# Dynamic Local Structure in Caesium Lead Iodide: Spatial Correlation and Transient Domains

*William J. Baldwin,\* Xia Liang, Johan Klarbring, Milos Dubajic, David Dell'Angelo, Christopher Sutton, Claudia Caddeo, Samuel D. Stranks, Alessandro Mattoni, Aron Walsh, and Gábor Csányi*

Metal halide perovskites are multifunctional semiconductors with tunable structures and properties. They are highly dynamic crystals with complex octahedral tilting patterns and strongly anharmonic atomic behavior. In the higher temperature, higher symmetry phases of these materials, several complex structural features are observed. The local structure can differ greatly from the average structure and there is evidence that dynamic 2D structures of correlated octahedral motion form. An understanding of the underlying complex atomistic dynamics is, however, still lacking. In this work, the local structure of the inorganic perovskite $CsPbI_3$ is investigated using a new machine learning force field based on the atomic cluster expansion framework. Through analysis of the temporal and spatial correlation observed during large-scale simulations, it is revealed that the low frequency motion of octahedral tilts implies a double-well effective potential landscape, even well into the cubic phase. Moreover, dynamic local regions of lower symmetry are present within both higher symmetry phases. These regions are planar and the length and timescales of the motion are reported. Finally, the spatial arrangement of these features and their interactions are investigated and visualized, providing a comprehensive picture of local structure in the higher symmetry phases.

## 1. Introduction

At finite temperatures metal halide perovskites exhibit complex and interesting structures. Methylammonium lead iodide ($CH_3NH_3PbI_3$) has been experimentally shown to exhibit static lattice twinning at room temperature,[1] as well as a stress-sensitive ferroelastic domain structure.[2] The average local strain of perovskite lattices has also been found to vary over large distances up to hundreds of nanometers or even micrometers.[3] There are also open questions about how local distortions of the lattice couple to electronic degrees of freedom and to light.[4,5]

Inorganic and organic perovskites have characteristic soft phonon modes which correspond to correlated octahedral tilting patterns. Along these modes, the potential energy surface can exhibit a shallow double well landscape, which gives rise to symmetry-breaking phase transitions.[6–8] During dynamics, these modes are overdamped with very short lifetimes.[9–12] In the high temperature, high symmetry perovskite phases, population of these modes results in interesting phenomena. For instance, it has been experimentally shown in $CsPbBr_3$ that excitations of these modes form a 2D, planar structure.[13] Similar structures have been found in $CH_3NH_3PbI_3$ using diffuse scattering experiments and classical molecular dynamics (MD) simulations.[14] These observations suggest that the local structure of metal

W. J. Baldwin, G. Csányi
Department of Engineering
University of Cambridge
Cambridge CB2 1PZ, UK
E-mail: wjb48@cam.ac.uk
X. Liang, J. Klarbring, A. Walsh
Department of Materials
Imperial College London
London SW7 2AZ, UK

J. Klarbring
Department of Physics
Chemistry and Biology (IFM)
Linköping University
Linköping SE-581 83, Sweden
M. Dubajic, S. D. Stranks
Department of Chemical Engineering and Biotechnology
University of Cambridge
Cambridge CB3 0AS, UK
D. Dell'Angelo, C. Caddeo, A. Mattoni
CNR-IOM, Unitá di Cagliari
Monserrato, Caligari 09042, Italy
C. Sutton
Department of Chemistry and Biochemistry
University of South Carolina
Columbia, SC 29208, USA
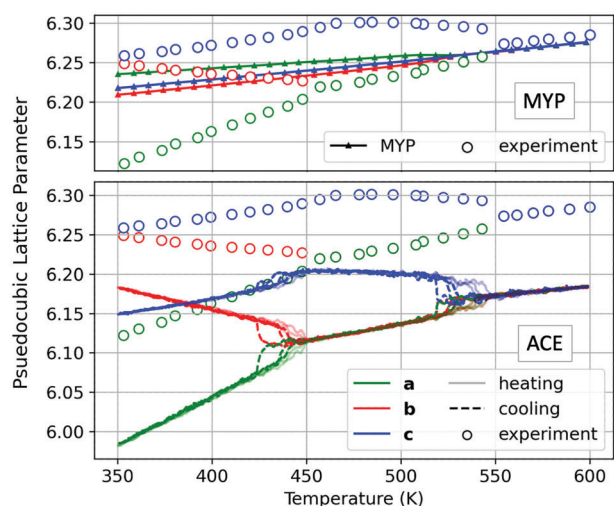
**Figure 1.** Evolution of the pseudocubic lattice constants of $CsPbI_3$ for the ACE potential and the MYP empirical force. These simulations were performed using a 13 720 atom simulation cell cooling over 4 ns. For ACE, three traces are shown for both heating and cooling. Experimental data reproduced with permission from Even and co-workers.[34]

halide perovskites can vary greatly from the average structure—a point that has also been explored computationally.[15]

