

<https://doi.org/10.1038/s41524-025-01825-w>

Accurate machine learning interatomic potentials for polyacene molecular crystals: application to single molecule host-guest systems

Check for updates

Burak Gurlek^{1,3}, Shubham Sharma^{1,3}, Paolo Lazzaroni¹, Angel Rubio^{1,2} & Mariana Rossi¹ ✉

Emerging machine learning interatomic potentials (MLIPs) offer a promising solution for large-scale accurate material simulations, but stringent tests related to the description of vibrational dynamics in molecular crystals remain scarce. Here, we develop a general MLIP by leveraging the graph neural network-based MACE architecture and active-learning strategies to accurately capture vibrational dynamics across a range of polyacene-based molecular crystals, namely naphthalene, anthracene, tetracene and pentacene. Through careful error propagation, we show that these potentials are accurate and enable the study of anharmonic vibrational features, vibrational lifetimes, and vibrational coupling. In particular, we investigate large-scale host-guest systems based on these molecular crystals, showing the capacity of molecular-dynamics-based techniques to explain and quantify vibrational coupling between host and guest nuclear motion. Our results establish a framework for understanding vibrational signatures in large-scale complex molecular systems and thus represent an important step for engineering vibrational interactions in molecular environments.

Organic molecular crystals, characterized by their long-range order and rich intermolecular interactions, are crucial in diverse applications, ranging from pharmaceuticals to electronics, and hold significant potential for emerging technologies, such as photovoltaics¹ and quantum information systems². While these applications primarily rely on the underlying electronic properties of these systems, molecular vibrations, encompassing both inter- and intramolecular modes, are equally important due to their role in determining the crystal structure and due to the pronounced electron-phonon coupling which is often observed^{3–7}.

Specifically, molecular vibrations and their anharmonic couplings play a pivotal role in determining the thermodynamic stability of crystal polymorphs^{8–12}, in enhancing or hindering charge transport by modulating carrier mobility through dynamic intermolecular coupling^{7,13,14}, in facilitating rapid singlet fission to improve solar cell efficiency^{1,15,16}, and even in offering a potential usage as quantum memory elements¹⁷. For example, polycyclic aromatic hydrocarbons embedded in large-bandgap host materials are being explored as single-photon sources, nonlinear quantum optical elements, and nanoscale sensors², as they exhibit narrow optical transitions at cryogenic temperatures, allowing highly coherent light-matter

interactions¹⁸. However, previous studies have predominantly focused on the electronic transitions, leaving the rich internal structures arising from vibrational degrees of freedom largely unexplored¹⁹. Despite their undeniable importance, accurately modeling vibrational dynamics that are affected by anharmonic mode-coupling and long-range van der Waals (vdW) interactions is hampered by the computational complexity of such simulations, which makes them prohibitively expensive with traditional first-principles methods such as density-functional theory (DFT).

Machine learning interatomic potentials (MLIPs) hold great promise in addressing the challenges associated with large-scale and long-time simulations of complex material systems, offering high computational efficiency without compromising accuracy^{20–27}. Recent developments in active-learning strategies and in strategies for sampling diverse atomic environments have further enabled the construction of smaller representative training datasets^{28–31} that deliver good training accuracy. These methods and training strategies have allowed the generalization of these potentials through the proposition of foundational models that can perform well for a large variety of systems, including those not represented in their training set^{32–38}.

¹Max Planck Institute for the Structure and Dynamics of Matter and Center for Free-Electron Laser Science, Hamburg, Germany. ²Initiative for Computational Catalysis (ICC), The Flatiron Institute, New York, NY, USA. ³These authors contributed equally: Burak Gurlek, Shubham Sharma.

✉ e-mail: mariana.rossi@mpsd.mpg.de

Despite the successes of MLIPs in modeling solids^{39,40}, solid-liquid interfaces⁴¹, and chemical reactions⁴², assessments of their performance and reliability in describing vibrational dynamics in molecular crystals remain limited. Existing studies have mostly focused on inorganic and covalently-bonded systems^{40,43–47,48}. Molecular crystals are particularly challenging due to their soft vibrational modes, which are impacted by intermolecular vdW interactions. The full long-range and non-local character of vdW interactions is normally not captured by MLIP architectures based on local atomic environments, which calls the accuracy of these architectures into question.

In this work, we develop MLIPs capable of accurately capturing harmonic and anharmonic vibrational dynamics in polyacene molecular crystals. Starting from naphthalene as a model system, we systematically develop MLIPs and assess their predictive accuracy across the polyacene series to pentacene, demonstrating that the potentials can generalize to larger acenes and to previously unseen host-guest configurations. We further assess the reliability of these MLIPs by showing how errors on forces propagate to phonon frequencies and anharmonic vibrational densities of states (VDOS). These measures allow for a rigorous predictive confidence of these quantities and for devising active learning targets for vibrational properties.

The extrapolative capabilities of these MLIPs to predict vibrational properties of host-guest systems, in particular pentacene molecules embedded in a naphthalene crystal host, provide important new insight for vibrational control. We present a clear assessment of host and guest vibrational mode assignment when anharmonic correlations play an important role, thus providing a foundation for the study of vibrational coherence and decoherence processes in molecular host-guest systems.

Results

Performance of VASP and MACE machine-learning potentials based on active learning

To investigate the accuracy of MLIPs for modeling vibrational dynamics in molecular crystals, we compare the performance of different MLIPs for the naphthalene crystal. We train a VASP machine-learning model and a MACE model (see “Methods”) using a dataset generated through the active learning strategy in VASP²², which employs on-the-fly sampling of structures from a molecular dynamics (MD) trajectory based on uncertainties in the predicted energy, forces and stresses, utilizing Bayesian regression.

We created the training dataset by running MD trajectories with VASP at 295 K for a $1 \times 2 \times 2$ naphthalene supercell. During the MD run, we monitored the changes in the number of training structures and the forces root mean square error (RMSE) of the model. We stopped the active learning process when these metrics showed negligible change, resulting in 1402 structures in the training dataset (see details in “Methods”). These structures were then used to train a MACE equivariant message-passing machine learning potential²⁴. The MACE MLIP training is continued until the RMSE of forces and energy on the validation set converged across different epochs (see details in “Methods”). To assess the stability of resulting MLIPs, we conducted a 1 ns NVT-MD run on a $4 \times 4 \times 4$ naphthalene supercell, which showed no signs of instability.

We compare the performance of the two MLIPs using the RMSE of energy and forces on the training dataset, as shown in Table 1. These metrics indicate that the MACE MLIP outperforms the VASP MLIP, particularly in predicting atomic forces, which is reflected on the accuracy with which it can

predict harmonic phonon frequencies, as shown in the Supplementary Note 1. We note that the MACE MLIP predicts energies more accurately for structures at temperatures close to its training temperature of 295 K (see relevant discussions in Supplementary Note 1). Overall, the improved performance of the MACE model on the transferred dataset could be due to its longer effective interaction range⁴⁹ or to the higher effective body order achieved with the message-passing procedure. The VASP model we trained contains up to 9-body order in the kernels and cutoffs of 8 Å for radial descriptors and 5 Å for angular descriptors, while the MACE model trained in this work goes up to 13 body-order and 12 Å effective cutoff. In addition, as shown in Supplementary Note 2, the performance of the VASP model on a transferred dataset obtained through a different active-learning strategy is relatively poor, indicating a limitation in data transferability similar to those observed in other MLIP architectures⁴⁹. We note that this transferred dataset, covering multiple temperatures, is the same dataset later used for the MACE potential, which is described in the next paragraph.

We next examine the impact of different active-learning strategies on the overall performance of MLIPs. To achieve this, we construct a new VASP MLIP by sequentially sampling MD trajectories at temperatures of 295 K, 220 K, 150 K, 120 K, and 80 K using VASP’s active learning algorithm, explained in ref. 22. The final training dataset consists of 1168 naphthalene structures, distributed as 940, 22, 145, 14, and 47 structures for each respective sampling temperature. We will refer to the resulting MLIP as VASP MLIP-multi for the remainder of the text. To construct the MACE MLIP, we employ a committee-based active learning strategy²⁹ explained in “Methods”. We simultaneously trained a committee of eight MACE MLIPs on an initial dataset of 100 structures, selected from a pool using farthest point sampling (FPS)^{28,50} (see details in “Methods, Computational Details”). At each active-learning iteration, 25 structures were added to the training dataset based on energy uncertainty, resulting in a total of 450 naphthalene crystal structures. The number of structures selected at each respective temperature were 290, 71, 30, 33, 26. We refer to the resulting model as MACE MLIP-committee in the following discussions.

In Table 2, we present a comparison of training and test errors for the VASP MLIP-multi and MACE MLIP-committee models, evaluated on an independent test dataset composed of 2100 naphthalene crystal structures (see details in “Methods, Computational Details”). The VASP MLIP-multi shows training errors comparable to those of the initial VASP MLIP (see Table 1) and may exhibit a sample-bias issue within the training dataset, as the test errors on forces are slightly smaller than the training errors. The MACE MLIP-committee outperforms the earlier MACE MLIP by achieving lower errors in predicting atomic forces. The close agreement between training and test errors for MACE MLIP-committee indicates that it is not overfitted, nor does it suffer from sample bias. Furthermore, in Fig. 1a, b, we compare the MLIP and DFT-predicted energies and force components of each structure in the test set, showing good predictive performance for both VASP MLIP-multi and MACE MLIP-committee.

In Fig. 1c, we compare the performance of the new MLIPs in predicting Γ -point phonon frequencies. The MACE MLIP-committee outperforms the VASP MLIP-multi, with mean percentage (absolute) frequency errors of 0.17% (0.98 cm^{-1}) and non-outlier maxima of 0.27%, corresponding to only

Table 1 | The training errors of VASP and MACE MLIPs for the naphthalene molecular crystal, based on a single-temperature training at 295 K

	Energy (meV/atom)	Force (meV/Å)
VASP	0.1	23.7
MACE	0.1	10.5

The errors are presented as RMSE for energies and forces.

