Editors' Suggestion

Bridging surface dynamics and phase transitions in $CsPbI_3$ perovskite with machine-learning force fields and dispersion interactions

Iryna Knysh,[1,*,†] Anton Charkin-Gorbulin,[1,*] Claudio Quarti,[2] David Beljonne,[2] Igor Poltavsky,[1,‡] and Alexandre Tkatchenko[1,§]

[1]*Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg City, Luxembourg*

[2]*Laboratory for Chemistry of Novel Materials, University of Mons, B-7000 Mons, Belgium*

(Received 28 April 2026; accepted 15 October 2026; published 16 November 2026)

In the rapidly evolving field of machine learning (ML) for atomistic simulations of materials and molecules, achieving transferability across different systems remains a key challenge. In this contribution, we demonstrate a case of transfer learning in which an ML model trained exclusively on surface slab data of the cubic (α) phase of $CsPbI_3$ perovskite can reproduce the key features of the dynamic behavior for different phases of the bulk material. Using an equivariant message-passing network (MACE) architecture, we construct an efficient ML force field trained on a dataset obtained from *ab initio* molecular dynamics simulations of a $CsPbI_3$ slab. The resulting model yields stable and reliable trajectories for both slab and bulk, reproducing the experimentally observed phase transitions of $CsPbI_3$. We demonstrate that pairwise dispersion methods (TS/D3) mispredict the correct phase of $CsPbI_3$ bulk at ambient conditions, while more advanced approaches (MBD-NL/rVV10) recover both the experimentally observed phase sequence with increasing temperature (orthorhombic–tetragonal–cubic) and the corresponding phase-transition temperatures. Through a local analysis of the rotational driving forces acting on PbI_6 octahedra, we show that this discrepancy can be traced to the strong sensitivity of Cs and I atomic forces to the treatment of van der Waals (vdW) dispersion. Overall, our results demonstrate that surface–slab datasets augmented with an accurate treatment of vdW dispersion provide a robust foundation for constructing transferable machine-learning force fields that consistently describe both surface and bulk behavior across multiple length scales.

DOI: 10.1103/qcvb-y84n

I. INTRODUCTION

Halide perovskites (HPs) have emerged as promising materials for next-generation photovoltaics, known for their exceptional properties such as tunable band gaps, high charge-carrier mobilities, and defect tolerance, together with comparatively low fabrication costs versus conventional silicon-based semiconductors [1–9]. HPs adopt the ABX_3 stoichiometry, where A is an inorganic or organic monovalent cation, B is a divalent metal, and X is a halide. Over the past decade, record power-conversion efficiencies (PCEs) for hybrid organic-inorganic perovskite solar cells (PSCs) have surpassed 25% [10,11], while inorganic $CsPbI_3$ PSCs have approached 22% [12]. Inorganic HPs are particularly attractive because they exhibit enhanced thermal stability and resistance toward atmospheric conditions [13]. However, long-term stability of HP devices remains a central challenge limiting their commercial use. HPs are soft materials and prone to phase instabilities that strongly affect their structural and electronic properties [14–17]. For example, $CsPbI_3$ transforms on cooling from the high-temperature perovskite cubic α phase through the tetragonal β phase to the orthorhombic γ phase, and near room temperature it spontaneously converts to the optically inactive, nonperovskite δ phase in the absence of stabilization techniques [14,18–21]. In parallel, surface defects in perovskite films are considered to be the driving force for the nonradiative charge recombination [22] and associated PCE losses, though their formation mechanisms remain not well understood [23–28]. Revealing structural dynamics and ionic transport at the atomistic level and their impact on electronic properties is therefore crucial for improving the stability and performance of PSCs.

First-principles calculations have provided key insights into how structure impacts the electronic properties of HPs [4,29–34]. Such studies are typically carried out at density functional theory (DFT) [35] or many-body Green's function (GW) [36] levels of theory, which impose limitations on the sizes of the employed atomistic models [4,30]. In practice, HPs are soft, dynamically disordered materials whose structure-property relationships unfold over timescales and length scales that remain inaccessible to *ab initio* simulations. In this context, machine-learning force fields (MLFFs) trained on accurate *ab initio* data provide a needed solution at affordable costs. For example, Jinnouchi *et al.* [37]

*These authors contributed equally to this work.
†Contact author: iryna.knysh@uni.lu
‡Contact author: igor.poltavskyi@uni.lu
§Contact author: alexandre.tkatchenko@uni.lu

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

developed an on-the-fly machine learning scheme to enable finite-temperature simulations of HPs over extended timescales. Braeckevelt *et al.* [38] proposed a protocol to determine the γ to δ phase-transition temperature of $CsPbI_3$ by combining highly accurate electronic energies from random phase approximation and Hartree-Fock (RPA+HF) with vibrational free energies obtained from molecular dynamics (MD) driven by a phase-transferable MLFF on a supercell of up to 2560 atoms. Fransson *et al.* [39] developed a third-generation neuroevolution potential for $CsPbX_3$ (with X = I, Br, and Cl), trained on *ab initio* data covering the γ, β, α, and δ phases, and used the GPUMD package to perform large-scale MD simulations of phase transitions in lead HPs (LHPs).

In LHPs, phase transitions are closely linked to the correlated tilting patterns of the PbX_6 octahedra, commonly referred to as octahedral tilting, which are governed by the soft lattice modes [40]. The temperature-dependent distributions and correlations of these tilts therefore serve as sensitive descriptors of polymorphism in LHPs, and a range of such descriptors has been proposed to characterize this tilting behavior [41–44]. At the same time, both experiment and theory have shown that the local structure in LHPs can deviate substantially from the average structure [45–47]. Therefore, understanding the local structure of these materials, and consequently their stability, requires accurate dynamical simulations over long length and timescales, which recent advances in MLFFs have made possible. In this context, Baldwin *et al.* [48] constructed an MLFF for $CsPbI_3$ using the atomic cluster expansion (ACE) framework to investigate its dynamic local structure. This ACE-based MLFF enables large-scale finite-temperature MD simulations, through which they analyze the spatial and temporal correlations of octahedral tilts, revealing transient planar domains and confirming that the local structure of $CsPbI_3$ remains highly dynamic in both space and time, even in the high-symmetry cubic phase, in line with earlier *ab initio* MD (AIMD) simulations performed on methylammonium lead iodide perovskite [49]. Recently, Liang *et al.* [50] employed an equivariant message-passing neural network potential within the MACE framework [51,52], trained with on-the-fly DFT data generation, to run large-scale MD simulations of three prototypical LHPs—$CsPbX_3$, $MAPbX_3$, and $FAPbX_3$. Their results show that the A-site cation, halide composition, and local atomic environments influence octahedral tilting patterns, which in turn affect phase stability and phase transitions in mixed-halide perovskites.

Notably, previous MLFFs were developed primarily for bulk materials, yet realistic modeling of PSCs requires a unified description of both bulk and surface regions [53]. Here, we explore slab-to-bulk transfer learning by training a MACE-based model on α-$CsPbI_3$ surface-slab data. The choice is physically motivated because slab structures naturally sample a broader spectrum of local environments: inner layers preserve bulk-like coordination, while surface layers introduce symmetry breaking and enhanced structural fluctuations. As a result, slab-based training sets can provide a richer representation of the local energetics governing finite-temperature phase behavior than bulk-only datasets. We show that our model trained exclusively on surface-slab data reproduces the experimentally observed orthorhombic-tetragonal-cubic phase transitions of $CsPbI_3$ bulk while also yielding stable surface dynamics.

Importantly, achieving such a result requires a reliable treatment of vdW dispersion interactions. Consistent with this, previous studies [39,43] have demonstrated that both the choice of exchange-correlation functional and the description of vdW dispersion critically affect the structural and dynamical properties of perovskites. However, the underlying mechanisms by which these factors influence the system's behavior remain largely unresolved. Here, we systematically investigate the role of vdW dispersion by comparing three dispersion treatments combined with the PBE functional [54]—D3 correction [55], the pairwise Tkatchenko–Scheffler (TS) scheme [56], and the nonlocal many-body dispersion approach (MBD-NL) [57]—alongside the vdW density functional r^2SCAN+rVV10 [58], with a focus on their impact on lattice parameters and phase-transition temperatures. We show that an accurate description of the phase diagram requires going beyond pairwise vdW treatments, as many-body dispersion significantly perturbs the Cs and I atomic forces and, consequently, the rotational driving forces acting on the PbI_6 framework, thereby influencing the structure and finite-temperature dynamics of the material.

II. DATASET GENERATION

In this section, we describe procedures used to generate datasets employed in this study. Details of the model architecture and training parameters are provided in the Supplemental Material (SM; see Sec. S2 for more details) [59].

Slab dataset. Throughout this work, slabs are denoted as $(n\sqrt{2} \times n\sqrt{2})$-*m*L, where the lateral dimensions are expressed relative to the pseudocubic unit cell, the in-plane lattice vectors are rotated by 45° relative to the pseudocubic axes, and *m*L denotes the number of PbI_6 octahedral layers. We generated two slab datasets using a CsI-terminated $(4\sqrt{2} \times 4\sqrt{2})$-5L $CsPbI_3$ slab containing 864 atoms, with its surface parallel to the *xy* plane [60]. We chose a symmetric CsI-terminated slab because previous *ab initio* calculations showed that the CsI termination is thermodynamically more stable than PbI_2 within the stability range of $CsPbI_3$ bulk [61]. Both datasets comprise the same 600 configurations but differ in their reference energy and force labels, calculated at the PBE+D3 and PBE+MBD-NL levels of theory.

The configurations were sampled from two AIMD trajectories performed at 300 and 600 K, each initiated from the relaxed slab structure with supercell dimensions $a = 35.07$ Å, $b = 35.29$ Å, and $c = 50.01$ Å, where c includes the vacuum region. Simulations were performed with CP2K [62,63] for 44 ps using a 4 fs time step in the isothermal-isobaric NPT ensemble, allowing fluctuations only in the in-plane cell components. The system was coupled to a Nosé-Hoover thermostat and barostat, with time constants set to 20 fs and 1000 fs, respectively, to maintain the target temperature (300 or 600 K) and a pressure of 1 bar [64,65]. The AIMD simulations employed the PBE exchange-correlation functional [54], the D3 dispersion correction, DZVP-MOLOPT-SR-GTH basis sets [66], and norm-conserving GTH-PBE pseudopo-

tentials [67,68]. The mixed Gaussian and plane-wave scheme employed a relative cutoff of 50 Ry, a plane-wave cutoff of 300 Ry, and 16 multigrid levels. A surface-dipole correction was included. A total of 300 configurations were selected from each trajectory using the energy-based sampling procedure implemented in the sGDML package [69–71], resulting in a total of 600 configurations.