In contrast to other Pb-I perovskites, less is known about $CsPbI_3$. The cubic phase of this material has a band gap that is well-suited for perovskite tandem solar cells.[16] However, poor thermodynamic stability of the black perovskite phases at ambient conditions[17] remains a significant challenge in photovoltaic applications. There is computational evidence to suggest that the planar correlation structures found in other materials are also present $CsPbI_3$. In particular, Yang and co-workers employed an effective Hamiltonian based approach and detected evidence of planar tilting correlation patterns.[18] Evidence of static symmetry broken domains has also been reported.[19] Evidently, the local structure $CsPbI_3$ is complex and a detailed picture of the structural dynamics is needed to tackle the stability of this material.

To analyze the structural dynamics of this system on the necessary length and time scales, a fast and accurate simulation approach is required. To date, several empirical force fields with a fixed functional form have been developed for organic and inorganic perovskites. The MYP (Model potential for hYbrid Perovskites) series of force fields[20,21] has proved useful for modeling hybrid and inorganic perovskites, and was used to study several ionic properties, such as the dielectric function of $CH_3NH_3PbI_3$[22] or the thermal conductivity in $CsPbI_3$.[23] The reactive force field ReaxFF and polarizable model AMEOBA have also been fitted to this material.[24,25] Both models are able to reproduce the phase transitions, but lack quantitative agreement for properties such as transition temperatures and ratios of lattice parameters. Alternatively, machine learning interatomic potentials (MLIPs) provide a flexible solution for accurate atomistic simulations.[26–29] For material science, MLIPs are now becoming mature and have been applied to tackle scientific questions which would have been out of reach using ab initio or empirical

methods.[30] In the context of halide perovskites, several modern MLIP architectures have been demonstrated to date.[31,32]

In this work, we characterize the dynamic local structure of $CsPbI_3$ using a new MLIP based on the atomic cluster expansion (ACE)[33] framework and large-scale simulations. Aiming to elucidate the three-dimensional structure and dynamics of $CsPbI_3$, we analyze spatial and temporal correlations of octahedral tilting angles for the three perovskite phases of the material. By investigating the temporal correlation of octahedral tilting in Section 2.2, we reveal that a double well effective potential is active in the cubic phase. In Section 2.3, we confirm that planar correlation structures observed in other materials also form in $CsPbI_3$. Furthermore, it is shown that the 3D octahedral tilting pattern of these local lower symmetry regions is inherited from the lower temperature phases. In Sections 2.4 and 2.5, the spatial arrangement and interactions of these lower symmetry regions is investigated. A detailed picture of the nature of the cubic and tetragonal phases is then presented. While planar regions of correlated octahedral motion exist, these alone do not fully describe the local structure. The spatial arrangement and overlapping nature of these features is needed to properly understand these symmetry-broken domains.

## 2. Results and Discussion

### 2.1. Machine Learning Potential Validation

A machine-learned interatomic potential based on the ACE framework has been trained for $CsPbI_3$ as described in Section 4. $CsPbI_3$ has four distinct phases which appear between room temperature and 600 K. There are three perovskite phases that form on cooling from 600 to 350 K: a cubic phase ($\alpha$), a tetragonal ($\beta$), and an orthorhombic ($\gamma$). At room temperature, the system transitions to a non-perovskite edge-sharing polymorph ($\delta$).[34] Since there is no continuous phase transition between the corner sharing perovskite phases and the $\delta$ phase, we have not studied the $\delta$ phase in this work.

Our model has been trained to describe these phases and validated by predicting experimental observables. In particular, we computed the variation of pseudocubic lattice parameters predicted by the model as a function of temperature. This was done by running a simulation of 13 720 atoms ($14^3$ pseudo cubic unit cells) in the NPT ensemble with a slowly varying temperature. The average pseudocubic lattice parameters can then be computed from the trajectory. **Figure 1** shows the predicted pseudocubic lattice parameter variation compared to experimental data obtained by in situ synchrotron X-ray diffraction provided by Even and co-workers.[34] The prediction of the MYP empirical force field[23] is also shown for comparison because force fields of this type have successfully been used to study organic perovskites. As it can be seen from Figure 1, the model is successful in predicting the overall experimental trends. There is an offset as the model under-predicts the lattice parameters by 1.6%. This is a feature of the density functional theory (DFT) reference data used to train the model, which utilized the local density approximation (LDA). In this 13 720 atom box and with a heating and cooling rate of 62.5 K ns$^{-1}$, some variability in the phase transition temperature is observed, but the nature of the transitions is correct. For