Table 2 | The training and validation errors of VASP MLIP-multi and MACE MLIP-committee for the naphthalene molecular crystal, based on training data at 80 K, 120 K, 150 K, 220 K and 295 K, accounting for thermal lattice expansion on each dataset

	Energy (meV/atom)		Force (meV/Å)	
	Training	Validation	Training	Validation
VASP	0.1	0.2	20.9	14.6
MACE	0.1	0.1	4.3	4.4

The errors are presented as RMSE for energies and forces.

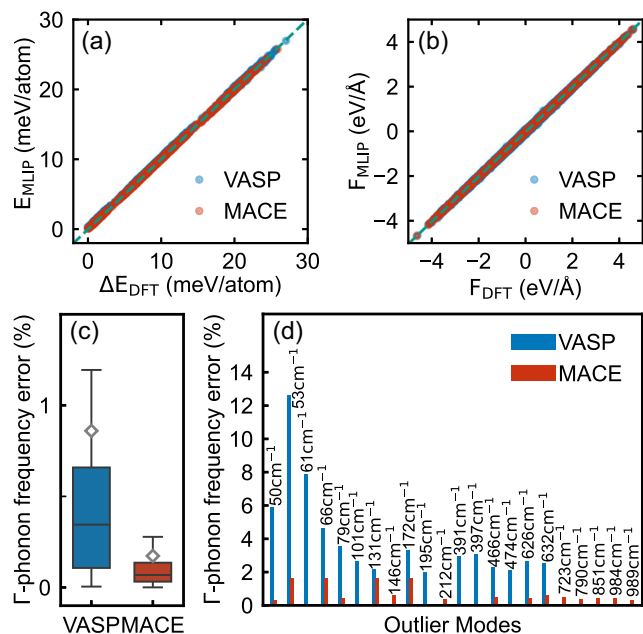


Fig. 1 | Comparison of VASP MLIP-multi and MACE MLIP-committee for the naphthalene molecular crystal. **a** Correlation plot of relative energies, $\Delta E_{\text{DFT(MLIP)}} = E_{\text{DFT(MLIP)}}^{\text{min}} - E_{\text{DFT}}^{\text{min}}$, and **b** forces predicted by DFT and MLIPs. **c** Error box plot of harmonic phonons (Γ -point) obtained with MLIPs. **d** Outliers identified from the box plot in (c). The wavenumbers of the modes correspond to those of the associated modes of the reference VASP DFT calculations.

2.88 cm^{-1} . Error distributions for outlier modes, shown in Fig. 1d, reveal that VASP MLIP-multi struggles with accurately predicting intermolecular vibrations. The MACE MLIP-committee achieves absolute frequency errors below 3.5 cm^{-1} with mean frequency errors of 0.48 cm^{-1} for intermolecular, 1.03 cm^{-1} for intramolecular and 1.39 cm^{-1} C-H stretching modes, surpassing the predictive capabilities of other MLIPs (see Supplementary Fig. 5b). Besides, the overall performance of MACE MLIP-committee is significantly improved compared to MACE MLIP, demonstrating the effectiveness of the committee-based active learning approach in capturing diverse atomic configurations.

These results demonstrate that both VASP and MACE MLIPs can achieve comparable accuracy in predicting the vibrational properties of naphthalene molecular crystals. However, when combined with a committee-based active-learning algorithm, the MACE model yields the best accuracy. Therefore, in the remainder of this work, we will focus on the MACE MLIP.

Committee uncertainty propagation for vibrational properties

Determining the reliability and confidence of any MLIP prediction relies on being able to calculate uncertainties for directly predicted quantities, as well as for quantities derived from such predictions. Various methodologies have been proposed for quantifying uncertainties in MLIP predictions of energies and forces^{22,29,51–56}, and for propagating these uncertainties to static observables^{55–57}. Quantifying uncertainties in dynamical (time-dependent) observables is also critical in the context where anharmonic vibrational couplings, vibrational lifetimes, and transport coefficients are derived from molecular dynamics simulations⁵⁸. In these simulations, the time-evolved atomic motion is governed by forces derived from the MLIP's potential energy surface. Errors in force predictions, especially for underrepresented or rare atomic configurations in the training dataset, will propagate to these observables. Therefore, we use our committee model to propagate uncertainties in MLIP predictions to the harmonic phonon frequencies and anharmonic mass-weighted VDOS, as discussed below. In this paper, when we refer to VDOS, we will always be referring to the mass-weighted quantity as written in Eq. (7).

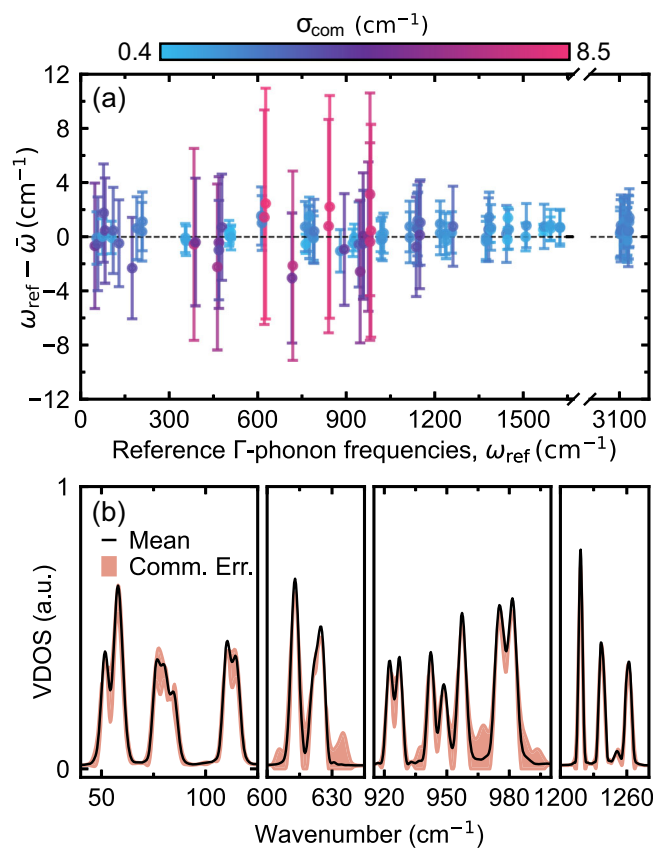


Fig. 2 | Uncertainty propagation of phonon frequencies and VDOS in a naphthalene crystal at 80 K. **a** The uncertainty estimations for the Γ -point phonon frequencies. The mean committee predictions ($\bar{\omega}$) are compared to reference DFT calculations (ω_{ref}), with error bars color-coded to represent the standard deviation among committee members. **b** Uncertainty estimations for the VDOS. The red shaded area and the black curve represent the committee error calculated by Eq. (10) and the mass-weighted VDOS calculated by using the committee mean force, respectively.

Employing the MACE MLIP-committee model, we first calculate the committee uncertainty for Γ -point harmonic phonon frequencies as detailed in “Methods, Uncertainty estimation and propagation to harmonic phonons”. In Fig. 2a, we show the distribution of relative errors between committee predictions and reference DFT calculations for the phonon spectrum. Overall, we conclude that the propagated uncertainty shows a good prediction capability of the real error across the whole frequency range. The largest uncertainties appear in the region between 600 and 1000 cm^{-1} . This region is dominated by modes that involve the in-plane and out-of-plane deformations of the fused benzene rings, as well as the in-plane and out-of-plane bending of CH groups (see Supplementary Fig. 6). As shown in Supplementary Fig. 7, sampling the displacements along these high-uncertainty modes in a committee-based active-learning procedure substantially improves both the predictive accuracy and uncertainty estimations of the corresponding normal modes, without the need for brute-force molecular dynamics. We do not observe instances of the uncertainty underestimating the real error for this model.

Next, we tackle a much harder problem, related to the propagation of the uncertainty in committee-MLIP predictions to the anharmonic VDOS, detailed in Methods. To illustrate the procedure for the naphthalene crystal, we perform molecular dynamics equilibration runs at 80 K for each committee member, propagating the dynamics using rescaled forces computed via Eq. (3). Afterwards, we perform 100 NVE simulation runs of 20 ps each, using the same forces for the propagation and calculate the VDOS_{F_i} corresponding to a given committee member (Eq. (7)). We call $\text{VDOS}_{\bar{F}}$ the VDOS computed by propagating trajectories using the committee's mean

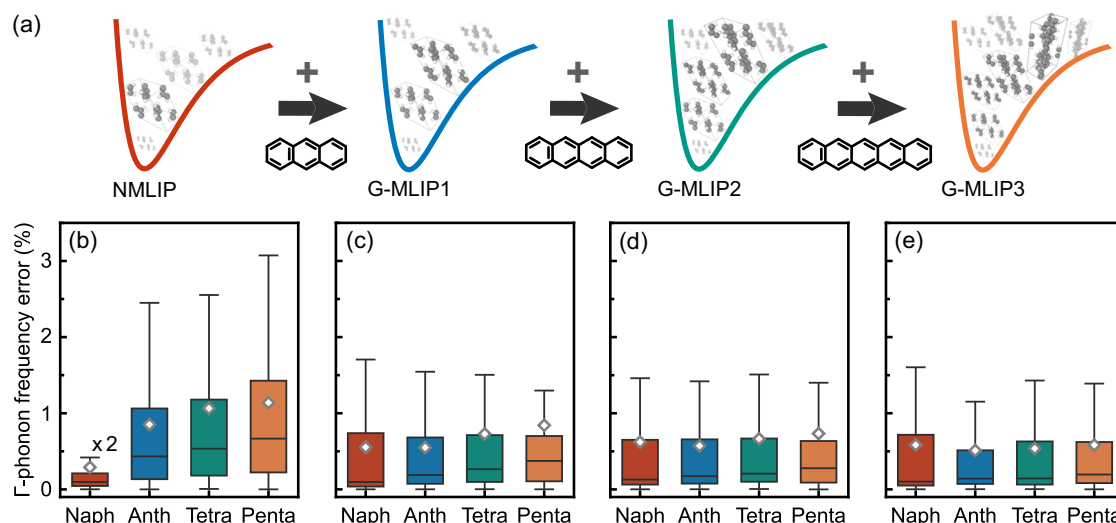


Fig. 3 | Generalized MLIPs for polyacene crystals: development and validation. a Sketch illustrating the active-learning scheme used to create generalized MLIPs. b–e Error box plots for the Γ -point phonon frequencies of naphthalene (Naph),

anthracene (Anth), tetracene (Tetra) and pentacene (Penta) molecular crystals, predicted by N-MLIP (b), G-MLIP1 (c), G-MLIP2 (d) and G-MLIP3 (e).

forces \bar{F} , which is the standard quantity computed when using committee models. Statistical uncertainties σ_{stat} for $\text{VDOS}_{\bar{F}}$ and each VDOS_{F_i} are obtained from the block average of 100 NVE runs, and σ_{com} is determined by averaging over all committee members (see “Methods, Uncertainty propagation to vibrational density of states”).