For `slab-PBE+D3`, the reference energies and forces were taken directly from the PBE+D3 AIMD trajectories. For `slab-PBE+MBD-NL`, the same configurations were recalculated at the PBE+MBD-NL level using the FHI-aims code [72]. For these DFT calculations, a $1 \times 1 \times 1$ Γ-centered k-point grid and the FHI-aims tight species defaults were used together with Gaussian smearing of 0.05 eV for the fractional occupations. The self-consistent field (SCF) procedure employed convergence criteria of 10^{-5} for the charge density, 10^{-5} eV for the eigenvalue sum, 10^{-6} eV for the total energy, and 10^{-4} eV Å^{-1} for the forces. A dipole correction was included.

Bulk dataset. For the large-scale bulk MD simulations, supercells are denoted as $n_a \times n_b \times n_c$ repetitions of a 20-atom orthorhombic-compatible reference cell, whose lattice vectors correspond to $\sqrt{2} \times \sqrt{2} \times 2$ in pseudocubic units. Thus, for example, a $5 \times 5 \times 5$ bulk supercell contains 2500 atoms.

The bulk dataset consists of 19 580 $CsPbI_3$ structures, including γ, β, α, and δ polymorphs, obtained from AIMD simulations with the r^{2}SCAN+rVV10 [58] functional, as described in Ref. [73]. AIMD simulations in the NPT ensemble were performed for each phase using 40-atom supercells over a temperature range 200–650 K with 100 K intervals. From this full dataset, we constructed `bulk-r`^{2}`SCAN+rVV10` by selecting 392 structures, for which energies and forces were kept at the r^{2}SCAN+rVV10 level of theory. In addition, we generated the `bulk-PBE+MBD-NL`, `bulk-PBE+TS`, and `bulk-PBE+D3` datasets by taking the same structures from `bulk-r`^{2}`SCAN+rVV10` and recalculating their energies and forces at the PBE+MBD-NL, PBE+TS, and PBE+D3 levels of theory, respectively, using FHI-aims. We employed the same FHI-aims settings as before for the `slab-PBE+MBD-NL` dataset, except for the k-point sampling. To match the k-point setup used in Ref. [73] and to ensure better uniformity across different phases, we used a `k_grid_density` keyword in the FHI-aims code, setting it to 5.0 along each of the three reciprocal axes of the first Brillouin zone.

DFT codes. In this work, we use datasets generated with CP2K and FHI-aims codes, together with a VASP dataset taken from the literature. The choice of CP2K for the slab AIMD simulations is motivated by its efficient MD implementation, which is particularly important for the computationally demanding slab systems considered here. The FHI-aims calculations are required because the MBD-NL method is not available in the CP2K code. The bulk r^{2}SCAN+rVV10 dataset is taken from Ref. [73] and was originally calculated using VASP. Although previous benchmark studies have demonstrated a high degree of reproducibility among DFT calculations performed with different codes [74], we explicitly assessed the possible influence of the code choice on our results. In Fig. S2 [59], we compare PBE forces computed with CP2K and FHI-aims, as well as PBE+D3 (CP2K) and PBE+MBD-NL (FHI-aims) forces. The deviation between the PBE forces obtained with CP2K and FHI-aims is substantially smaller (12 meV/Å) than the difference between the PBE+D3 and PBE+MBD-NL forces (48 meV/Å). These results demonstrate that the use of different DFT codes does not impact our conclusions regarding the influence of the vdW methods.

III. RESULTS AND DISCUSSION

The predictive accuracy of an MLFF is intrinsically linked to the quality of its reference training data. Perovskites are soft, strongly anharmonic materials with a shallow octahedral-tilt energy landscape [75], where small changes in long-range interactions can shift equilibrium structures and relative phase stability [76]. Accurate DFT modeling, therefore, requires an explicit treatment of dispersion to reproduce lattice parameters and tilt energetics quantitatively. We therefore begin by examining how the choice of vdW dispersion treatment influences relaxed structure and atomic forces of the $CsPbI_3$ slab. We then perform finite-temperature MD simulations using MLFFs trained on PBE+D3 and PBE+MBD-NL single-phase slab datasets, and show that only the latter stabilizes the expected γ-$CsPbI_3$ phase under ambient conditions. Because slab datasets capture a broader range of structural motifs and local environments, we assess whether an MLFF trained on a single-phase slab dataset can transfer to bulk $CsPbI_3$ and reproduce its temperature-driven phase transitions. Finally, we systematically compare pairwise (D3 and TS), beyond-pairwise (MBD-NL), and nonlocal density-based (rVV10) methods, demonstrating that more advanced dispersion treatments are essential for accurately capturing the bulk dynamics, octahedral tilting patterns, and the resulting phase-transition temperatures of $CsPbI_3$.

A. Dispersion effects in $CsPbI_3$ slabs

In slab geometries, surface formation reduces atomic coordination numbers and introduces an abrupt change in dielectric screening toward vacuum, which perturbs bulk tilt motifs and drives a competition between surface relaxations and bulk-like tilting. Thus, $CsPbI_3$ slabs provide a sensitive probe of dispersion effects [77].

Figure 1 compares the bulk reference structures with the corresponding relaxed CsI-terminated slab structures computed at the PBE level using either the pairwise D3 correction or the screening-aware many-body MBD-NL treatment of dispersion. The PBE+D3 structures were relaxed using CP2K, whereas the PBE+MBD-NL structures were relaxed using FHI-aims, since MBD-NL is not available in CP2K (see Sec. S1 in the SM) [59]. The bulk reference structures are shown in Fig. 1(a), while the layer-resolved relaxed slab geometries are shown in Fig. 1(b). Even for the bulk reference structures, PBE+MBD-NL yields a unit-cell volume that is 7.7% smaller than that of PBE+D3, consistent with prior benchmarks showing that MBD-NL typically produces more compact structures for soft materials [57]. After introducing the surfaces and relaxing the slab, both dispersion treatments lead to a contraction of the equivalent CsI-plane separation [Fig. 1(c)]. For PBE+D3, this separation decreases from 31.8

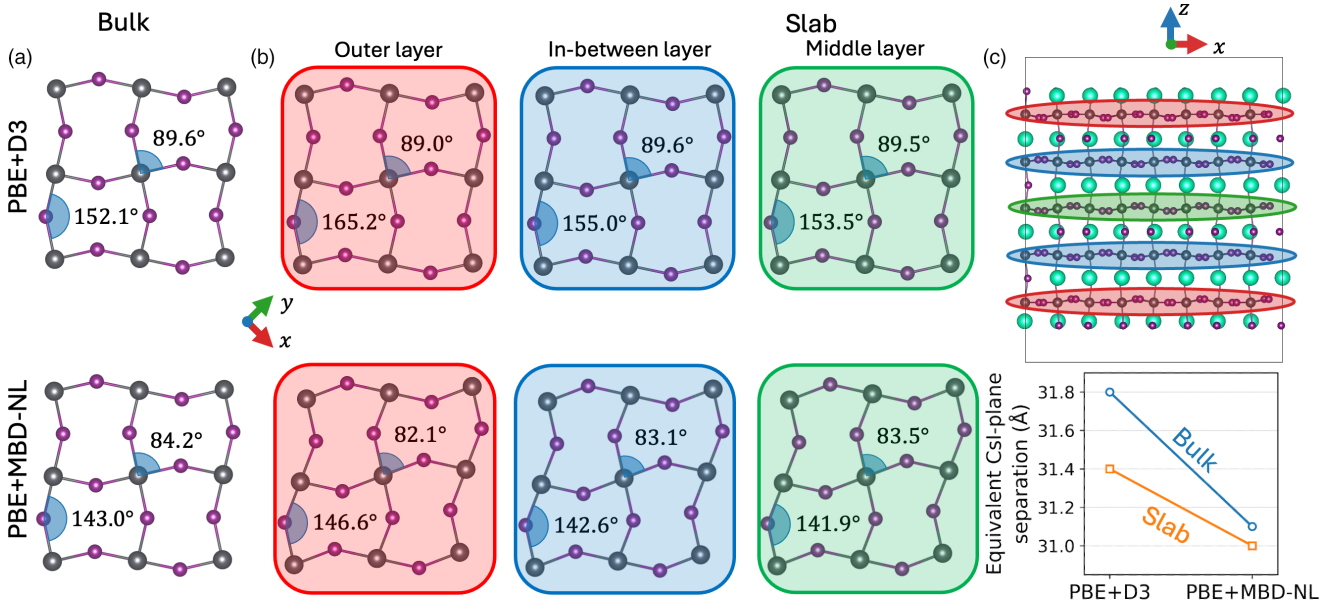


FIG. 1. Structural relaxation of CsI-terminated $CsPbI_3$ slabs computed with PBE+D3 and PBE+MBD-NL. (a) Corresponding bulk reference structures and I–Pb–I and Pb–I–Pb angles. (b) Relaxed slab structures with layer-resolved I–Pb–I and Pb–I–Pb angles for the outermost, intermediate, and middle layers, highlighted in red, blue, and green, respectively. The first row in panels (a) and (b) corresponds to the results obtained with PBE+D3, whereas the second row shows the results obtained with PBE+MBD-NL. (c) Layer assignment in the slab along the z axis (direction of the vacuum layer) and comparison of the equivalent CsI-plane separation in the bulk reference and relaxed slab. The equivalent CsI-plane separation is defined as the distance between the CsI planes corresponding to the two slab surfaces.

to 31.4 Å, corresponding to a contraction of 1.3% (0.4 Å). For PBE+MBD-NL, the contraction is smaller, from 31.1 to 31.0 Å (0.1 Å). The surface relaxation also modifies the layer-resolved octahedral angles. For PBE+D3, the outermost-layer I–Pb–I angle decreases from the corresponding bulk reference value of ∼89.6° to ∼89.0°, while the Pb–I–Pb angle increases from ∼152.1° to ∼165.2°. For PBE+MBD-NL, the corresponding changes are from ∼84.2° to ∼82.1° for I–Pb–I and from ∼143.0° to ∼146.6° for Pb–I–Pb. The central-layer I–Pb–I and Pb–I–Pb angles deviate from the corresponding bulk reference values by 0.1° and 1.4° for PBE+D3, and by 0.7° and 1.1° for PBE+MBD-NL, respectively. These deviations remain below 1.5° for both dispersion treatments, showing that, for the structural parameters considered, the interior region closely approaches the corresponding bulk geometry. Inclusion of MBD-NL yields a denser equilibrium structure and a smaller slab-normal contraction than D3. The surface-induced angular response is method-dependent: D3 produces a larger change in the outermost-layer Pb–I–Pb angle, whereas MBD-NL produces a larger change in the outermost-layer I–Pb–I angle. Note that these 0 K relaxations neglect thermal expansion and anharmonic effects present in the experimental data.