**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

NANO · MICRO
**small**
www.small-journal.com

**Table 1.** Finite temperature structural predictions for $CsPbI_3$ compared to experimental data from Even and coworkers.[34] Experimental values are shown in parentheses. Definitions of angles $\beta$ and $\delta$ are from Kepenekian et al.[35]

| | Temperature [K] | | |
| --- | --- | --- | --- |
| | 325 | 510 | 645 |
| a (Å) | 5.95 | 6.14 | 6.19 |
| a rescaled | 6.05 (6.095) | 6.24 (6.214) | 6.29 (6.297) |
| b (Å) | 6.20 | 6.14 | 6.19 |
| b rescaled | 6.30 (6.259) | 6.24 (6.241) | 6.29 (6.297) |
| c (Å) | 6.14 | 6.20 | 6.19 |
| c rescaled | 6.24 (6.250) | 6.30 (6.299) | 6.29 (6.297) |
| $\beta$ (degrees) | 12.4 (11.5) | 9.0 (8.6) | 0.0 (0.0) |
| $\delta$ (degrees) | 11.2 (9.9) | 0.0 (0.0) | 0.0 (0.0) |

discussion of finite time effects in such phase transitions, readers are referred to recent work by Fransson and co-workers.[32]

In **Table 1**, we show a numerical comparison of the experimental structural characteristics of the pseudocubic unit cell and those computed by simulations of a 1000 atom system in the NPT ensemble and averaging the atomic positions over time. We also report the re-scaled lattice parameters corresponding to a uniform 5% re-scaling of the volume. Overall, the model performs well, capturing the ratios of lattice constants and tilting angles accurately. Readers are referred to the supplementary material for further MLIP validation. The fact that the MYP force field is unable to capture the above structural features illustrates the need for machine learning potentials for studying this material.

## 2.2. Multiple Timescales of Octahedral Motion

Structural phase transitions of the general $ABX_3$ perovskite can occur through three types of distortions:[36] 1) octahedral deformations, 2) B-cation displacements, and 3) octahedral rotations. The latter commonly act as an order parameter that drives structural phase transitions in halide perovskites. Consequently, our focus will be primarily on these octahedral rotations. We define three tilting angles for each $PbI_6$ octahedron. This is done by first finding a best fit regular octahedron to a particular Pb site and then taking the Euler angles that map this octahedron to a perfect one in the laboratory reference frame. One can choose a convention to order the Euler angles such that in the limit of small rotations, the three angles correspond to signed tilting about the three coordinate axes. This is illustrated in **Figure 2**. Further detail about the tilting angle computation is given in the Supporting Information. We introduce the following notation to refer to tilting angles at a given time and position:

$$\theta_i(t; n_x, n_y, n_z) \tag{1}$$

where $i$ is the direction of tilt ($i = x$ refers to tilting around the $x$-axis), $t$ is the given time and $(n_x, n_y, n_z)$ are integer coordinates indexing the pseudocubic lattice site of the octahedron. All angles are reported in degrees.
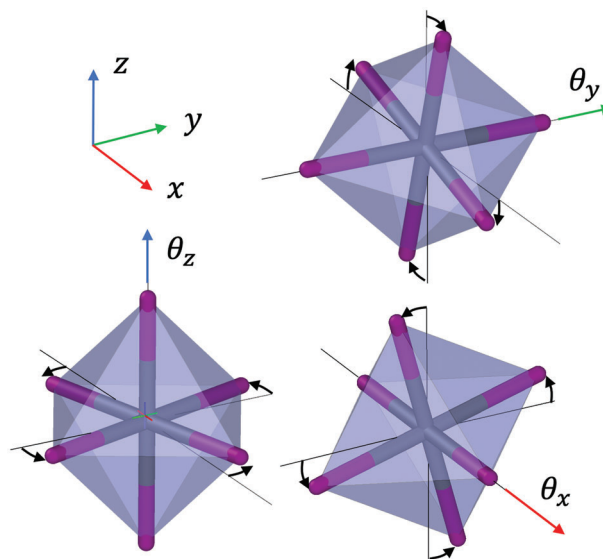


**Figure 2.** Visualizations of the $PbI_6$ octahedral tilting angles used in this study. In the limit of small rotations, the Euler angles that map a rotated octahedron to one aligned with the coordinate axes correspond to the rotations shown here.

To understand the timescales of tilting dynamics, we have examined the temporal autocorrelation function of local octahedral tilts. This describes how a single tilting angle is correlated with itself at a future time. With the above notation, the temporal autocorrelation function is defined as

$$A_i^{temporal}(\tau) = \frac{1}{\mathcal{N}} \mathbb{E}_{t,\mathbf{n}}[\, \theta_i(t; n_x, n_y, n_z)\, \theta_i(t + \tau; n_x, n_y, n_z)\, ] \tag{2}$$

where $\mathbb{E}_{t,\mathbf{n}}$ denotes the expectation over time and lattice sites and $\mathcal{N}$ is such that $A_i^{temporal}(0) = 1$. **Figure 3**b shows the three components of $A_i^{temporal}$ as a function of temperature. These data were collected from simulations of 69 120 atoms ($24^3$ pseudo cubic unit cells).