An issue with this procedure is that due to the non-linear dependence of the VDOS on forces, the VDOS obtained from averaging the predictions of all committees, $\overline{\text{VDOS}_{F_i}}$, is not equal to $\text{VDOS}_{\bar{F}}$. However, as we show in Supplementary Fig. 8, both spectra are quite similar. We can formally only calculate the uncertainty on VDOS_{F_i} with the procedure we follow in this work, and we therefore take this uncertainty as a proxy for the uncertainty in $\text{VDOS}_{\bar{F}}$. The standard error on the mean of the spectra is reported.

In Fig. 2b, we show these results in distinct frequency regions corresponding to intermolecular and intramolecular vibrations. We confirm the robustness of the statistical sampling across the entire frequency range and find that the statistical error is overall small (see Supplementary Fig. 9). Furthermore, the committee error is generally comparable in magnitude to the statistical error, indicating that the variability among committee members is on par with statistical fluctuations. However, around 600 cm^{-1} and within the range $900\text{--}1000\text{ cm}^{-1}$, the committee error shows that different committees would predict peak positions differently, leading to larger errors along the frequency axis. Interestingly, these regions correlate with the regions of largest uncertainties on harmonic phonon frequencies shown in Fig. 2a.

Such a careful uncertainty quantification shows that the MACE-MLIP committee model can deliver accurate harmonic and anharmonic vibrational properties of the naphthalene molecular crystal. It also defines the limits within which peak positions and widths can be interpreted, based on committee uncertainty. This approach separates model uncertainty from statistical noise, which mainly affects spectral intensities, allowing estimation of MLIP-related errors. We find that the uncertainties on harmonic modes correlate with the errors in the anharmonic dynamical vibrational spectra, making them useful for assessing VDOS-prediction reliability, at least at lower temperatures.

Generalizing machine-learning potentials for polyacene molecular crystals

Next, we investigate the capability of the MACE MLIP to generalize across acene-based molecular crystals for the prediction of vibrational dynamics. We choose to train our own general potentials, instead of using a

foundational model such as MACE-OFF³⁴, because in this work an important goal is to carefully benchmark the quality of potentials for vibrational properties. Therefore, training a model from scratch with data coming from codes we can fully control, including a uniform definition of the DFT functional, basis sets and other numerical settings, is paramount. In addition, the committee-based uncertainty quantification requires the ability to generate and subsample datasets, which we can create specifically for this purpose.

We employ a systematic active-learning strategy based on the MACE MLIP-committee developed in the previous section. The MLIP is progressively generalized by sequentially incorporating molecular crystal structures with an increasing number of fused benzene rings, i.e., anthracene, tetracene, and pentacene molecular crystals, as schematically shown in Fig. 3a. The similarity of the crystals across the acene-based series makes it more likely that the generalization of the MLIP will be successful. The total pool of training structures for the each molecular crystal are reported in Table 9. After each active-learning step, we assess the performance of the resulting MLIP in predicting harmonic Γ -point phonon frequencies for naphthalene, anthracene, tetracene, and pentacene molecular crystals.

We begin by evaluating the generalization capabilities of the MACE MLIP-committee developed in section “Performance of VASP and MACE machine-learning potentials based on active learning” without including any additional data. Hereafter, we label this MLIP as N-MLIP. To evaluate the stability of N-MLIP, we performed 2 ns long NVT MD runs at 295 K on $1 \times 2 \times 2$ supercells of anthracene, tetracene and pentacene molecular crystals, observing no signs of instabilities. We then analyzed the performance of N-MLIP in predicting phonon frequencies as shown in Fig. 3b. Compared to its performance on the naphthalene molecular crystal, the accuracy of N-MLIP is nearly five to ten times lower for anthracene, tetracene and pentacene molecular crystals (see also Supplementary Fig. 10 for outlier modes and Supplementary Table 2 for absolute errors). This limitation is particularly pronounced in the maximum absolute frequency errors reaching up to 40 cm^{-1} in the case of the tetracene molecular crystal. Despite the stability of the N-MLIP, its predictive accuracy for vibrational dynamics within the acene family is very limited.

To address this limitation and improve the generalization of the potential, anthracene molecular crystal structures are added to the training dataset using the active-learning strategy. Following five consecutive active-learning steps, 125 anthracene structures were incorporated into the dataset, resulting in an updated MLIP, denoted as G-MLIP1. Stability tests confirm

that G-MLIP1 is robust across all studied molecular crystals (see details in “Methods, Committee-based active learning strategy”). As illustrated in Fig. 3c, G-MLIP1 achieves nearly twofold improvements in phonon frequency predictions for anthracene, tetracene, and pentacene crystals (see also Supplementary Table 2). However, this generalization comes at the cost of reduced accuracy for naphthalene, attributed to a trade-off between specificity and broader applicability across the acene family. Nevertheless, the inclusion of anthracene configurations significantly enhances the extrapolative capacity of the MLIP.

We then examine the impact of expanding the training dataset with additional tetracene molecular structures. G-MLIP2, built by incorporating 150 primitive-cell structures of tetracene through six successive active-learning steps, demonstrates stability across all tested systems. While it shows a slight improvement in overall performance, as illustrated in Fig. 3d, it notably reduces maximum phonon frequency errors by nearly 10 cm^{-1} for tetracene and pentacene molecular crystals, highlighting its enhanced generalization to more complex molecular environments (see Supplementary Table 2).

Building on this observation, we further expanded the training dataset by incorporating 125 pentacene structures obtained through five consecutive active-learning steps. The resulting G-MLIP3 remains stable across all acene molecular crystals and demonstrates slightly improved performance over earlier models (see Fig. 3e). Importantly, it achieves a consistent average error of $\sim 2.8\text{ cm}^{-1}$ across all systems, highlighting its ability to generalize effectively to diverse molecular environments while avoiding overfitting to any specific system (see Supplementary Table 2). This robust generalization is further evident in its threefold reduction of the maximum phonon frequency error for the pentacene molecular crystal.

A closer examination of the errors associated with G-MLIPs reveals distinct performance trends for intermolecular and intramolecular vibrational modes throughout the molecular crystals investigated. G-MLIP3 stands out with the lowest mean errors, consistently below 3.0 cm^{-1} (Supplementary Table 2), effectively capturing the key physical characteristics of both high-frequency intramolecular modes and low-frequency intermolecular modes. This good performance is particularly prominent for intramolecular vibrations in the range of $1000\text{--}1600\text{ cm}^{-1}$, while for the lower-frequency range of $100\text{--}250\text{ cm}^{-1}$, all the G-MLIPs exhibit similar error performance (see Supplementary Fig. 11). As a generalization test for G-MLIP3, we further assessed its performance on the vibrations of crystal polymorphs of tetracene and pentacene, as shown in Supplementary Fig. 12. The potential is as accurate for different polymorphs as it is for the polymorph it was trained on. In addition, we analyze whether the committee uncertainties remain predictive of the potential’s accuracy by propagating the uncertainties to harmonic phonon frequencies of naphthalene, anthracene, tetracene, and pentacene molecular crystals, following section “Committee uncertainty propagation for vibrational properties”. As shown in Supplementary Fig. 13, G-MLIP3 generally exhibits smaller uncertainties for the naphthalene crystal compared to MACE MLIP-multi (see Fig. 2a), despite its lower predictive performance. This indicates that the committee model is slightly overconfident and underestimates the actual error, particularly in the $150\text{--}1000\text{ cm}^{-1}$ range. However, all errors we quantify are already very small.

These results demonstrate that the generalized MLIPs derived from the multi-acene active-learning strategy exhibit good performance and robust MD stability. The inclusion of larger molecular structures into the training dataset results in a clear accuracy improvement for vibrational properties, highlighting the benefits of diversifying the dataset with closely related molecular structures. Next, we use G-MLIP3 to study the vibrational dynamics in acene-based host-guest systems, which include atomic environments not represented during training, testing the model’s ability to extrapolate to new, unseen structures.

Vibrational correlations in a host-guest system

An attractive property of single-molecule host-guest systems is that, once engineered, the vibrational levels in these systems could feature long

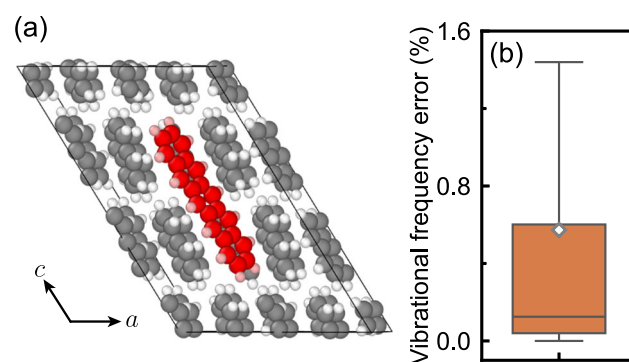


Fig. 4 | Extrapolative performance of G-MLIP3 in host-guest systems. **a** Crystal structure of pentacene-doped $2 \times 2 \times 3$ naphthalene molecular crystal, with a and c representing the crystal axes in monoclinic symmetry. Two naphthalene molecules along the front-facing plane are removed for illustration purposes. **b** Box plot showing the prediction errors for vibrational frequencies of the host-guest system.

coherence times that exceed those of the electronic transitions, potentially paving the way for the realization of quantum memories and efficient optomechanical interactions¹⁷. Therefore, to fully harness the potential of molecular host-guest systems, it is important to develop a deeper understanding of their vibrational dynamics—a challenge that traditional *ab-initio* methods struggle to address. Here, we validate the G-MLIP3 potential developed in the previous section for host-guest systems and apply it to explore their vibrational properties, at this point keeping a classical description of anharmonic nuclear motion.