To identify the origin of the structural differences discussed above, Fig. 2 compares atomic-force components obtained with PBE+D3 and PBE+MBD-NL for identical slab configurations (datasets `slab-PBE+D3` and `slab-PBE+MBD-NL`, see Sec. II for details). The dashed black line indicates perfect agreement ($y = x$); each panel reports root-mean-square deviation (RMSD), mean absolute deviation (MAD), and the squared Pearson correlation coefficient (R^2). Species-resolved parity plots (top row) show close agreement for Pb (RMSD 0.035 eV Å^{-1}, MAD 0.028 eV Å^{-1}, and $R^2 = 0.991$) and I (RMSD 0.053 eV Å^{-1}, MAD 0.043 eV Å^{-1}, and $R^2 = 0.959$). In contrast, Cs exhibits substantially larger deviations (RMSD 0.094 eV Å^{-1} and MAD 0.077 eV Å^{-1}) and significantly reduced correlation ($R^2 = 0.739$). This is consistent with its large ionic radius and weak, nondirectional bonding environment, making Cs markedly sensitive to long-range dispersion effects.

Figure 2(b) shows that the agreement between the pairwise and many-body vdW forces is poorest for Cs atoms in the surface layers, with the RMSD of 0.108 eV Å^{-1}. The dominant discrepancy is along the slab-normal z direction with the RMSD of 0.146 eV Å^{-1} [Fig. 2(c)], whereas the in-plane x and y components are smaller and comparable (RMSD 0.082 eV Å^{-1}). We further isolated the dispersion-induced ΔF^{D3} and $\Delta F^{\mathrm{MBD-NL}}$ contributions by subtracting the PBE forces from the PBE+D3 and PBE+MBD-NL forces (see Fig. S3) [59]. The two corrections are strongly anticorrelated and mostly act in opposite directions. Because the dispersion contributions are small relative to the underlying PBE forces, the PBE+D3 and PBE+MBD-NL forces remain largely aligned (see Fig. S4) [59]. Nevertheless, the different dispersion treatment can strengthen, weaken, or slightly reorient total forces. Moreover, the PBE+D3 force components along the z direction are generally larger in magnitude than those from PBE+MBD-NL, with negative forces on the top surface layer and positive forces on the bottom surface layer (for more details, see Fig. S5) [59]. This observation is consistent with the stronger slab contraction after surface relaxation predicted by the pairwise dispersion approach in Fig. 1. Overall, these layer- and component-resolved trends suggest that the choice of dispersion treatment impacts atomic

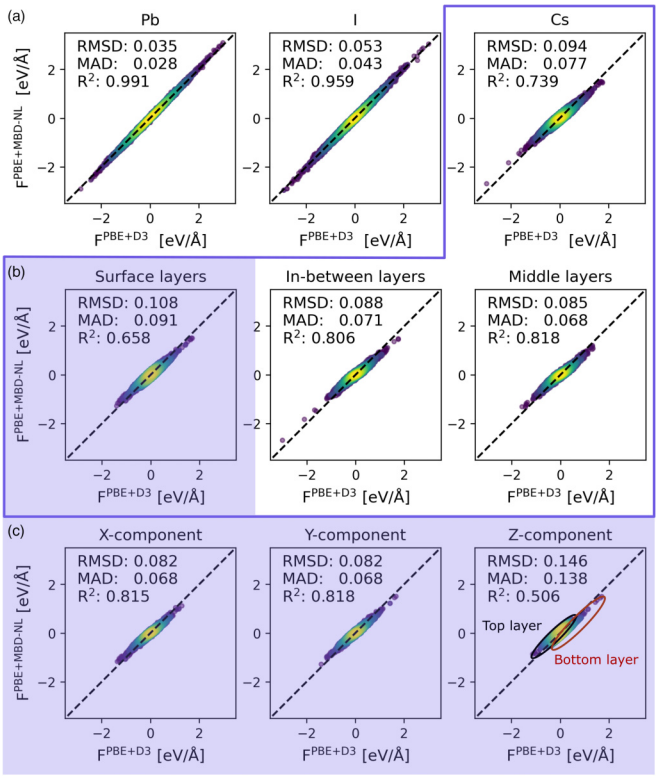


FIG. 2. Comparison of force components computed with PBE+D3 and PBE+MBD-NL for identical slab configurations: (a) species-resolved forces for Pb, I, and Cs; (b) Cs forces by slab region; (c) force components for surface-layer Cs. Colors represent point density (blue → yellow for increasing density), and the dashed black line denotes perfect agreement. Insets report RMSD, MAD, and R^2, with deviations given in eV Å^{-1}.

forces more strongly for Cs in the weakly coordinated and weakly screened surface environment. This is consistent with D3 augmenting PBE via a pairwise atom–atom correction, whereas MBD-NL captures screened, collective many-body dispersion, which can be especially important in heterogeneous regions such as slab surfaces.

These force differences are expected to have an influence on the structure at finite temperature, since Cs occupies the soft A-site cavity and couples to the PbI_6 tilt framework through weak, long-range interactions. We therefore test whether dispersion-induced changes in Cs forces translate into distinct slab lattice responses and tilt-driven phase behavior by running MD with MLFFs trained on `slab-PBE+D3` and `slab-PBE+MBD-NL` datasets. Using these two MLFF models, finite-temperature slab behavior was probed via the NPT MD protocol and heating schedule detailed in Sec. S3 of the SM [59] for an $(8\sqrt{2} \times 8\sqrt{2})$-5L $CsPbI_3$ slab ($N = 3456$ atoms). Figure 3 shows the temperature evolution of the in-plane lattice parameters a and b, together with time-averaged structures evaluated over selected temperature windows. For the MACE@slab-PBE+MBD-NL model [Fig. 3(a)], a and b lattice parameters respond anisotropically to heating and intersect near 375 K. The time-averaged snapshots indicate an orthorhombic structure at 300–325 K and a more cubic-like (i.e., reduced octahedral tilting) structure at 375–400 K, consistent with a gradual $\gamma \rightarrow \alpha$ transition without a clearly resolved intermediate β-phase. The earlier appearance of a cubic-like structure with MBD-NL is qualitatively consistent with reports that confinement, strain, and interfacial constraints can stabilize the α phase of $CsPbI_3$ at lower temperatures than in the bulk [78,79], where the α-phase is stabilized near 600 K [14]. In contrast, the MACE@slab-PBE+D3 [Fig. 3(b)] remains cubic-like over the full 300–550 K range, implying a stronger stabilization of the high-symmetry perovskite structure than suggested by bulk measurements. Analysis of the PBE+D3 AIMD trajectories at 300 and 600 K, which were used to construct the `slab-PBE+D3` and `slab-PBE+MBD-NL` datasets, shows that the slab remains cubic-like at both temperatures. The similar behavior obtained with the MACE@slab-PBE+D3 therefore demonstrates consistency between the MLFF and its underlying reference data. Notably, even though the MACE@slab-PBE+MBD-NL was

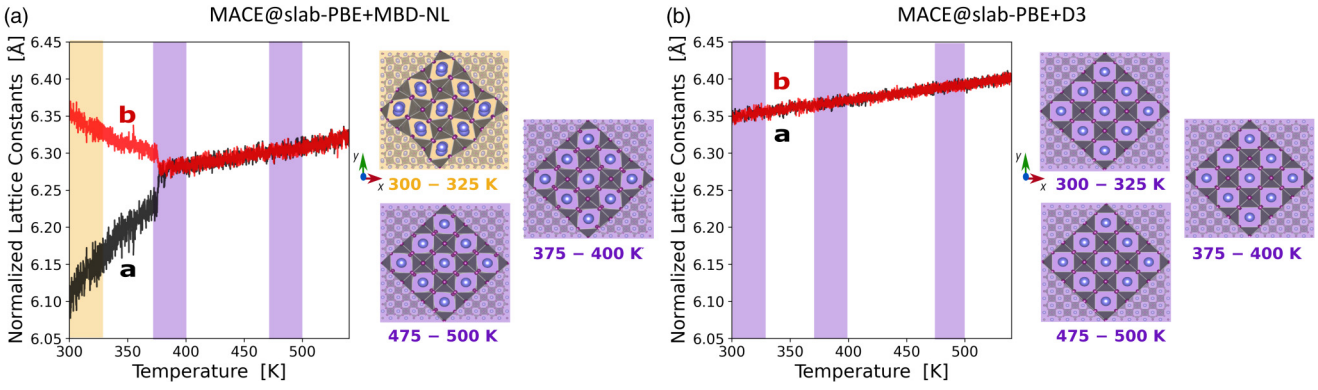


FIG. 3. Temperature evolution of normalized lattice parameters a (black) and b (red) obtained from an NPT MD heating trajectory with heating rate 50 K ns^{-1} of slab $CsPbI_3$ $(8\sqrt{2} \times 8\sqrt{2})$-5L using (a) the MACE@slab-PBE+MBD-NL and (b) MACE@slab-PBE+D3 MLFFs. The lattice parameters are time-averaged over 2 ps windows (corresponding to $\Delta T = 0.1$ K). The right panels show representative time-averaged structures for selected windows (300–325 K, 375–400 K, and 475–500 K), with background colors matching the shaded regions in the lattice-constant plot; insets showing magnified regions are included to better visualize the octahedral tilting. Shaded bands indicate phase regions: orange, γ phase; purple, α phase.

trained on the same cubic-like structures as the MACE@slab-PBE+D3, it nevertheless stabilizes orthorhombic-like order at room temperature, highlighting that an accurate treatment of dispersion can meaningfully reshape the finite-temperature free-energy landscape governing octahedral tilts.

B. Slab-to-bulk transfer learning

Given that slab structures include a wider range of local environments, arising from the coexistence of surface, subsurface, and interior layers that retain the bulk coordination environment, we next examine whether the MACE@slab-PBE+MBD-NL transfers reliably to finite-temperature MD simulations of bulk $CsPbI_3$. First, we verified long-time dynamical stability by running 10 ns of NPT MD at 300 K for supercells up to $8 \times 8 \times 8$: all trajectories remained numerically stable and retained the expected γ-phase tilt statistics (Fig. S6) [59]. Such long-timescale, large-supercell stability checks are important because MLFFs can exhibit instabilities or nonphysical behavior during extended MD simulations [80,81].

