. Müller, *Science Advances*, 2023, **9** (2), eadf0873.
- 60 S. Luo, T. Chen and A. Krishnapriyan, *ICLR*, 2024
- 61 Z. Guo, D. Lu, Y. Yan, S. Hu, R. Liu, G. Tan, N. Sun, W. Jiang, L. Liu, and Y. Chen, *Proceedings of the 27th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, 2022, 205–218.
- 62 J. Li, B. Li, Z. Guo, M. Li, E. Li, L. Liu, G. Yuan, Z. Wang, G. Tan, and W. Jia, *SC24: International Conference for High Performance Computing, Networking, Storage and Analysis*, 2024, 1–15.
- 63 Y. Zhao, M. Schwab, A. Kiersnowski, W. Pisula, M. Baumgarten, L. Chen, K. Müllen and C. Li, *C. Trifluoromethyl-functionalized bathocuproine for polymer solar cells. Journal of Materials Chemistry C*, 2016, **4** (21), 4640-4646.
- 64 M. Van Vleet, A. Misquitta, A. Stone and J. Schmidt, *Journal of Chemical Theory and Computation*, 2016, **12**, 3851-3870.

View Article Online
DOI: 10.1039/D5DD00579E



Data availability

The data and code relevant to this study are deposited in Zenodo [DOI:10.5281/zenodo.19078768]. The deposit includes the raw HDNNP reference datasets in the n2p2 input.data format; the explicit train/test split files corresponding to each trained model; the deposited input.nn files specifying the random_seed values used for the internal n2p2 split; the trained model artifacts (scaling.data and trained weight files) for the OLED reference and relabeled committees and for the aspirin and DWCNT reference/relabeled models; the FPS selection script; additional dataset-preparation scripts; and representative CP2K input files (aimd.inp and endpoint_recalculation.inp). Additional methodological details and supplementary figures are provided in the ESI. The external benchmark datasets used in this study are publicly available. The aspirin benchmark was taken from the revised MD17 (rMD17) dataset on Figshare (DOI : 10.6084/m9.figshare.12672038.v3), and the double-walled carbon nanotube benchmark was obtained from the MD22 dataset portal hosted by sGDML (accessed March 10, 2026).