We investigate the pentacene-doped naphthalene molecular crystal due to its compatibility with the generalized MLIP and its relevance in both experimental and theoretical studies^{59,60}. A schematic visualization of this system is shown in Fig. 4a, where a pentacene molecule replaces two naphthalene molecules in a $2 \times 2 \times 3$ naphthalene supercell. Using G-MLIP3 alongside reference DFT calculations, we first relax the host-guest molecular crystal and evaluate the accuracy of G-MLIP3. The insertion energy error for the guest molecule in the host crystal is found to be 0.1 meV per atom, consistent with the test set error (see Table 2). Furthermore, we compute the vibrational frequencies of the host-guest system and benchmark these results against reference DFT calculations. As illustrated in Fig. 4b, the mean vibrational frequency error is less than 1%, and its overall performance is on par with G-MLIP3 (see Fig. 3), demonstrating the ability of MLIPs to generalize to unseen atomic configurations, with absolute errors consistently below 15 cm^{-1} (see Supplementary Fig. 14).

As shown in Supplementary Note 3, we find that G-MLIP3 rapidly loses accuracy under isotropic cell expansion due to its inability to capture long-range vdW interactions, whereas NMLIP retains accuracy for small expansions when predicting properties of the naphthalene crystal. When we analytically include vdW interactions in the potential, long-range effects can be accurately described even for very large lattice expansions in both generalized and specialized MACE MLIPs (see Supplementary Note 3). In the following simulations, we fix the lattice constants to the experimental values of the 295 K naphthalene structure.

In order to gauge the reliability of the vibrational property predictions of G-MLIP3 on the host-guest system, we analyse the propagated uncertainties on the harmonic phonons. In Supplementary Fig. 15, we show that even for the significantly larger $4 \times 4 \times 5$ pentacene-naphthalene supercell containing 8637 vibrational modes, the committee uncertainty remains below 11 cm^{-1} across the entire vibrational spectrum.

After confirming the reliability of the generalized MLIP for the host-guest system, we apply it to the analysis of its vibrational dynamics. The vibrational landscape of combined host-guest systems can be rationalized by

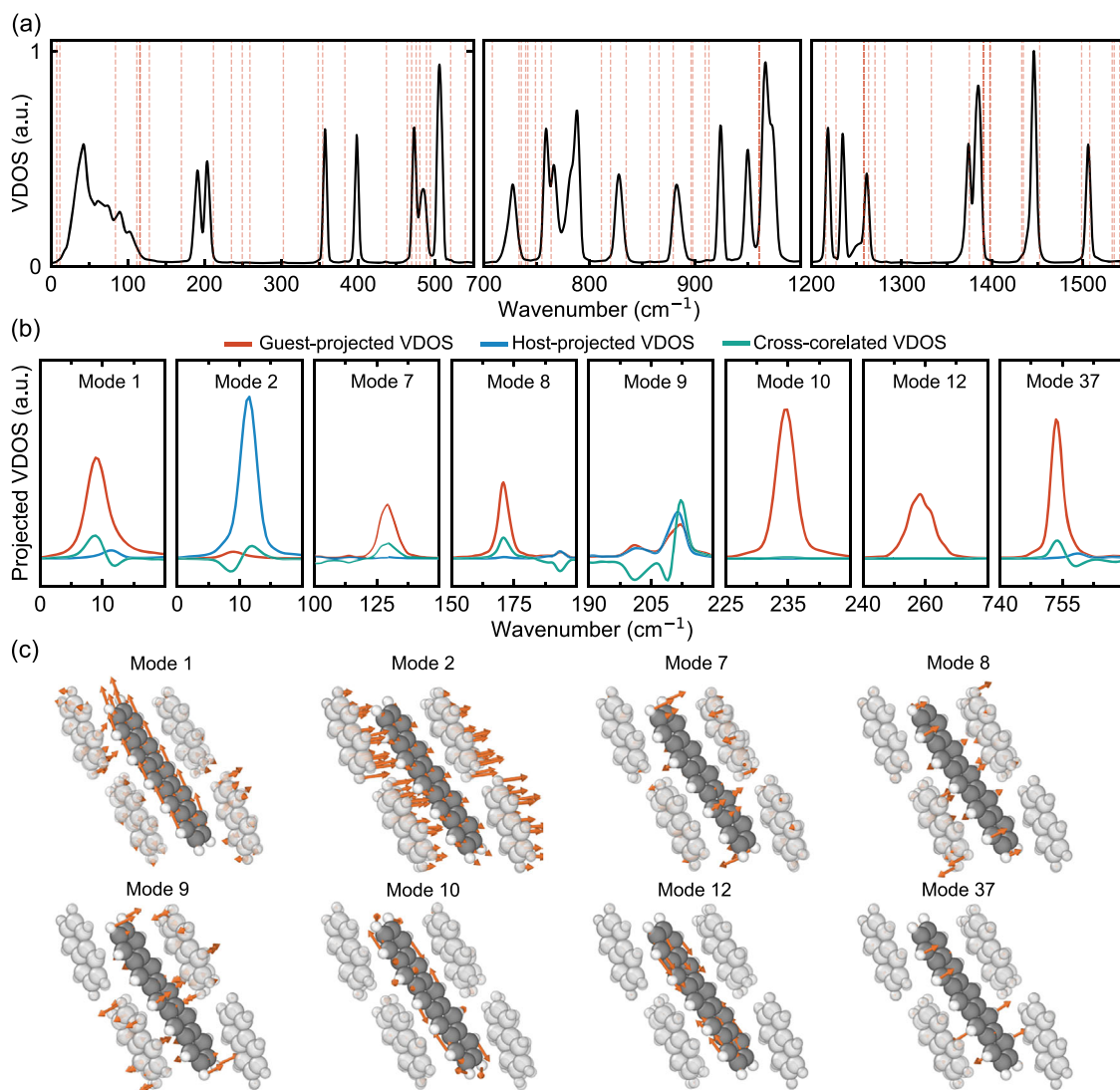


Fig. 5 | Vibrational dynamics in host-guest systems captured by G-MLIP3. **a** VDOS for the $4 \times 4 \times 5$ host-guest system. The red dashed lines indicate the frequencies of vibrational modes where the atomic displacements of the guest molecule dominate over those of the host molecules. **b** Normal-mode-projected VDOS for the host-guest system, with the red, blue and green lines representing guest-projected, host-projected and cross-correlated VDOSs, respectively.

c Illustrations of the normal modes of the host-guest system discussed in **(b)** in Cartesian space, along the ac crystal plane, showing only the most relevant naphthalene molecules from the $4 \times 4 \times 5$ supercell. For clarity, the displacement vectors of the host and guest molecules are scaled by a factor of 3:1, except for modes 2 and 9, which are scaled by 1:1. For more detailed illustrations, see Supplementary Fig. 19.

considering the vibrations of the isolated host and isolated guest systems. The host crystal exhibits a continuum of intermolecular vibrational modes (often named phonons) up to 150 cm^{-1} , and a series of intramolecular vibrational bands, while the guest molecule has discrete intramolecular vibrational modes (often named vibrons) starting at frequencies around 30 cm^{-1} , depending on its size¹⁹. When these two systems are combined, their vibrational modes hybridize, resulting in guest-like, host-like, and mixed modes. The guest-like modes in the phonon spectral region are referred to as *pseudolocal modes*, characterized by a decaying vibrational amplitude away from the guest molecule¹⁹. These pseudolocal modes can play a crucial role in explaining the temperature-dependent dephasing of electronic transitions⁶¹.

In high-resolution vibronic spectroscopy of single-molecule host-guest systems, optical scattering is observed at frequencies corresponding to transitions to the continuum of host phonons, guest-like vibrations, and pseudolocal modes^{59,62,63}. Transitions to high-frequency host-like modes above the phonon cut-off frequency of 150 cm^{-1} , are negligible due to weak electron-vibration coupling strengths⁶⁴. Despite previous studies on specific

vibrational modes in these systems^{63,65–67}, a comprehensive understanding and characterization of these modes and their correlations remains largely unexplored.

We calculate the VDOS of a $4 \times 4 \times 5$ pentacene-doped naphthalene molecular crystal (2880 atoms) at 100 K (see details in “Methods, Computational Details”). The VDOS in Fig. 5a is analyzed in three spectral regions where prominent vibronic transitions have been observed⁶³. To rationalize these features, we also calculate harmonic phonons and identify 98 harmonic normal modes with significant atomic displacements in the guest molecules. These are highlighted as potential guest-like modes in Fig. 5a with red dashed lines. It is worth noting that a visual comparison of harmonic and anharmonic spectra of this host-guest system at 100 K does not immediately show any clear signs of anharmonicity below 800 cm^{-1} , see Supplementary Fig. 17.