To probe the temperature-driven structural phase transitions of bulk $CsPbI_3$, we performed NPT MD using the MACE@slab-PBE+MBD-NL model on a $5 \times 5 \times 5$ γ-phase $CsPbI_3$, heating from 300 to 700 K with a rate of $50\,\mathrm{K\,ns^{-1}}$. Since experimental transition temperatures are commonly inferred from crossings of the lattice parameters (e.g., a and b for the β phase and a, b, and c for the α phase), Fig. 4 summarizes the evolution of the normalized lattice parameters with temperature, together with time-averaged structures for selected temperature windows. Because lattice parameters during MD exhibit sizable thermostat- and barostat-driven fluctuations, the lattice parameters in Fig. 4 are time-averaged over a 2 ps interval. As shown in Fig. S7 in the SM [59], using a shorter averaging window produces noticeably noisier plots and slightly shifts the crossing points between the lattice constants. Lattice parameters alone may be insufficient to assign phases unambiguously, particularly in soft, anharmonic materials such as $CsPbI_3$, where transitions can be gradual [82]. As discussed in Sec. I, octahedral tilt angles and their temperature-dependent distributions and correlations, obtained from the MD simulations, are now routinely used to distinguish different polymorphs [41–44]. We therefore complement the time-averaged structures with the tilt-angle distributions (see Fig. 4) using the (φ, θ, ψ) proposed by Liang *et al.* (see Fig. 4 in Ref. [44]) to assign the phases: the γ phase by bimodal distributions for all three angles (φ, θ, ψ), the β phase by a bimodal ψ distribution together with unimodal φ and θ peaks centered at 0°, and for the α by unimodal distributions centered at 0° for all three angles (see Sec. S4 in the SM for details) [59]. Noteworthy, the (φ, θ, ψ) angles can be associated with the octahedral tilting patterns expressed by the Glazer notation; the results in Fig. S8 in the SM [59] confirm the $a^-a^-c^+$ and $a^0a^0c^+$ octahedral tilting for the orthorhombic and tetragonal phases.

Using these metrics, the MACE@slab-PBE+MBD-NL model reproduces the expected phase sequence of bulk $CsPbI_3$: γ phase at 300 K, a $\gamma \rightarrow \beta$ transition around 525 K, and a $\beta \rightarrow \alpha$ transition around 600 K, where all three rotation-angle distributions become unimodal and

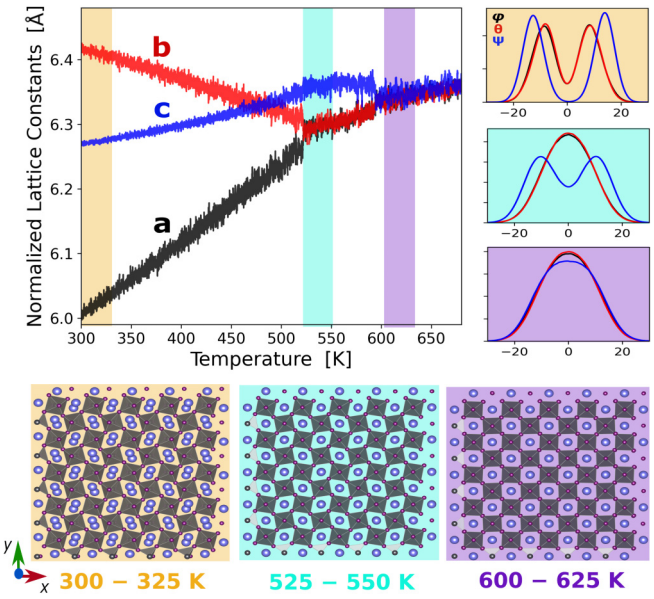


FIG. 4. Temperature evolution of normalized lattice parameters a (black), b (red), and c (blue) obtained from an NPT MD heating trajectory with heating rate $50\,\mathrm{K\,ns^{-1}}$ of bulk $CsPbI_3$ ($5 \times 5 \times 5$ supercell) using the MACE@slab-PBE+MBD-NL MLFF. The lattice parameters are time-averaged over 2 ps windows (corresponding to $\Delta T = 0.1$ K). Shaded bands indicate the temperature windows used for structural analysis. Right panels report the probability densities of PbI_6 octahedral rotation angles (in degrees) about the x, y, and z axes (labeled φ, θ, and ψ), with background colors matching the shaded regions in the lattice-constant plot. Bottom panels show representative time-averaged structures for selected windows (300–325 K, 525–550 K, and 600–625 K). Shaded bands indicate phase regions: orange, γ phase; navy, β phase; purple, α phase.

centered near 0°. The corresponding time-averaged structures in the selected temperature windows are consistent with these assignments. Quantitatively, however, the lattice parameters deviate from the experiment of Marronnier *et al.* [14], most notably for the in-plane parameters: a is typically slightly underestimated (less than 0.1 Å), while b is slightly overestimated (≈0.1 Å). These lattice-constant deviations result in the shifts in the apparent transition temperatures: $\gamma \rightarrow \beta$ is shifted by ≈65 K relative to experiment ($T^{\mathrm{exp}}_{\gamma\rightarrow\beta} \approx 460$ K), and $\beta \rightarrow \alpha$ is higher by ≈40 K ($T^{\mathrm{exp}}_{\beta\rightarrow\alpha} \approx 560$ K). Note that such small deviations in lattice parameters (less than 2%) and energy shifts of approximately 0.1 kcal/mol lie within the typical error range of the most advanced MLFFs. Because these parameters can also be sensitive to finite-size effects and the heating/cooling rate, we repeated the NPT ramps with a larger supercell and slower heating rate (Fig. S9) [59]. Within the tested ranges (see Sec. S3 of the SM for more details) [59], neither factor measurably changes the transition signatures, indicating that our MD protocol is unlikely to be the primary source of the remaining offsets.

We emphasize that most previous MLFFs developed for HPs to reproduce bulk phase behavior were trained on bulk reference data spanning multiple perovskite polymorphs [37,39,48,73], whereas our model, MACE@slab-PBE+MBD-NL, was trained exclusively on α-phase slab

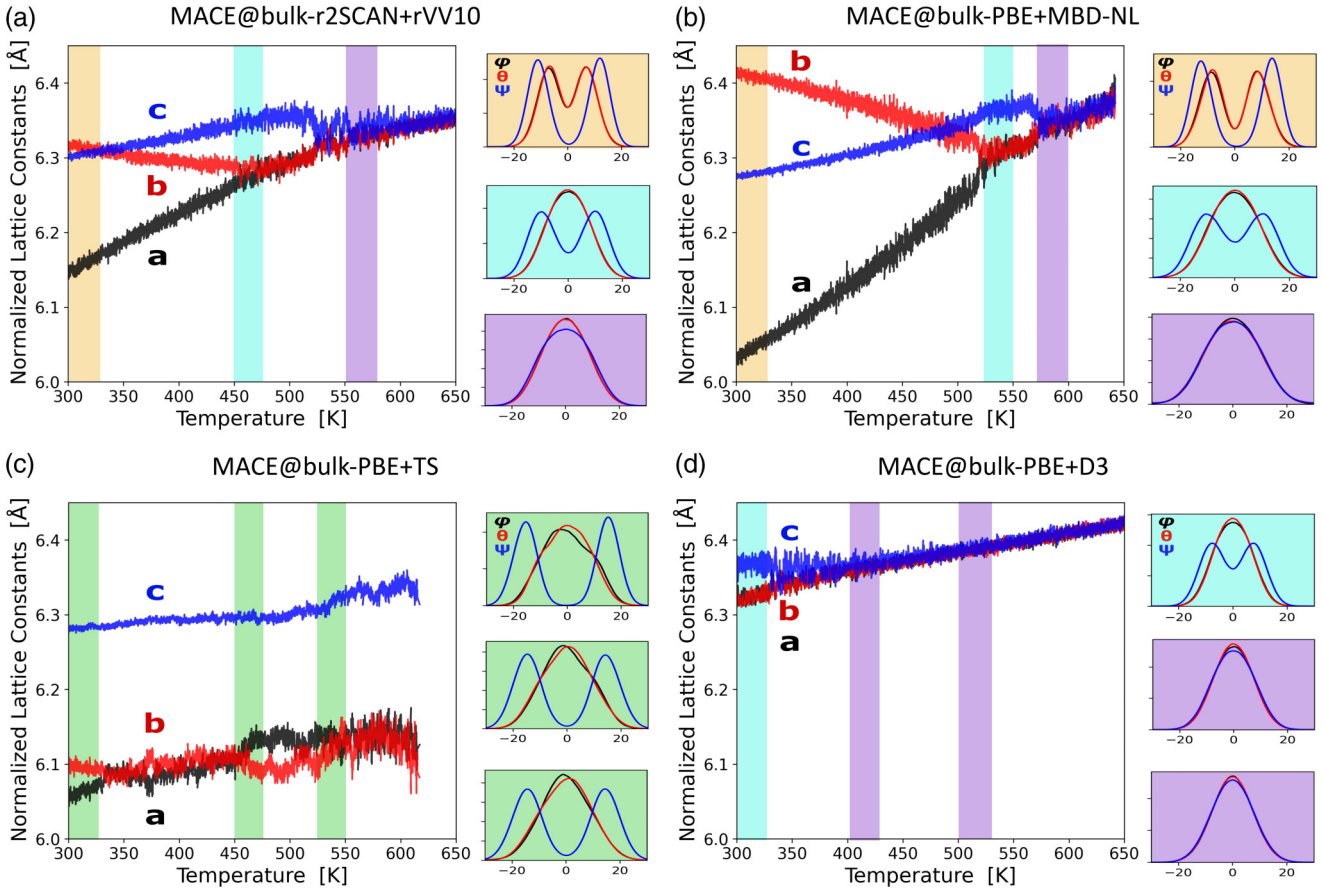


FIG. 5. Temperature dependence of lattice parameters a (black), b (red), and c (blue) obtained from MD upon heating with 50 K ns^{-1} rate for $CsPbI_3$ $5 \times 5 \times 5$ bulk with (a) MACE@bulk-r^{2}SCAN+rVV10, (b) MACE@bulk-PBE+MBD-NL, (c) MACE@bulk-PBE+TS, and (d) MACE@bulk-PBE+D3. The lattice parameters were averaged over 2 ps windows (corresponding to a $\Delta T = 0.1$ K). Right panels report the probability densities of PbI_6 octahedral rotation angles (in degrees) about the x, y, and z axes (labeled φ, θ, and ψ), with background colors matching the shaded regions in the lattice-constant plot. Shaded bands indicate phase regions: orange, γ phase; navy, β phase; purple, α phase; green, mixed γ/β phase.

configurations and nevertheless reproduces bulk phase transitions in agreement with experiment.

C. Reference-method dependence: Bulk-trained models

Beyond the MD settings, the choice of reference method used to generate the training data is another user-controlled factor and a potential source of disagreements. Previous work (Ref. [39]) showed that the choice of exchange-correlation functional can strongly influence the lattice parameters and, consequently, the phase-transition temperatures. Here, we focus on the role of the vdW dispersion by comparing many-body MBD-NL with the pairwise TS and D3 corrections, as well as with the nonlocal r^{2}SCAN+rVV10 functional. To demonstrate that the previously observed differences in lattice parameters originate from the level of theory rather than from the slab structures, we perform a benchmark on bulk datasets (see Sec. II for more details). To this end, we trained four MACE models on identical configurations from a bulk dataset (obtained with r^{2}SCAN+rVV10) with energies and forces recomputed with different levels of theory: r^{2}SCAN+rVV10, PBE+MBD-NL, PBE+TS, and PBE+D3. The resulting models were then used to perform the same NPT MD heating protocol (300–700 K at 50 K ns^{-1}) on a $5 \times 5 \times 5$ supercell of the orthorhombic phase of bulk $CsPbI_3$, and the temperature evolution of lattice parameters, octahedral tilts, and time-averaged structures were analyzed as before. Figure 5 summarizes temperature-driven lattice parameters and octahedral tilts for these four models. The time-averaged structures for these models are available in the SM (see Fig. S10) [59].

One can see that both the MACE@bulk-r^{2}SCAN+rVV10 and MACE@bulk-PBE+MBD-NL preserve the correct $\gamma \rightarrow \beta \rightarrow \alpha$ sequence. The two models differ only slightly in the in-plane lattice parameters a and b (≈ 0.1 Å), yet these small differences shift the phase-transition temperatures. The MACE@bulk-r^{2}SCAN+rVV10 [Fig. 5(a)] yields tilt-angle distributions consistent with the β phase over 450–500 K, followed by a transition to the cubic-like α phase around 550–575 K, in close agreement with experimental transition temperatures reported by Marronnier *et al.* [14]. The MACE@bulk-PBE+MBD-NL [Fig. 5(b)] reproduces the same qualitative behavior as the MACE@bulk-r^{2}SCAN+rVV10, but shifts the $\gamma \rightarrow \beta$ transition to 525 K

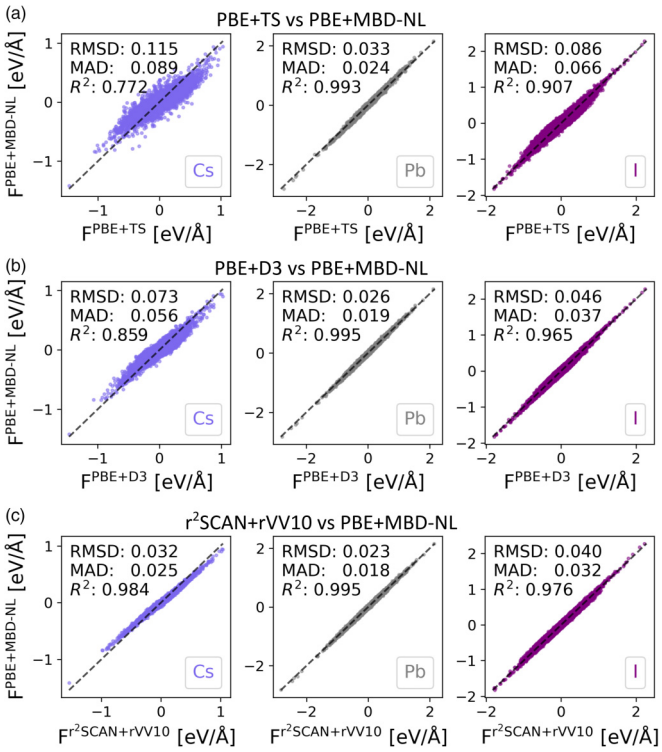


FIG. 6. Comparison of force components on the identical structures from the bulk dataset (392 structures) computed with different reference methods: (a) PBE+TS vs PBE+MBD-NL, (b) PBE+D3 vs. PBE+MBD-NL, and (c) r^{2}SCAN+rVV10 vs. PBE+MBD-NL. RMSD and MAD are reported in eV Å^{-1}. The dashed line denotes perfect agreement. R^2 is the squared Pearson correlation coefficient.

and $\beta \rightarrow \alpha$ to 575 K. In contrast, MLFFs trained with pairwise dispersion methods exhibit significantly different behavior. The MACE@bulk-PBE+TS [Fig. 5(c)] remains in a mixed/intermediate state between γ and β across the 300–650 K range, in clear disagreement with experiment. The MACE@bulk-PBE+D3 [Fig. 5(d)], on the other hand, overstabilizes the β phase at room temperature and predicts a $\beta \rightarrow \alpha$ transition already around 400 K, more than 150 K below the experimental value. These results therefore show that more advanced vdW dispersion methods are essential for correctly capturing the phase transitions of $CsPbI_3$.

To identify the atomic-level origin of the reference-dependent differences in phase transitions, we compare force components on each atom across methods in Fig. 6. The near-zero mean signed deviation observed across all cases indicates little systematic bias, as force overestimations and underestimations largely cancel each other out across the different methods. As in the slab analysis, the strongest dispersion sensitivity is observed for Cs atoms. PBE+TS forces strongly deviate from PBE+MBD-NL ones [Fig. 6(a)], with an RMSD of 0.115 eV Å^{-1} and a relatively low correlation ($R^2 = 0.77$). Similarly, PBE+D3 forces also exhibit large deviations from PBE+MBD-NL forces, although with slightly improved R^2 and lower RMSD/MAD values. The isolated dispersion force contributions in Fig. S11 [59] confirm that ΔF^{D3} and ΔF^{TS} are predominantly anticorrelated with ΔF^{MBD-NL}, with TS contributions exhibiting the larger deviation from MBD-NL. By contrast, r^{2}SCAN+rVV10 and PBE+MBD-NL forces correlate closely [Fig. 6(c)] with $R^2 = 0.98$ and significantly smaller RMSD/MAD than those of PBE+TS and PBE+D3. Nevertheless, even the modest Cs-force discrepancies between r^{2}SCAN+rVV10 and PBE+MBD-NL appear sufficient for measurable changes in tilt dynamics, contributing to the observed differences in the a and b lattice parameters and the phase-transition temperatures.

Consistent with the slab analysis, Pb and I atomic forces show higher R^2 and lower RMSDs and MAD than Cs atomic forces. However, among the considered methods, PBE+TS yields the largest deviations for the I atomic forces relative to PBE+MBD-NL. This might correlate with the tendency of the MACE@bulk-PBE+TS MLFF to stabilize the mixed phase of $CsPbI_3$.

To examine how the dispersion method influences octahedral tilting, we selected 2500 structures of $5 \times 5 \times 5$ CsPbI3 bulk from the 300–325 K temperature window of the slow heating trajectory obtained with the MACE@slab-PBE+MBD-NL model. For those structures, energies and forces were computed with the MACE@bulk-PBE, MACE@bulk-PBE+MBD-NL, and MACE@bulk-PBE+D3 models. The PBE model was included to isolate the dispersion-induced contributions, which are otherwise small compared to the full forces. As shown in Fig. S12 in the SM [59], the MLFF dispersion-induced force corrections,

$$\Delta F^{D3}_{MLFF} = F_{MACE@bulk-PBE+D3} - F_{MACE@bulk-PBE}, \quad (1)$$

and

$$\Delta F^{MBD-NL}_{MLFF} = F_{MACE@bulk-PBE+MBD-NL} - F_{MACE@bulk-PBE}, \quad (2)$$

reproduce the same qualitative trends as the reference bulk calculations. To connect these force differences with octahedral rotations, we computed the local torque acting on a PbI_6 octahedron,

$$\tau = \sum_i r_{Pb \rightarrow I} \times F_{I_i}. \quad (3)$$

This quantity represents the instantaneous rotational driving force acting on the iodine cage about the central Pb atom. The torque distributions obtained from the full forces for two corner-sharing octahedra overlap substantially because the PBE contribution dominates (see Fig. S13 in the SM) [59]. However, when the dispersion-induced force contributions are considered separately, the resulting torque distributions show that MBD-NL and D3 generate oppositely directed rotational contributions along all three Cartesian axes (Fig. 7). In addition, the signs of these torque corrections reverse between neighboring connected octahedra, consistent with the opposite signs of their local tilt angles (Fig. S14 in the SM) [59]. These results show that the treatment of dispersion modifies the iodine forces and thereby changes the local rotational driving-force field of the corner-sharing PbI_6 framework.

Because the forces acting on Cs are particularly sensitive to the dispersion treatment, we further examined the local coupling between Cs displacement and the rotational response of neighboring PbI_6 octahedra. One Cs atom and its eight

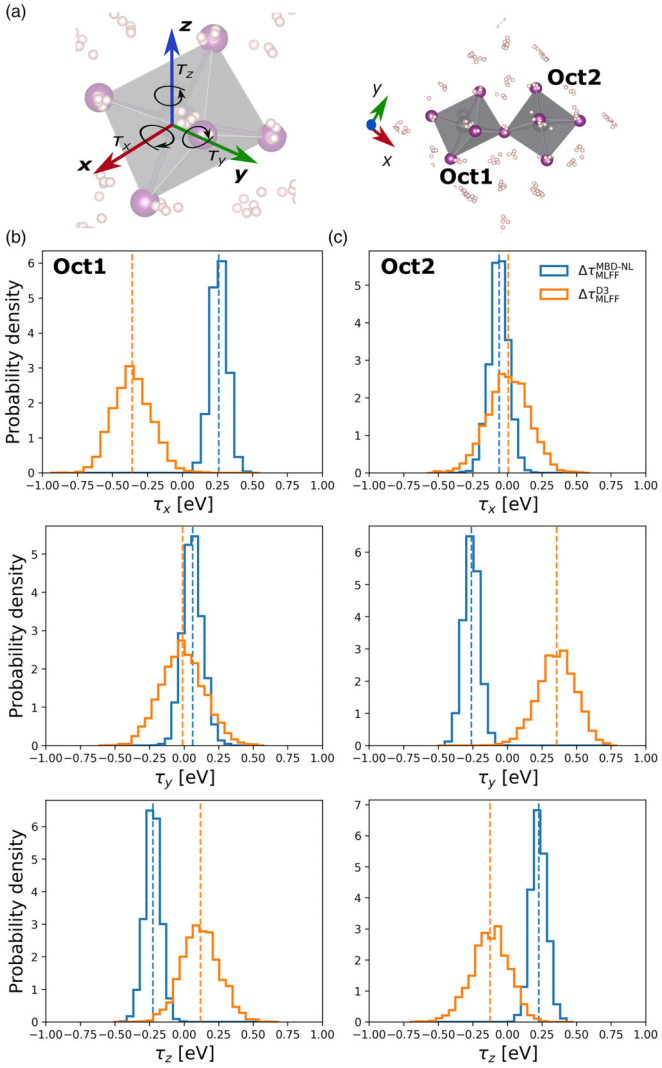


FIG. 7. Local PbI_6 torque distributions computed from dispersion-induced force corrections, $\Delta F^{\mathrm{D3}}_{\mathrm{MLFF}}$ and $\Delta F^{\mathrm{MBD-NL}}_{\mathrm{MLFF}}$. (a) Schematic illustration of torque components and locations of two selected corner-sharing octahedra, **Oct1** and **Oct2**. All the atoms except those selected are shown as small pale spheres for visualization only. (b),(c) Probability density distributions of τ_x, τ_y, and τ_z for **Oct1** and **Oct2**, respectively, calculated using dispersion-induced force corrections $\Delta F^{\mathrm{D3}}_{\mathrm{MLFF}}$ [see Eq. (1)] and $\Delta F^{\mathrm{MBD-NL}}_{\mathrm{MLFF}}$ [see Eq. (2)] from the same set of 2500 structures. Dashed vertical lines indicate the mean values of each distribution. See Fig. S13 in the SM for more details [59].