The continuum of phonon modes is clearly observed below 150 cm^{-1} , where the spectrum peaks around 45 cm^{-1} , corresponding to the low group velocities of naphthalene optical phonon modes. This peak gradually decays toward the phonon cut-off frequency, in agreement with experimental

observations^{59,63}. In this region, we identify seven modes with the potential to represent pseudolocal modes. The remainder of the spectrum above 150 cm⁻¹ exhibits vibrational bands originating from intermolecular and intramolecular vibrational modes of naphthalene molecules (isolated bands) and mixed vibrational modes (bands coinciding with guest-like vibrational modes). We also observe minor peaks between these bands, coinciding with the red dashed lines, which may correspond to guest-like vibrations, e.g., around 230 cm⁻¹ and 430 cm⁻¹. Note that, due to the large number of degrees of freedom in the host molecules, the contributions from the guest molecules are not easily discernible in the total VDOS, but can be more clearly seen in the projected VDOS of the guest molecule (see Supplementary Fig. 18).

To gain a deeper understanding of the vibrational mode structure, we project the VDOS onto the 98 potential guest-like normal-modes of the host-guest system, and separate the contributions from the host and the guest degrees of freedom as outlined in Methods. This projection yields: (i) a partial projected VDOS arising solely from the guest molecular motions; (ii) a partial projected VDOS accounting for vibrational contributions solely from the host molecules; and (iii) the cross-correlated VDOS, which contains the cross contributions of host and guest vibrations in the normal mode basis. The cross-correlated VDOS term reflects potential hybridization effects in the vibrational spectrum.

In Fig. 5b, we present an example of this projection for several guest-like modes across various spectral regions, including pseudolocal vibrational modes (below 150 cm⁻¹) and the most relevant guest-like modes, illustrating their mode displacements in Fig. 5c. Among the seven candidate pseudolocal modes (red-dashed lines below 150 cm⁻¹ in Fig. 5a), only modes 1 and 7 exhibit strong localization on the guest molecule with small contributions from the host atoms as evidenced by the cross-correlated VDOS. These modes correspond to hindered backbone bending and wave-like out-of-plane deformations, with corresponding frequencies of 36 cm⁻¹ and 100 cm⁻¹, respectively, for pentacene in vacuum. Mode 1 is coupled to the nearby host-like mode 2 in an anti-correlated manner, likely arising from the strong hybridization of the same backbone bending mode of pentacene with host phonons (see Fig. 5c). This anti-correlation behavior suggests anharmonic coupling between these two modes. Interestingly, the mode at 130 cm⁻¹ lies at the phonon band edge and exhibits a non-Lorentzian lineshape, consistent with experimental observations⁶³ that report sharp spectral features in this region.

Moreover, modes 10 and 12, which involve in-plane torsional and stretching motions, do not overlap with any host spectral features and remain largely localized on the guest molecule. In contrast, mode 8, which exhibits out-of-plane torsional motion, shows a mixed character and couples to host intermolecular vibrations near 200 cm⁻¹, despite also lacking any direct spectral overlap with host vibrations. This highlights that mode coupling is influenced not only by spectral overlap but also by nonlinear effects in the potential energy landscape. Such anharmonic couplings could explain discrepancies between earlier spectroscopic mode assignments and *ab initio* results⁶³.

Finally, modes 9 and 37 exemplify the effect of spectral overlap with host bands around 200 cm⁻¹ and 760 cm⁻¹, respectively. Mode 9, involving wave-like out-of-plane deformations, exhibits mixed character with strong anharmonic coupling to nearby host modes. On the other hand, mode 37, featuring rocking deformations of the guest molecule, retains its guest character and shows only weak anharmonic interactions despite the energy overlap with the host modes.

Using the VDOS normal-mode projections, we classified 80 vibrational modes as guest-dominated (with an integrated guest-projected VDOS greater than 90% of total VDOS) and 11 mixed modes (including 2 pseudolocal modes) that exhibit strong guest character, based on the relative contributions from both guest and host components. Additionally, by fitting the normal-mode projected VDOS with a Lorentzian lineshape, we obtained the lifetimes of the guest-dominated vibrational modes, which averaged 3.5 ps (see Supplementary Fig. 20). This analysis is only expected to be accurate for low-frequency modes, below ≈ 300 cm⁻¹, for which

employing classical dynamics for the nuclei at 100 K is expected to be valid. Nevertheless, these lifetimes are shorter for modes with frequencies below twice the phonon cut-off frequency of the naphthalene molecular crystal, and longer above the cutoff, as previously discussed by Dlott et al.⁶⁸.

Discussion

Our work demonstrates the potential of MLIPs in accurately modeling the vibrational dynamics of polyacene molecular crystals, with a particular focus on their ability to generalize across related chemical spaces and predict vibrational properties and vibrational correlations of host-guest systems. Recent developments in MLIPs have led to the so-called foundational models, also based on the MACE architecture, which are transferable across large variety of systems. In particular, it is important to put the results presented in this paper within the context of MACE-OFF³⁴, which is a foundational model targeted at the description of organic molecules and molecular condensed-phase systems.

The potential we developed achieves errors comparable to the large models of MACE-OFF for energies and forces while keeping the maximum angular momentum of equivariant features equal to the small model (i.e., only invariant features of $L = 0$). By leveraging active-learning strategies, an excellent accuracy on energies, forces and vibrational properties was achieved for naphthalene, anthracene, tetracene and pentacene crystals with only a few hundred DFT calculations in total.

With this strategy, we could propagate and quantify errors in dynamical quantities such as anharmonic vibrational spectra. Our incorporation of uncertainty propagation into vibrational properties enables a quantitative assessment of the MLIP predictions of these quantities. We show that the error quantification for the harmonic spectra, when done carefully, serves as a good proxy for identifying the vibrational motions for which the potential is least accurate also in anharmonic spectra derived from molecular-dynamics. Even though we expect this relationship to degrade with increasing temperature, we propose that harmonic phonon uncertainties are incorporated on active learning strategies in order to improve the potential for anharmonic vibrational properties, as it is computationally quite challenging to calculate the uncertainties in the latter case. Indeed, we expect to use this strategy on top of general foundational models to fine-tune them for more reliable anharmonic vibrational analysis.

We also note that while the general MLIPs we developed for the polyacene crystals do not contain explicit long-range interactions, they show good accuracy also for low-frequency phonon modes where intermolecular forces dominated by vdW interactions play a dominant role. However, as demonstrated in Supplementary Note 3, explicit inclusion of vdW corrections is essential to accurately capture lattice expansion and contraction at larger scales.

The training strategy we have followed produced potentials that are also able to very accurately extrapolate to host-guest systems not contained in the training set, in particular regarding their anharmonic vibrational properties. We have presented an analysis of a pentacene molecule embedded into a naphthalene crystal matrix, unraveling guest-host vibrational mode correlations in exquisite detail and providing a measure of anharmonic coupling that goes beyond a simple analysis of mode-energy overlap or lineshape. These findings demonstrate that it is now possible to investigate the previously unexplored vibrational dynamics of host-guest systems that cannot be captured in small unit cells, leading to a rationalization of mode coupling and energy transfer pathways between host and guest vibrational modes. The ability to accurately predict vibrational properties in these systems opens new possibilities for engineering materials with tailored characteristics, such as optimized vibrational lifetimes or energy transfer pathways.

While our results show promising extrapolation within PAH-based host-guest systems, extending this approach to chemically or structurally diverse host-guest systems will benefit from general ML potentials that can address a wide range of systems with reasonable accuracy and uncertainty-aware predictions that allow targeting refinement for accurate vibrational properties.

Methods

Computational details

The reference calculations for geometry relaxations, atomic forces, energies, and stresses of the molecular crystal structures are performed at the DFT level. All DFT calculations are conducted using the FHI-aims⁶⁹ and VASP^{70–73} codes, employing the Perdew-Burke-Ernzerhof exchange-correlation functional⁷⁴ with vdW corrections^{75,76} for FHI-AIMS and VASP. We confirm that both VASP and FHI-aims calculations produce vibrational frequency calculations that differ by no more than 1.9 cm⁻¹, 1.3 cm⁻¹, and 4.6 cm⁻¹ for intermolecular, intramolecular, and C-H stretching modes, respectively. This ensures consistency and reliability for comparative analysis.

The on-the-fly machine learning algorithm in VASP dynamically constructs an accurate ML potential during AIMD simulations. It trains a kernel-based Gaussian process regression model using DFT-calculated energies, forces, and stresses for a selected subset of configurations. The algorithm yields predictive uncertainty through Bayesian error estimates for each configuration encountered during the simulation²². If the uncertainty exceeds a user-defined threshold, new DFT calculations are performed, and the resulting data are added to the training set to further refine the potential. The resulting DFT energies, forces and stresses are printed in the ML_AB file, which contains the reference values for all selected configurations. These values serve as the ground-truth DFT data for training the potential, while the MLIP itself guides efficient sampling of relevant configurations. Structuring DFT reference data into the ML_AB file allows VASP MLIP to be trained on external datasets without performing additional DFT calculations, as demonstrated in Methods.

In the MACE machine learning architecture, atomic structures are represented as graphs, where nodes correspond to atoms in three-dimensional space, and any two nodes within a cutoff distance are connected. Multiple high body-order message-passing iterations are performed to update the features associated with each node in the network. The final atomic site energy prediction is obtained as a function of all node states generated throughout the iterations. In this work we used VASP version 6.4.2 and MACE version 0.3.10.

A $2 \times 2 \times 1$ Monkhorst-Pack k-mesh is used to perform single-point calculations.

For VASP calculations, an energy cutoff of 1000 eV is applied to the plane-wave basis set, while a tight basis set is used for FHI-aims calculations. Geometry relaxations are performed within a fixed primitive cell, with a force convergence criterion of 4×10^{-4} eV/Å. Phonon calculations are performed using the phonopy code with a $2 \times 2 \times 2$ supercell⁷⁷.

The key hyperparameters used to train the VASP MLIP, in addition to the default values, are ML_MRB2 = 12, ML_SION1 = 0.3 and ML_WTSIF = 2. Before the evaluations, the MLIP is refitted with sparsification with ML_MODE = refit. The active-learning step in section “Performance of VASP and MACE machine-learning potentials based on active learning” resulted in 1402 structures, and the learning iteration was terminated when negligible changes were observed in the number of structures, as well as in the mean forces and energies, as shown in Supplementary Fig. 21.