neighboring octahedra were selected from a representative structure used in the torque analysis. The Cs atom was displaced by ±0.01 Å along each Cartesian direction, while all other atoms were kept fixed, and the forces were evaluated using the MACE@bulk-PBE, MACE@bulk-PBE+D3, and MACE@bulk-PBE+MBD-NL models. For each octahedron, we calculated the displacement–torque response matrix

$$K_{ij} = \frac{\tau_i(+u_j) - \tau_i(-u_j)}{2u_j}, \tag{4}$$

where u_j is the Cs displacement along $j \in \{x, y, z\}$, and τ_i is the corresponding octahedral torque component $i \in \{x, y, z\}$.

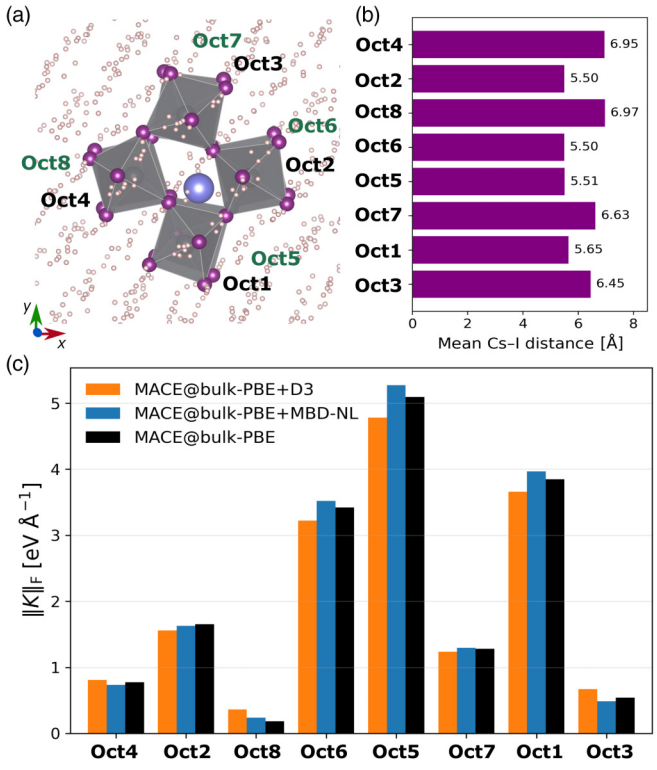


FIG. 8. Local coupling between Cs displacement and octahedra torque. (a) Selected Cs atom and its eight neighboring PbI_6 octahedra. All the atoms except those selected are shown as small pale spheres for visualization only. (b) Mean Cs-I distance for each neighboring octahedron in the structure without Cs displacement. (c) Frobenius norm, $\|K\|_{\mathrm{F}}$, of the Cs displacement and octahedral torque response matrix calculated using the MACE@bulk-PBE, MACE@bulk-PBE+MBD-NL, and MACE@bulk-PBE+D3 forces.

The diagonal elements describe same-axis displacement-torque coupling, whereas offdiagonal elements describe crossaxis coupling. The overall coupling strength was characterized using the Frobenius norm,

$$\|K\|_{\mathrm{F}} = \sqrt{\sum_{ij} K_{ij}^2}. \tag{5}$$

In Fig. 8, we show the selected Cs atom and its eight neighboring PbI_6 octahedra, the mean Cs-I distance for each octahedron calculated for the structure without displacement, and displacement-torque coupling strength $\|K\|_{\mathrm{F}}$, calculated using the MACE@bulk-PBE, MACE@bulk-PBE+MBD-NL, and MACE@bulk-PBE+D3 forces. The strongest responses are obtained for **Oct1**, **Oct5**, and **Oct6**, which also have the shortest mean Cs-I distances among the selected octahedra. For these three octahedra, the MACE@bulk-PBE+MBD-NL consistently predicts a larger displacement-torque response than the MACE@bulk-PBE+D3. The response matrices are highly anisotropic and contain substantial crossaxis couplings (see Fig. S15) [59]. This local finite-displacement analysis shows that changes in Cs position modify the rotational driving forces acting on the surrounding PbI_6 octahedra, providing a local coupling mechanism between A-site dynamics and the inorganic framework.

Taken together, the results for the slab and bulk clearly indicate that the choice of the dispersion model primarily affects Cs forces, accompanied by distinct octahedral-tilt dynamics and the resulting phase behavior. Both MBD-NL and rVV10 preserve the experimentally observed $\gamma \rightarrow \beta \rightarrow \alpha$ phase transitions over the 300–700 K range, TS stabilizes the wrong mixed phase between γ and β, and D3 favors the α phase at unrealistically low temperatures.

One should note that although the MACE@bulk-PBE+MBD-NL [Fig. 5(b)] recovers the correct $\gamma \rightarrow \beta \rightarrow \alpha$ phase-transition sequence and reproduces the same qualitative behavior as the MACE@slab-PBE+MBD-NL model (Fig. 4), inspection of time-averaged snapshots (Fig. S10b) [59] at 525–550 K and 575–600 K reveals nonphysical MLFF behavior. This includes atomic displacements where atoms move too close to one another or deviate excessively from their expected positions, disrupting the anticipated symmetry and bonding patterns and reflecting the limited diversity of the bulk training set. To investigate this further, we trained an additional MLFF model on a larger bulk dataset containing 3916 structures, with energies and forces computed at the same PBE+MBD-NL level of theory (see Sec. S7 of the SM for more details) [59]. The results in Fig. S16 [59] show that the stability of dynamics improves with increasing training set size. Notably, the MACE@slab-PBE+MBD-NL yields stable dynamics across the entire temperature range studied for both bulk and slab materials, with no observed structural disruption. This suggests that large slab supercells, containing 864 atoms, together with the simultaneous presence of surface, subsurface, and interior layers, enrich the training dataset with a broad range of local atomic environments despite the relatively small number of configurations (600 structures).

IV. CONCLUSIONS

We presented a practical strategy for developing transferable and accurate MLFFs for modeling dynamic behavior of $CsPbI_3$ across slab interfaces and bulk. Training on the large slab supercell (864 atoms) produced MLFFs that yield stable MD for both slabs and bulk over multi-nanosecond timescales. Notably, a model trained only on α-phase slab data with energies and forces calculated at the PBE+MBD-NL level of theory reproduces the correct phase-transition behavior upon heating – a level of transferability that is typically achieved using multi-phase bulk training sets. This suggests an efficient protocol: training on surface-enriched single-phase slab data to obtain a robust MLFF for modeling both bulk materials and surfaces.

By systematically comparing reference methods, we found that the treatment of vdW dispersion is the primary determinant of the predicted finite-temperature behavior of $CsPbI_3$. Pairwise dispersion schemes (TS/D3) can qualitatively distort the tilt-driven phase evolution, whereas beyond-pairwise dispersion (MBD-NL) and nonlocal density-based (rVV10) approaches recover the expected $\gamma \rightarrow \beta \rightarrow \alpha$ sequence. Analysis of forces highlights that the choice of dispersion method has a significant impact on Cs forces, consistent with its large and polarizable nature. The torque analysis further demonstrates that D3 and MBD-NL produce different, often oppositely directed, contributions to the rotational driving forces acting on the PbI_6 framework. A local finite-displacement analysis also shows that Cs motion is coupled to the torque on neighboring octahedra, with the strongest response obtained for the closest ones. These results indicate that dispersion-dependent changes in both Cs and I forces modify octahedral rotational dynamics and thereby contribute to the distinct phase-behavior predicted by the different dispersion treatments.

Taken together, our findings establish that slab datasets computed with beyond-pairwise vdW dispersion capture the essential physics required to describe both bulk materials and interfaces, opening the way to robust simulations of complete solar-cell devices under realistic operating conditions.

ACKNOWLEDGMENTS

We are thankful to Dr. Paramvir Ahlawat for providing us with a dataset of bulk structures. The authors also thank Dr. Simon Thébaud and Dr. Evgeny Moerman for insightful discussions. This research was funded in whole or in part by the Luxembourg National Research Fund (FNR), grant Reference No. INTER/DFG/25/18971124/CLARINET. A.C. acknowledges support from the European Research Council (ERC) under the European Union's Horizon Europe programme through the FITMOL project (Grant Agreement No. 101054629). The simulations were performed on the Luxembourg national supercomputer MeluXina. The authors gratefully acknowledge the LuxProvide teams for their expert support. Computational resources were provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the F.R.S.-FNRS under Grant No. 2.5020.11. The present research also benefited from access to Lucia, the Tier-1 supercomputer of the Walloon Region, an infrastructure funded by the Walloon Region under Grant Agreement No. 1910247. C.Q. is an FNRS research associate, and D.B. is an FNRS research director.

We acknowledge the use of generative AI (ChatGPT) for language editing of the manuscript and for assistance in editing and debugging Python scripts used for data analysis.

The authors have no conflicts to disclose.

DATA AVAILABILITY

The data supporting the findings of this study are available on Zenodo at [84]. Additional data are available from the corresponding authors upon reasonable request.

[1] A. Kojima, K. Teshima, Y. Shirai, and T. Miyasaka, Organometal halide perovskites as visible-light sensitizers for photovoltaic cells, J. Am. Chem. Soc. **131**, 6050 (2009).