The hyperparameters used for training the MACE MLIPs are summarized in Table 3.

The primitive lattice vectors and space groups of the molecular crystals used for training the generalized MACE MLIPs are provided in Tables 4–8. The dataset pools used for training generalized MACE MLIPs for acenes were generated as follows: For naphthalene, the dataset pool was created by performing 50 ps ab-initio MD simulations on a $1 \times 2 \times 2$ supercell using FHI-aims at temperatures of 80 K, 120 K, 150 K, 220 K, and 295 K⁷⁸. For anthracene, the dataset pool was generated for a $1 \times 2 \times 2$ supercell using the universal forcefield MACE-OFF³⁴ at temperatures of 100 K and 295 K. For tetracene and pentacene, the dataset pools were generated for $1 \times 1 \times 1$ cells using MACE-OFF at temperature 295 K. The total number of structures in the dataset pool for each molecular crystal is listed in Table 9. Single-point calculations for

Table 3 | Hyperparameters for MACE MLIPs

Hyperparameter	Value
Number of committee members	8
Number of interactions	2
Number of channels	256
Maximum L	0
Correlation	3
r_{\max} (Å)	6.0
Forces weight	1000
Energy weight	10
Train:Validation split	9:1

Table 4 | Number of acene molecular crystals in generalized potentials

MLIP	Naph.	Anth.	Tetra.	Penta.
N-MLIP	450	-	-	-
G-MLIP 1	450	125	-	-
G-MLIP 2	450	125	125	-
G-MLIP 3	450	125	125	125

Table 5 | Lattice parameters of the naphthalene crystal

T (K)/CCDC entry	a	b	c	α	β	γ	Space group
80 ⁷⁸ NAPHTA33	8.10	5.94	8.64	90	124.46	90	P 2 ₁ /a (14)
120 ⁷⁸ NAPHTA27	8.12	5.94	8.65	90	124.34	90	P 2 ₁ /a (14)
150 ⁷⁸ NAPHTA34	8.14	5.95	8.65	90	124.08	90	P 2 ₁ /a (14)
220 ⁷⁸ NAPHTA35	8.19	5.96	8.66	90	123.57	90	P 2 ₁ /a (14)
295 ⁷⁸ NAPHTA36	8.25	5.98	8.67	90	122.72	90	P 2 ₁ /a (14)

the selected structures in active learning iterations were performed using FHI-aims.

For both MACE and VASP MLIPs, vdW interactions are implicitly accounted for in the results presented in the main text, whereas in Supplementary Note 3, we include them explicitly, similar to other models^{79,80}.

Committee-based active learning strategy

To enhance training data efficiency and improve MLIP accuracy, we implemented a committee-based active learning strategy^{29,81–84}. A committee of 8 MACE MLIPs was trained simultaneously using the same dataset. The only differences among them were the initialization of weight parameters and the random split between the training and validation sets. All other hyperparameters were kept the same throughout the active learning process and are detailed in Table 3. This approach enables each MLIP to capture different landscapes of the potential energy surface, enhancing the diversity in predictions.

A pool of atomic structures for acenes was generated at different temperatures, a detailed description is provided in Computational Details. The workflow for the committee-based active learning begins with selecting a small subset of labeled data from the pool to train the ensemble of MLIPs. Each MLIP is trained until the training and validation errors stabilize at sufficiently low values for both energy and forces. Following this training phase, long MD simulations are performed using the mean of the committee forces to assess the stability of the MLIPs across various temperature ranges and supercell sizes. For all MACE MLIPs, we use the accuracy of the Γ -point phonon frequencies, computed using the mean of the committee forces, as

Table 6 | Lattice parameters of the anthracene crystal

T (K)/CCDC entry	a	b	c	α	β	γ	Space group
100 ⁸⁶ ANTGEN25	9.28	5.99	8.41	90	102.52	90	P 2 ₁ /c (14)
293 ⁸⁶ ANTGEN26	9.45	6.00	8.54	90	103.51	90	P 2 ₁ /c (14)

Table 7 | Lattice parameters of the tetracene crystal

T (K)/CCDC entry	a	b	c	α	β	γ	Space group
295 ⁸⁷ TETCEN	7.90	6.03	13.53	100.3	113.2	86.3	P $\bar{1}$ (2)

Table 8 | Lattice parameters of the pentacene crystal

T (K)/CCDC entry	a	b	c	α	β	γ	Space group
295 ⁸⁷ PENCEN	7.90	6.06	16.01	101.9	112.6	85.8	P $\bar{1}$ (2)

Table 9 | Dataset pool for molecular crystals

Molecular crystal	Number of structures
Naphthalene	36859
Anthracene	200000
Tetracene	183123
Pentacene	85578

Only a subset is used for training the potentials.

the stopping criterion for active learning, ensuring consistency with the reference DFT method. The number of AL-iterations, the corresponding number of training structures, and the achieved errors in energies and forces for the G-MLIPs are reported in Supplementary Tables 1–6. The total number of acene molecular structures in generalized potentials is also summarized in Table 4.

Supplementary Fig. 2 presents the maximum and mean errors in Γ -point phonon frequencies observed throughout the active-learning steps of G-MLIPs. Additionally, for comparison with the stopping criterion employed in VASP MLIP, we evaluate the mean and maximum energy and force errors per active-learning step for N-MLIP, as shown in Supplementary Fig. 2. Once the initial training is complete, the committee of MLIPs makes predictions on the dataset pool. The 25 most uncertain atomic structures, based on the standard deviation of the predicted energies, are then selected. Their energies and forces are recomputed using the reference method and subsequently added to the training set. The committee is retrained on this expanded dataset, and the process is repeated to iteratively improve the model’s performance and accuracy.

Uncertainty estimation and propagation to harmonic phonons

In committee of M machine learning potentials, for any molecular structure A , the mean of committee is given by $\bar{y}(A)$ and the committee uncertainty is represented by its standard deviation $\sigma(A)$ ^{29,55}. The active learning strategy in this work uses the committee uncertainty on energy to select new training structure from the pool.

For the error propagation, we estimate committee uncertainties with more care. For a given atomic structure A , we calculated the mean committee prediction, $\bar{y}(A)$, and the associated committee uncertainty, $\sigma(A)$, along with the reference values $y_{\text{ref}}(A)$ for M committee members. Due to the limited number of training structures available to each committee member and to the small amount of committee members, the uncertainty is

rescaled using a factor α , computed as⁵⁵

$$\alpha^2 \equiv -\frac{1}{M} + \frac{M-3}{M-1} \frac{1}{N_{\text{test}}} \sum_{A \in \text{test}} \frac{|y_{\text{ref}}(A) - \bar{y}(A)|^2}{\sigma^2(A)}, \quad (1)$$

where N_{test} represents the total number of structure in the test set. Using this scaling factor, we rescale the predictions of each committee member as

$$y'_i(A) = \bar{y}(A) + \alpha[y_i(A) - \bar{y}(A)], \quad (2)$$

which are used to compute vibrational properties, enabling robust uncertainty estimation due to the finite sampling of atomic configuration space.

Rescaling is performed on the committee forces for each member such that the uncertainty is accurately propagated onto the observables using equation Eq. (2)

$$F'_i(A) = \bar{F}(A) + \alpha[F_i(A) - \bar{F}(A)]. \quad (3)$$

As an example, we determine α for forces as 2.8, on a validation set of 2100 diverse $1 \times 2 \times 2$ naphthalene crystal structures selected using FPS^{28,50}.

To propagate this uncertainty to the harmonic phonons, we compute the phonon frequencies individually for each committee member using these scaled forces. This method enables the intrinsic propagation of uncertainty from the forces to the squared-phonon frequencies as

$$F'_i(A) \longrightarrow [\omega_i(A)]^2. \quad (4)$$

The phonon frequency prediction of committee, is given as the average,

$$\bar{\omega}' = \frac{1}{M} \sum_{i=1}^M \omega_i(A). \quad (5)$$

$$\sigma_{\omega^2} = \sqrt{\frac{1}{M-1} \sum_{i=1}^M [\omega_i^2 - \bar{\omega}^2]^2}. \quad (6)$$

Since forces are directly proportional to the square of the frequencies, the uncertainty in forces propagates to the harmonic frequencies in the following way: $F \propto \omega^2 \rightarrow dF/d\omega = 2\omega$, so $\sigma_{\omega} = \sigma_{\omega^2}/|2\omega|$, where σ_{ω^2} is given by Eq. (6).

Uncertainty propagation to vibrational density of states

The mass-weighted VDOS is calculated for each committee member i with the corresponding scaled forces F_i from the velocity auto-correlation function (VACF)

$$C_{F_i}(t) = \lim_{T \rightarrow \infty} \frac{1}{T-t} \sum_j \int_0^{T-t} m_j \mathbf{v}_{F_i}^j(\tau) \cdot \mathbf{v}_{F_i}^j(\tau+t) d\tau, \quad (7)$$

where $\mathbf{v}_{F_i}^j(t)$ is a shorthand notation for the velocity of the j^{th} atom in Cartesian coordinates at time t , as predicted by the i^{th} committee member. The VDOS spectrum ($C_{F_i}(\omega)$) is obtained from the Fourier transform of the corresponding VACF. Atomic trajectories are obtained from NVE simulations following equilibration runs in the NVT ensemble. For the naphthalene molecular crystal, 100 simulations per committee member are run for 20 ps each at 80 K. In addition, with mean of the committee forces, 100 simulations are run for 20 ps at 80 K for naphthalene molecular crystal, while for the host-guest system, 27 simulations are conducted for 15 ps each at 100 K.

The statistical error, σ_{stat} is calculated by block averaging of VDOS over NVE runs, i.e.,

$$\sigma_{\text{stat}}(\omega) = \sqrt{\frac{1}{N-1} \sum_{k=1}^N [C_{\bar{F}}^k(\omega) - C_{\bar{F}}(\omega)]^2}, \quad (8)$$

where $\overline{C_{\mathbf{F}}(\omega)}$ is the block averaged VDOS calculated with mean of the committee forces $\overline{\mathbf{F}}$ and $C_{\mathbf{F}}^k(\omega)$ is the VDOS corresponding to k th trajectory obtained with the mean of the committee forces $\overline{\mathbf{F}}$, and N is the total number of NVE runs.

We estimate the committee error on the VDOS by calculating the standard deviation across all committee members as

$$\sigma_{\text{com}}(\omega) = \sqrt{\frac{1}{(M-1)} \sum_{i=1}^M \left[C_{\mathbf{F}_i}^i(\omega) - \overline{C_{\mathbf{F}}(\omega)} \right]^2}, \quad (9)$$

where $M = 8$ and we take the last term in the expression above as the average over the VDOS obtained by all committee members. In principle, the expression above also contains statistical error. We have checked that the statistical error is vastly smaller than the committee error.

Therefore we simply compute the corresponding committee error as,

$$\sigma_{\text{com}}^{e(\omega)} = \frac{\sigma_{\text{com}}(\omega)}{\sqrt{M}}. \quad (10)$$

Normal-mode projected VDOS

We follow the procedure outlined in ref. 85 to compute the normal-mode projected VV-ACF and the corresponding power spectra. We first calculate the normal-mode projected atomic velocities ($v_s(t)$) from the AIMD trajectories as

$$v_s(t) = \sum_j \sqrt{m_j} \mathbf{e}_j^*(s) \cdot \mathbf{v}_j(t), \quad (11)$$

where m_j is the mass of j th atom, $\mathbf{e}_j^*(s)$ is the polarization vector of the harmonic phonon at Γ -point, and $\mathbf{v}_j(t)$ are the atomic velocities in x , y and z directions. The normal-mode projected VDOS can then be calculated via Eq. (7) as $C_s(\omega) = \mathcal{F}\{\langle v_s(0)v_s(t) \rangle\}$. The total VDOS can be obtained via $C(\omega) = \sum_s C_s(\omega)$.

Here, we calculate the contributions of host and guest atoms to the VDOS by separating the normal-mode projected velocities as $v_s(t) = v_s^h(t) + v_s^g(t)$, where the last two terms are the contributions of host and guest atoms to the projected velocities, i.e., $v_s^{\text{host(guest)}}(t) = \sum_{j \in \text{host(guest)}} \sqrt{m_j} \mathbf{e}_j^*(s) \cdot \mathbf{v}(t)$. Hence, the power spectrum can be decomposed to three terms as

$$C_s(\omega) = C_s^{\text{host}}(\omega) + C_s^{\text{guest}}(\omega) + C_s^{\text{cross}}(\omega), \quad (12)$$

where the host(guest)-projected VDOS is $C_s^{\text{host(guest)}}(\omega) = \mathcal{F}\{\langle v_s^{\text{host(guest)}}(0)v_s^{\text{host(guest)}}(t) \rangle\}$, and the cross-correlated VDOS is $C_s^{\text{cross}}(\omega) = \mathcal{F}\{\langle v_s^{\text{guest}}(0)v_s^{\text{host}}(t) \rangle + \langle v_s^{\text{host}}(0)v_s^{\text{guest}}(t) \rangle\}$. The cross terms can be interpreted as a measure of the coupling between host and guest normal modes in the harmonic phonon basis.

Data availability

A Zenodo repository is available (<https://zenodo.org/records/17100564>) containing data and machine-learning potential models reported in this manuscript.

Received: 15 April 2025; Accepted: 2 October 2025;

Published online: 23 October 2025

References

- Congreve, D. N. et al. External quantum efficiency above 100% in a singlet-exciton-fission-based organic photovoltaic cell. *Science* **340**, 334–337 (2013).
- Toninelli, C. et al. Single organic molecules for photonic quantum technologies. *Nat. Mater.* **20**, 1615–1628 (2021).
- Alvertis, A. M. & Engel, E. A. Importance of vibrational anharmonicity for electron-phonon coupling in molecular crystals. *Phys. Rev. B* **105**, L180301 (2022).
- Devos, A. & Lannoo, M. Electron-phonon coupling for aromatic molecular crystals: possible consequences for their superconductivity. *Phys. Rev. B* **58**, 8236–8239 (1998).
- Vukmirović, N., Bruder, C. & Stojanović, V. M. Electron-phonon coupling in crystalline organic semiconductors: microscopic evidence for nonpolaronic charge carriers. *Phys. Rev. Lett.* **109**, 126407 (2012).
- Kato, T., Yoshizawa, K. & Hirao, K. Electron-phonon coupling in negatively charged acene- and phenanthrene-edge-type hydrocarbon crystals. *J. Chem. Phys.* **116**, 3420–3429 (2002).
- Neef, A. et al. Frontier orbitals control dynamical disorder in molecular semiconductors. arXiv <https://doi.org/10.48550/arXiv.2412.06030> (2024).
- Nyman, J. & Day, G. M. Static and lattice vibrational energy differences between polymorphs. *CrystEngComm* **17**, 5154–5165 (2015).
- Rossi, M., Gasparotto, P. & Ceriotti, M. Anharmonic and quantum fluctuations in molecular crystals: a first-principles study of the stability of paracetamol. *Phys. Rev. Lett.* **117**, 115702 (2016).
- Kryncki, M. & Rossi, M. Efficient Gaussian process regression for prediction of molecular crystals harmonic free energies. *npj Comput. Mater.* **7**, 169 (2021).
- Kapil, V. & Engel, E. A. A complete description of thermodynamic stabilities of molecular crystals. *Proc. Natl. Acad. Sci. USA* **119**, e2111769119 (2022).
- Hoja, J. et al. Reliable and practical computational description of molecular crystal polymorphs. *Sci. Adv.* **5**, eaau3338 (2019).
- Coropceanu, V. et al. Charge transport in organic semiconductors. *Chem. Rev.* **107**, 926–952 (2007).
- Chang, B. K., Zhou, J.-J., Lee, N.-E. & Bernardi, M. Intermediate polaronic charge transport in organic crystals from a many-body first-principles approach. *Npj Comput. Mater.* **8**, 63 (2022).
- Seiler, H. et al. Nuclear dynamics of singlet exciton fission in pentacene single crystals. *Sci. Adv.* **7**, eabg0869 (2021).
- Neef, A., Rossi, M., Wolf, M., Ernstorfer, R. & Seiler, H. On the role of nuclear motion in singlet exciton fission: the case of single-crystal pentacene. *Phys. Status Solidi A* **221**, 2300304 (2024).
- Gurlek, B., Sandoghdar, V. & Martin-Cano, D. Engineering long-lived vibrational states for an organic molecule. *Phys. Rev. Lett.* **127**, 123603 (2021).
- Basché, T., Moerner, W. E., Orrit, M. & Wild, U. P. *Single-Molecule Optical Detection, Imaging and Spectroscopy* (Verlag-Chemie, 1997).
- Gurlek, B. & Wang, D. Small but large: single organic molecules as hybrid platforms for quantum technologies. *Phys. Rev. Res.* **7**, 021001 (2025).
- Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
- 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).
- Jinnouchi, R., Karsai, F. & Kresse, G. On-the-fly machine learning force field generation: application to melting points. *Phys. Rev. B* **100**, 014105 (2019).
- Batzner, S. et al. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453 (2022).
- Batatia, I., Kovacs, D. P., Simm, G. N. C., Ortner, C. & Csanyi, G. MACE: higher order equivariant message passing neural networks for fast and accurate force fields. *Adv. Neural Inf. Process.* **35**, 11423–11436 (2022).