[2] M. A. Green, A. Ho-Baillie, and H. J. Snaith, The emergence of perovskite solar cells, Nat. Photon. **8**, 506 (2014).

[3] W.-J. Yin, T. Shi, and Y. Yan, Unusual defect physics in $CH_3NH_3PbI_3$ perovskite solar cell absorber, Appl. Phys. Lett. **104**, 063903 (2014).

[4] A. Amat, E. Mosconi, E. Ronca, C. Quarti, P. Umari, M. K. Nazeeruddin, M. Grätzel, and F. De Angelis, Cation-induced band-gap tuning in organohalide perovskites: Interplay of spin–orbit coupling and octahedra tilting, Nano Lett. **14**, 3608 (2014).

[5] T. M. Brenner, D. A. Egger, L. Kronik, G. Hodes, and D. Cahen, Hybrid organic—inorganic perovskites: Low-cost semiconductors with intriguing charge-transport properties, Nat. Rev. Mater. **1**, 15007 (2016).

[6] R. E. Beal, D. J. Slotcavage, T. Leijtens, A. R. Bowring, R. A. Belisle, W. H. Nguyen, G. F. Burkhard, E. T. Hoke, and M. D. McGehee, Cesium lead halide perovskites with improved stability for tandem solar cells, J. Phys. Chem. Lett. **7**, 746 (2016).

[7] Y. Rong, Y. Hu, A. Mei, H. Tan, M. I. Saidaminov, S. I. Seok, M. D. McGehee, E. H. Sargent, and H. Han, Challenges for commercializing perovskite solar cells, Science **361**, eaat8235 (2018).

[8] L. Pedesseau, D. Sapori, B. Traore, R. Robles, H.-H. Fang, M. A. Loi, H. Tsai, W. Nie, J.-C. Blancon, A. Neukirch, S. Tretiak, A. D. Mohite, C. Katan, J. Even, and M. Kepenekian, Advances and promises of layered halide hybrid perovskite semiconductors, ACS Nano **10**, 9776 (2016).

[9] L. Yuan, Q. Xue, F. Wang, N. Li, G. I. N. Waterhouse, C. J. Brabec, F. Gao, and K. Yan, Perovskite solar cells and light emitting diodes: Materials chemistry, device physics and relationship, Chem. Rev. **125**, 5057 (2025).

[10] S. Liu, J. Li, W. Xiao, R. Chen, Z. Sun, Y. Zhang, X. Lei, S. Hu, M. Kober-Czerny, J. Wang, F. Ren, Q. Zhou, H. Raza, Y. Gao, Y. Ji, S. Li, H. Li, L. Qiu, W. Huang, Y. Zhao, *et al.*, Buried interface molecular hybrid for inverted perovskite solar cells, Nature (London) **632**, 536 (2024).

[11] H. Chen, C. Liu, J. Xu, A. Maxwell, W. Zhou, Y. Yang, Q. Zhou, A. S. R. Bati, H. Wan, Z. Wang, L. Zeng, J. Wang, P. Serles, Y. Liu, S. Teale, Y. Liu, M. I. Saidaminov, M. Li, N. Rolston, S. Hoogland, *et al.*, Improved charge extraction in inverted perovskite solar cells with dual-site-binding ligands, Science **384**, 189 (2024).

[12] J. Wang, Y. Che, Y. Duan, Z. Liu, S. Yang, D. Xu, Z. Fang, X. Lei, Y. Li, and S. F. Liu, 21.15%-efficiency and stable γ-$CsPbI_3$ perovskite solar cells enabled by an acyloin ligand, Adv. Mater. **35**, 2210223 (2023).

[13] J. Duan, H. Xu, W. E. I. Sha, Y. Zhao, Y. Wang, X. Yang, and Q. Tang, Inorganic perovskite solar cells: An emerging member of the photovoltaic community, J. Mater. Chem. A **7**, 21036 (2019).

[14] A. Marronnier, G. Roma, S. Boyer-Richard, L. Pedesseau, J.-M. Jancu, Y. Bonnassieux, C. Katan, C. C. Stoumpos, M. G. Kanatzidis, and J. Even, Anharmonicity and disorder in the black phases of cesium lead iodide used for stable inorganic perovskite solar cells, ACS Nano **12**, 3477 (2018).

[15] R. J. Sutton, M. R. Filip, A. A. Haghighirad, N. Sakai, B. Wenger, F. Giustino, and H. J. Snaith, Cubic or orthorhombic? Revealing the crystal structure of metastable black-phase $CsPbI_3$ by theory and experiment, ACS Energy Lett. **3**, 1787 (2018).

[16] W. Chu, W. A. Saidi, J. Zhao, and O. V. Prezhdo, Soft lattice and defect covalency rationalize tolerance of β-$CsPbI_3$ perovskite solar cells to native defects, Angew. Chem. Int. Ed. **59**, 6435 (2020).

[17] H. Lyu, H. Su, and Z. Lin, Two-stage dynamic transformation from δ- to α-$CsPbI_3$, J. Phys. Chem. Lett. **15**, 2228 (2024).

[18] C. C. Stoumpos, C. D. Malliakas, and M. G. Kanatzidis, Semiconducting tin and lead iodide perovskites with organic cations: Phase transitions, high mobilities, and near-infrared photoluminescent properties, Inorg. Chem. **52**, 9019 (2013).

[19] C. C. Stoumpos and M. G. Kanatzidis, The renaissance of halide perovskites and their evolution as emerging semiconductors, Acc. Chem. Res. **48**, 2791 (2015).

[20] A. Marronnier, H. Lee, B. Geffroy, J. Even, Y. Bonnassieux, and G. Roma, Structural instabilities related to highly anharmonic phonons in halide perovskites, J. Phys. Chem. Lett. **8**, 2659 (2017).

[21] F. Ke, C. Wang, C. Jia, N. R. Wolf, J. Yan, S. Niu, T. P. Devereaux, H. I. Karunadasa, W. L. Mao, and Y. Lin, Preserving a robust $CsPbI_3$ perovskite phase via pressure-directed octahedral tilt, Nat. Commun. **12**, 461 (2021).

[22] J.-S. Park, J. Calbo, Y.-K. Jung, L. D. Whalley, and A. Walsh, Accumulation of deep traps at grain boundaries in halide perovskites, ACS Energy Lett. **4**, 1321 (2019).

[23] J. Haruyama, K. Sodeyama, L. Han, and Y. Tateyama, Surface properties of $CH_3NH_3PbI_3$ for perovskite solar cells, Acc. Chem. Res. **49**, 554 (2016).

[24] C. Li, S. Tscheuschner, F. Paulus, P. E. Hopkinson, J. Kießling, A. Köhler, Y. Vaynzof, and S. Huettner, Iodine migration and its effect on hysteresis in perovskite solar cells, Adv. Mater. **28**, 2446 (2016).

[25] H. Uratani and K. Yamashita, Charge carrier trapping at surface defects of perovskite solar cell absorbers: A first-principles study, J. Phys. Chem. Lett. **8**, 742 (2017).

[26] X. Gu, W. Xiang, Q. Tian, and S. F. Liu, Rational surface-defect control via designed passivation for high-efficiency inorganic perovskite solar cells, Angew. Chem. Int. Ed. **60**, 23164 (2021).

[27] M. Pols, J. M. Vicent-Luna, I. Filot, A. C. T. van Duin, and S. Tao, Atomistic insights into the degradation of inorganic halide perovskite $CsPbI_3$: A reactive force field molecular dynamics study, J. Phys. Chem. Lett. **12**, 5519 (2021).

[28] M. Pols, T. Hilpert, I. A. Filot, A. C. van Duin, S. Calero, and S. Tao, What happens at surfaces and grain boundaries of halide perovskites: Insights from reactive molecular dynamics simulations of $CsPbI_3$, ACS Appl. Mater. Interfaces **14**, 40841 (2022).

[29] J. Even, L. Pedesseau, J.-M. Jancu, and C. Katan, Importance of spin–orbit coupling in hybrid organic/inorganic perovskites for photovoltaic applications, J. Phys. Chem. Lett. **4**, 2999 (2013).

[30] J. M. Frost, K. T. Butler, F. Brivio, C. H. Hendon, M. van Schilfgaarde, and A. Walsh, Atomistic origins of high-performance in hybrid halide perovskite solar cells, Nano Lett. **14**, 2584 (2014).

[31] S. Meloni, G. Palermo, N. Ashari-Astani, M. Grätzel, and U. Rothlisberger, Valence and conduction band tuning in halide perovskites for solar cell applications, J. Mater. Chem. A **4**, 15997 (2016).

[32] J. Wiktor, U. Rothlisberger, and A. Pasquarello, Predictive determination of band gaps of inorganic halide perovskites, J. Phys. Chem. Lett. **8**, 5507 (2017).

[33] W. Li, A. S. Vasenko, J. Tang, and O. V. Prezhdo, Anharmonicity extends carrier lifetimes in lead halide perovskites at elevated temperatures, J. Phys. Chem. Lett. **10**, 6219 (2019).

[34] Y.-K. Jung, M. Dubajic, and S. D. Stranks, Electronic consequences of the formation of local tilting in halide perovskites, J. Phys. Chem. Lett. **16**, 5862 (2025).

[35] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. **140**, A1133 (1965).

[36] L. Hedin, New method for calculating the one-particle Green's function with application to the electron-gas problem, Phys. Rev. **139**, A796 (1965).

[37] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, and M. Bokdam, Phase transitions of hybrid perovskites simulated by machine-learning force fields trained on the fly with Bayesian inference, Phys. Rev. Lett. **122**, 225701 (2019).

[38] T. Braeckevelt, R. Goeminne, S. Vandenhaute, S. Borgmans, T. Verstraelen, J. A. Steele, M. B. J. Roeffaers, J. Hofkens, S. M. J. Rogge, and V. Van Speybroeck, Accurately determining the phase transition temperature of $CsPbI_3$ via random-phase approximation calculations and phase-transferable machine learning potentials, Chem. Mater. **34**, 8561 (2022).

[39] E. Fransson, J. Wiktor, and P. Erhart, Phase transitions in inorganic halide perovskites from machine-learned potentials, J. Phys. Chem. C **127**, 13773 (2023).

[40] M. A. Carignano, S. A. Aravindh, I. S. Roqan, J. Even, and C. Katan, Critical fluctuations and anharmonicity in lead iodide perovskites from molecular dynamics supercell simulations, J. Phys. Chem. C **121**, 20729 (2017).

[41] C. Quarti, E. Mosconi, P. Umari, and F. De Angelis, Chlorine incorporation in the $CH_3NH_3PbI_3$ perovskite: Small concentration, big effect, Inorg. Chem. **56**, 74 (2017).

[42] Y. Dou, F. Cao, T. Dudka, Y. Li, S. Wang, C. Zhang, Y. Gao, X. Yang, and A. L. Rogach, Lattice distortion in mixed-anion lead halide perovskite nanorods leads to their high fluorescence anisotropy, ACS Mater. Lett. **2**, 814 (2020).

[43] J. Wiktor, E. Fransson, D. Kubicki, and P. Erhart, Quantifying dynamic tilting in halide perovskites: Chemical trends and local correlations, Chem. Mater. **35**, 6737 (2023).

[44] X. Liang, J. Klarbring, W. J. Baldwin, Z. Li, G. Csányi, and A. Walsh, Structural dynamics descriptors for metal halide perovskites, J. Phys. Chem. C **127**, 19141 (2023).

[45] T. Baikie, Y. Fang, J. M. Kadro, M. Schreyer, F. Wei, S. G. Mhaisalkar, M. Graetzel, and T. J. White, Synthesis and crystal chemistry of the hybrid perovskite $CH_3NH_3PbI_3$ for solid-state sensitised solar cell applications, J. Mater. Chem. A **1**, 5628 (2013).

[46] C. Quarti, E. Mosconi, and F. De Angelis, Structural and electronic properties of organo-halide hybrid perovskites from *ab initio* molecular dynamics, Phys. Chem. Chem. Phys. **17**, 9394 (2015).

[47] O. Yaffe, Y. Guo, L. Z. Tan, D. A. Egger, T. Hull, C. C. Stoumpos, F. Zheng, T. F. Heinz, L. Kronik, M. G. Kanatzidis, J. S. Owen, A. M. Rappe, M. A. Pimenta, and L. E. Brus, Local polar fluctuations in lead halide perovskite crystals, Phys. Rev. Lett. **118**, 136001 (2017).

[48] W. J. Baldwin, X. Liang, J. Klarbring, M. Dubajic, D. Dell'Angelo, C. Sutton, C. Caddeo, S. D. Stranks, A. Mattoni, A. Walsh, and G. Csányi, Dynamic local structure in caesium lead iodide: Spatial correlation and transient domains, Small **20**, 2303565 (2024).

[49] C. Quarti, E. Mosconi, J. M. Ball, V. D'Innocenzo, C. Tao, S. Pathak, H. J. Snaith, A. Petrozza, and F. De Angelis, Structural and optical properties of methylammonium lead iodide across the tetragonal to cubic phase transition: Implications for perovskite solar cells, Energy Environ. Sci. **9**, 155 (2016).

[50] X. Liang, J. Klarbring, and A. Walsh, Phase stability and transformations in lead mixed halide perovskites from machine learning force fields, Chem. Mater. **37**, 9392 (2025).

[51] I. Batatia, D. P. Kovacs, G. Simm, C. Ortner, and G. Csanyi, MACE: Higher order equivariant message passing neural networks for fast and accurate force fields, in *Advances in Neural Information Processing Systems*, edited by S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Curran Associates, Inc. Red Hook, NY, 2022), Vol. 35, pp. 11423–11436.

[52] I. Batatia, S. Batzner, D. P. Kovács, A. Musaelian, G. N. C. Simm, R. Drautz, C. Ortner, B. Kozinsky, and G. Csányi, The design space of E(3)-equivariant atom-centred interatomic potentials, Nat. Mach. Intell. **7**, 56 (2025).

[53] P. Schulz, D. Cahen, and A. Kahn, Halide perovskites: Is it all about the interfaces? Chem. Rev. **119**, 3349 (2019).

[54] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. **77**, 3865 (1996).

[55] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. **132**, 154104 (2010).

[56] A. Tkatchenko and M. Scheffler, Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data, Phys. Rev. Lett. **102**, 073005 (2009).

[57] J. Hermann and A. Tkatchenko, Density functional model for van der Waals interactions: Unifying many-body atomic approaches with nonlocal functionals, Phys. Rev. Lett. **124**, 146401 (2020).

[58] J. Ning, M. Kothakonda, J. W. Furness, A. D. Kaplan, S. Ehlert, J. G. Brandenburg, J. P. Perdew, and J. Sun, Workhorse minimally empirical dispersion-corrected density functional with tests for weakly bound systems: r^2SCAN+rVV10, Phys. Rev. B **106**, 075422 (2022).

[59] See Supplemental Material at http://link.aps.org/supplemental/10.1103/qcvb-y84n for computational details on structural optimizations, description of the ML architecture and training procedures, molecular dynamics settings, calculations of octahedral tilting angles, additional data for MLFFs trained on the slab and bulk datasets, which also includes Ref. [83].

[60] C. Quarti, F. De Angelis, and D. Beljonne, Influence of surface termination on the energy level alignment at the $CH_3NH_3PbI_3$ perovskite/C60 interface, Chem. Mater. **29**, 958 (2017).

[61] A. Seidu, M. Dvorak, P. Rinke, and J. Li, Atomic and electronic structure of cesium lead triiodide surfaces, J. Chem. Phys. **154**, 074712 (2021).

[62] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach, Comput. Phys. Commun. **167**, 103 (2005).

[63] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, CP2K: Atomistic simulations of condensed matter systems, WIREs Comput. Mol. Sci. **4**, 15 (2014).

[64] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, J. Chem. Phys. **81**, 511 (1984).
[65] W. G. Hoover, Constant-pressure equations of motion, Phys. Rev. A **34**, 2499 (1986).
[66] J. VandeVondele and J. Hutter, Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases, J. Chem. Phys. **127**, 114105 (2007).
[67] S. Goedecker, M. Teter, and J. Hutter, Separable dual-space Gaussian pseudopotentials, Phys. Rev. B **54**, 1703 (1996).
[68] M. Krack, Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals, Theor. Chem. Acc. **114**, 145 (2005).
[69] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, Machine learning of accurate energy-conserving molecular force fields, Sci. Adv. **3**, e1603015 (2017).
[70] S. Chmiela, H. E. Sauceda, K.-R. Müller, and A. Tkatchenko, Towards exact molecular dynamics simulations with machine-learned force fields, Nat. Commun. **9**, 3887 (2018).
[71] S. Chmiela, H. E. Sauceda, I. Poltavsky, K.-R. Müller, and A. Tkatchenko, sGDML: Constructing accurate and data efficient molecular force fields using machine learning, Comput. Phys. Commun. **240**, 38 (2019).
[72] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, *Ab initio* molecular simulations with numeric atom-centered orbitals, Comput. Phys. Commun. **180**, 2175 (2009).
[73] P. Ahlawat, Size dependent solid-solid crystallization of halide perovskites, arXiv:2404.05644.
[74] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. Dal Corso, S. de Gironcoli, T. Deutsch, J. K. Dewhurst, I. Di Marco, C. Draxl, M. Dułak, O. Eriksson, J. A. Flores-Livas, K. F. Garrity, L. Genovese, *et al.*, Reproducibility in density functional theory calculations of solids, Science **351**, aad3000 (2016).
[75] R. X. Yang, J. M. Skelton, E. Lora da Silva, J. M. Frost, and A. Walsh, Spontaneous octahedral tilting in the cubic inorganic cesium halide perovskites $CsSnX_3$ and $CsPbX_3$ (X = F, Cl, Br, I), J. Phys. Chem. Lett. **8**, 4720 (2017).
[76] J.-H. Lee, N. C. Bristowe, J. H. Lee, S.-H. Lee, P. D. Bristowe, A. K. Cheetham, and H. M. Jang, Resolving the physical origin of octahedral tilting in halide perovskites, Chem. Mater. **28**, 4259 (2016).
[77] Y. Yang, Y.-J. Chen, C. Hou, and T.-X. Liang, Ground-state surface of all-inorganic halide perovskites, J. Phys. Chem. C **126**, 21155 (2022).
[78] S. Ma, S. H. Kim, B. Jeong, H.-C. Kwon, S.-C. Yun, G. Jang, H. Yang, C. Park, D. Lee, and J. Moon, Strain-mediated phase stabilization: A new strategy for ultrastable α-$CsPbI_3$ perovskite by nanoconfined growth, Small **15**, 1900219 (2019).
[79] J. A. Steele, T. Braeckevelt, V. Prakasam, G. Degutis, H. Yuan, H. Jin, E. Solano, P. Puech, S. Basak, M. I. Pintor-Monroy, H. Van Gorp, G. Fleury, R. X. Yang, Z. Lin, H. Huang, E. Debroye, D. Chernyshov, B. Chen, M. Wei, Y. Hou, *et al.*, An embedded interfacial network stabilizes inorganic $CsPbI_3$ perovskite thin films, Nat. Commun. **13**, 7513 (2022).
[80] I. Poltavsky, A. Charkin-Gorbulin, M. Puleva, G. Fonseca, I. Batatia, N. J. Browning, S. Chmiela, M. Cui, J. Thorben Frank, S. Heinen, B. Huang, S. Käser, A. Kabylda, D. Khan, C. Müller, A. J. A. Price, K. Riedmiller, K. Töpfer, T. W. Ko, M. Meuwly, *et al.*, Crash testing machine learning force fields for molecules, materials, and interfaces: Model analysis in the TEA challenge 2023, Chem. Sci. **16**, 3720 (2025).
[81] I. Poltavsky, M. Puleva, A. Charkin-Gorbulin, G. Fonseca, I. Batatia, N. J. Browning, S. Chmiela, M. Cui, J. Thorben Frank, S. Heinen, B. Huang, S. Käser, A. Kabylda, D. Khan, C. Müller, A. J. A. Price, K. Riedmiller, K. Töpfer, T. W. Ko, M. Meuwly, *et al.*, Crash testing machine learning force fields for molecules, materials, and interfaces: Molecular dynamics in the TEA challenge 2023, Chem. Sci. **16**, 3738 (2025).
[82] C. M. Hylton-Farrington and R. C. Remsing, Dynamic local symmetry fluctuations of electron density in halide perovskites, Chem. Mater. **36**, 9442 (2024).
[83] A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, *et al.*, The atomic simulation environment—a Python library for working with atoms, J. Phys.: Condens. Matter **29**, 273002 (2017).
[84] I. Knysh, A. Charkin-Gorbulin, C. Quarti, D. Beljonne, I. Poltavsky, and A. Tkatchenko, Bridging surface dynamics and phase transitions in CsPbI3 Perovskite with machine-learning force fields and dispersion interactions [Dataset], Zenodo, 2026, https://doi.org/10.5281/zenodo.19686711.