25. Deringer, V. L., Caro, M. A. & Csányi, G. Machine learning interatomic potentials as emerging tools for materials science. *Adv. Mater.* **31**, 1902765 (2019).
26. Smith, J. S. et al. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nat. Commun.* **10**, 2903 (2019).
27. Ko, T. W. & Ong, S. P. Recent advances and outstanding challenges for machine learning interatomic potentials. *Nat. Comput. Sci.* **3**, 998–1000 (2023).
28. Cersonsky, R. K., Helfrecht, B. A., Engel, E. A., Kliavinek, S. & Ceriotti, M. Improving sample and feature selection with principal covariates regression. *Mach. Learn. Sci. Technol.* **2**, 035038 (2021).
29. Schran, C., Brezina, K. & Marsalek, O. Committee neural network potentials control generalization errors and enable active learning. *J. Chem. Phys.* **153**, 104105 (2020).
30. Karabin, M. & Perez, D. An entropy-maximization approach to automated training set generation for interatomic potentials. *J. Chem. Phys.* **153**, 094110 (2020).
31. Allotey, J., Butler, K. T. & Thiyagalingam, J. Entropy-based active learning of graph neural network surrogate models for materials properties. *J. Chem. Phys.* **155**, 174116 (2021).
32. Chen, C. & Ong, S. P. A universal graph deep learning interatomic potential for the periodic table. *Nat. Comput. Sci.* **2**, 718–728 (2022).
33. Batatia, I. et al. A foundation model for atomistic materials chemistry. arXiv: <https://doi.org/10.48550/arXiv.2401.00096> (2023).
34. Kovács, D. P. et al. Mace-off: Short-range transferable machine learning force fields for organic molecules. *J. Am. Chem. Soc.* **147**, 17598–17611 (2025).
35. Merchant, A. et al. Scaling deep learning for materials discovery. *Nature* **624**, 80–85 (2023).
36. Allen, A. E. et al. Learning together: towards foundation models for machine learning interatomic potentials with meta-learning. *npj Comput. Mater.* **10**, 154 (2024).
37. Deng, B. et al. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nat. Mach. Intell.* **5**, 1031–1041 (2023).
38. Yang, H. et al. Mattersim: a deep learning atomistic model across elements, temperatures and pressures arXiv: <https://doi.org/10.48550/arXiv.2405.04967> (2024).
39. Schmidt, J., Marques, M. R. G., Botti, S. & Marques, M. A. L. Recent advances and applications of machine learning in solid-state materials science. *Npj Comput. Mater.* **5**, 83 (2019).
40. Verdi, C., Karsai, F., Liu, P., Jinnouchi, R. & Kresse, G. Thermal transport and phase transitions of zirconia by on-the-fly machine-learned interatomic potentials. *Npj Comput. Mater.* **7**, 156 (2021).
41. Wan, K., He, J. & Shi, X. Construction of high accuracy machine learning interatomic potential for surface/interface of nanomaterials—a review. *Adv. Mater.* **36**, e2305758 (2024).
42. Yang, Y., Zhang, S., Ranasinghe, K. D., Isayev, O. & Roitberg, A. E. Machine learning of reactive potentials. *Annu. Rev. Phys. Chem.* **75**, 371–395 (2024).
43. Bartók, A. P., Kermode, J., Bernstein, N. & Csányi, G. Machine learning a general-purpose interatomic potential for silicon. *Phys. Rev. X* **8**, 041048 (2018).
44. Loew, A., Wang, H.-C., Cerqueira, T. F. T. & Marques, M. A. L. Training machine learning interatomic potentials for accurate phonon properties. *Mach. Learn. Sci. Technol.* **5**, 045019 (2024).
45. Bandi, S., Jiang, C. & Marianetti, C. A. Benchmarking machine learning interatomic potentials via phonon anharmonicity. *Mach. Learn. Sci. Technol.* **5**, 030502 (2024).
46. Lee, H., Hegde, V. I., Wolverton, C. & Xia, Y. Accelerating high-throughput phonon calculations via machine learning universal potentials. *Mater. Today Phys.* **53**, 101688 (2025).
47. Monserrat, B., Brandenburg, J. G., Engel, E. A. & Cheng, B. Liquid water contains the building blocks of diverse ice phases. *Nat. Commun.* **11**, 5757 (2020).
48. George, J., Hautier, G., Bartók, A. P., Csányi, G. & Deringer, V. L. Combining phonon accuracy with high transferability in Gaussian approximation potential models. *J. Chem. Phys.* **153**, 044104 (2020).
49. Niblett, S. P., Kourtis, P., Magdau, I.-B., Grey, C. P. & Csányi, G. Transferability of data sets between machine-learned interatomic potential algorithms. *J. Chem. Theory Comput.* **21**, 6096–6112 (2025).
50. Eldar, Y., Lindenbaum, M., Porat, M. & Zeevi, Y. The farthest point strategy for progressive image sampling. *IEEE Trans. Image Process.* **6**, 1305–1315 (1997).
51. Zhu, A., Batzner, S., Musaelian, A. & Kozinsky, B. Fast uncertainty estimates in deep learning interatomic potentials. *J. Chem. Phys.* **158**, (2023).
52. Heid, E., Schörghuber, J., Wanzenböck, R. & Madsen, G. K. H. Spatially resolved uncertainties for machine learning potentials. *J. Chem. Inf. Model.* **64**, 6377–6387 (2024).
53. Venturi, S., Jaffe, R. L. & Panesi, M. Bayesian machine learning approach to the quantification of uncertainties on ab initio potential energy surfaces. *J. Phys. Chem. A* **124**, 5129–5146 (2020).
54. Musil, F., Willatt, M. J., Langovoy, M. A. & Ceriotti, M. Fast and accurate uncertainty estimation in chemical machine learning. *J. Chem. Theory Comput.* **15**, 906–915 (2019).
55. Imbalzano, G. et al. Uncertainty estimation for molecular dynamics and sampling. *J. Chem. Phys.* **154**, 074102 (2021).
56. Kellner, M. & Ceriotti, M. Uncertainty quantification by direct propagation of shallow ensembles. *Mach. Learn. Sci. Technol.* **5**, 035006 (2024).
57. Bauer, S. et al. Roadmap on data-centric materials science. *Model. Simul. Mater. Sci. Eng.* **32**, 063301 (2024).
58. Kubo, R. Statistical-mechanical theory of irreversible processes. I. general theory and simple applications to magnetic and conduction problems. *J. Phys. Soc. Jpn.* **12**, 570–586 (1957).
59. Kummer, S., Bräuchle, C. & Basché, T. Optical spectroscopy of single pentacene molecules in a naphthalene crystal. *Mol. Cryst. Liq. Cryst.* **283**, 255–260 (1996).
60. Steiner, M. O. E., Pedernales, J. S. & Plenio, M. B. Pentacene-doped naphthalene for levitated optomechanics. arXiv: <https://doi.org/10.48550/arXiv.2405.13869> (2024).
61. Skinner, J. L. Theory of pure dephasing in crystals. *Annu. Rev. Phys. Chem.* **39**, 463–478 (1988).
62. Myers, A. B., Tchenio, P., Zgierski, M. Z. & Moerner, W. E. Vibronic spectroscopy of individual molecules in solids. *J. Chem. Phys.* **98**, 10377–10390 (1994).
63. Zirkelbach, J. et al. High-resolution vibronic spectroscopy of a single molecule embedded in a crystal. *J. Chem. Phys.* **156**, 104301 (2022).
64. Nazir, A. & McCutcheon, D. P. Modelling exciton–phonon interactions in optically driven quantum dots. *J. Phys. Condens. Matter* **28**, 103002 (2016).
65. Fleischhauer, H.-C., Kryschi, C., Wagner, B. & Kupka, H. Pseudolocal phonons in p-terphenyl: pentacene single crystals. *J. Chem. Phys.* **97**, 1742–1749 (1992).
66. Bordat, P. & Brown, R. Elucidation of optical switching of single guest molecules in terylene/p-terphenyl mixed crystals. *Chem. Phys. Lett.* **331**, 439–445 (2000).
67. Deperasinska, I. & Kozankiewicz, B. Non-planar distortion of terylene molecules in a naphthalene crystal. *Chem. Phys. Lett.* **684**, 208–211 (2017).
68. Dlott, D. D. Dynamics of molecular crystal vibrations. In *Laser Spectroscopy of Solids II*, (ed. Yen, W. M.) 167–200 (Springer-Verlag, 1989).
69. Blum, V. et al. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **180**, 2175–2196 (2009).

70. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558 (1993).
71. Kresse, G. & Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metal–amorphous–semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251 (1994).
72. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
73. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).
74. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
75. Tkatchenko, A., DiStasio, R. A., Car, R. & Scheffler, M. Accurate and efficient method for many-body van der Waals interactions. *Phys. Rev. Lett.* **108**, 236402 (2012).
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. Togo, A. First-principles phonon calculations with phonopy and phono3py. *J. Phys. Soc. Jpn.* **92**, 012001 (2023).
78. Capelli, S. C., Albinati, A., Mason, S. A. & Willis, B. T. Molecular motion in crystalline naphthalene: analysis of multi-temperature x-ray and neutron diffraction data. *J. Phys. Chem. A*. **110**, 11695–11703 (2006).
79. Deringer, V. L., Caro, M. A. & Csányi, G. A general-purpose machine-learning force field for bulk and nanostructured phosphorus. *Nat. Commun.* **11**, 5461 (2020).
80. Anstine, D. M., Zubatyuk, R. & Isayev, O. Aimnet2: a neural network potential to meet your neutral, charged, organic, and elemental-organic needs. *Chem. Sci.* **16**, 10228–10244 (2025).
81. Sivaraman, G. et al. Machine-learned interatomic potentials by active learning: amorphous and liquid hafnium dioxide. *Npj Comput. Mater.* **6**, 104 (2020).
82. Stolte, N., Daru, J., Forbert, H., Marx, D. & Behler, J. Random sampling versus active learning algorithms for machine learning potentials of quantum liquid water. *J. Chem. Theory Comput.* **21**, 886–899 (2025).
83. Kulichenko, M. et al. Uncertainty-driven dynamics for active learning of interatomic potentials. *Nat. Comput. Sci.* **3**, 230–239 (2023).
84. Jinnouchi, R., Miwa, K., Karsai, F., Kresse, G. & Asahi, R. On-the-fly active learning of interatomic potentials for large-scale atomistic simulations. *J. Phys. Chem. Lett.* **11**, 6946–6955 (2020).
85. Sun, T., Zhang, D.-B. & Wentzcovitch, R. M. Dynamic stabilization of cubic CaSiO₃ perovskite at high temperatures and pressures from ab initio molecular dynamics. *Phys. Rev. B* **89**, 094109 (2014).
86. Asher, M. et al. Anharmonic lattice vibrations in small-molecule organic semiconductors. *Adv. Mat.* **32**, 1908028 (2020).
87. Campbell, R., Robertson, J. M. & Trotter, J. The crystal structure of hexacene, and a revision of the crystallographic data for tetracene. *Acta Crystallogr.* **15**, 289–290 (1962).

Acknowledgements

We acknowledge support from the Cluster of Excellence “CUI: Advanced Imaging of Matter” —EXC 2056—project ID 390715994, BiGmax, the Max Planck Society Research Network on Big-Data-Driven Materials-Science and the Max Planck-New York City Center for Non-Equilibrium Quantum Phenomena. The Flatiron Institute is a division of the Simons Foundation. We also acknowledge support from the European Research Council MSCA-ITN TIMES under grant agreement 101118915. S.S. and P.L. acknowledge support from the UFAST International Max Planck Research School.

Author contributions

B.G. and S.S. trained the potentials, implemented workflows, calculated, plotted and analyzed results. B.G., S.S., P.L., and M.R. discussed and interpreted the results. B.G. and M.R. designed and supervised research. M.R. and A.R. acquired funding. B.G. and S.S. wrote the first draft of the manuscript. B.G. and M.R. finalized writing the manuscript with input from all authors.

Funding

Open Access funding enabled and organized by Projekt DEAL.

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/s41524-025-01825-w>.

Correspondence and requests for materials should be addressed to Mariana Rossi.

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

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

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

© The Author(s) 2025