The asymptotes of the correlation functions reflect the Glazer notation[37] of each phase: In the cubic phase at 550 K, all three correlations decay to zero with time since all tilts are disordered with zero mean over long time scales ($a^0a^0a^0$). In the tetragonal phase, which has a tilt-configuration of $a^0a^0c^+$, only $\theta_z$ has a non-zero mean ($c^+$). This is reflected in the autocorrelation function at 480 K where only the $A_z(\tau)$ converges to a non-zero value. The same holds at 400 K in the orthorhombic phase, where the correlation of all three tilts converge to non-zero values and the $\theta_x$ and $\theta_y$ correlations are identical, in agreement with its $a^-a^-c^+$ tilt configuration.

At 400 K, the autocorrelation for $\theta_x$ and $\theta_y$ show evidence of a damped oscillatory mode with a time period on the order of 2 ps. In fact, as shown in the Supporting Information, the $A_i^t$ for the ordered tilts can be fitted accurately to damped harmonic oscillators. For the disordered tilts, at 480 and 550 K, we find an initial relaxation on the same timescale, followed by a longer decay component. To explain the shape of these correlation functions, we examine the tilt of a single octahedron tilt as a function of time. In Figure 3a, the $\theta_x$ value of a single octahedron at 550K

**2303565 (3 of 10)**
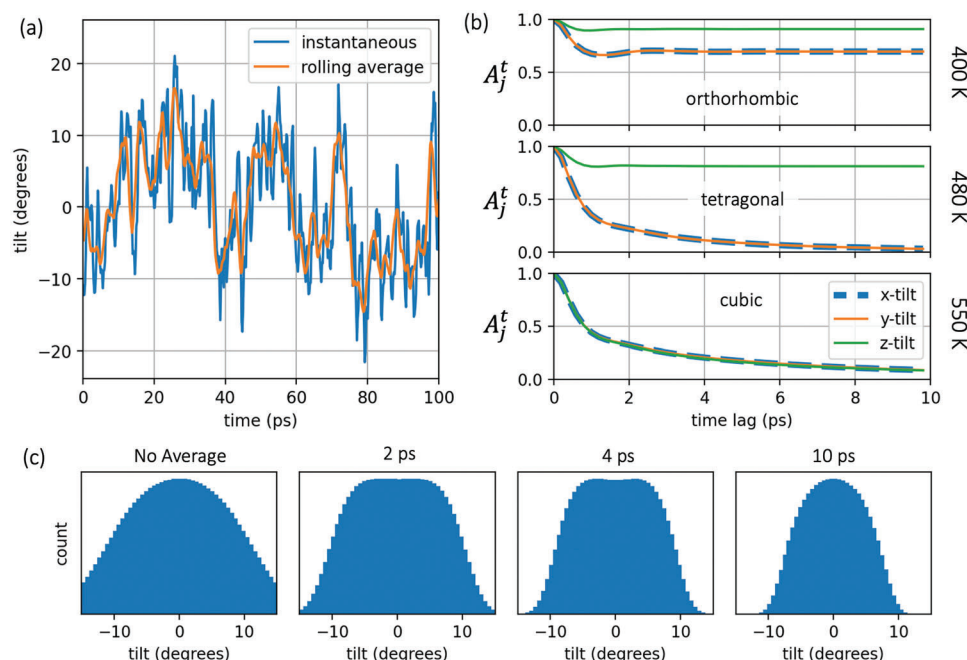
**Figure 3.** Temporal correlation structure of $PbI_6$ octahedral tilts in $CsPbI_3$. a) The $x$-direction tilt ($\theta_x$) of a single octahedron over the course of 100 ps at 550 K with a 200 fs sampling interval, and after applying a 2 ps rolling average. b) Temporal autocorrelation of all octahedra in the three perovskite phases. c) Histograms of instantaneous $\theta_x$ values at 550 K, as well as the same tilt after performing a rolling average over various timescales.

is plotted over the course of 100 ps. Also shown is the rolling average of this trajectory. A square window is used to perform the average—the angle at a given time is replaced by the average value over the next 2 ps. This is described in Equation (3). One can see evidence of a long time scale wandering motion of the averaged value, as well as high-frequency oscillations around the average. This behavior is reminiscent of analysis by Fransson [10] of octahedral tilting phonon modes in $CsPbBr_3$.
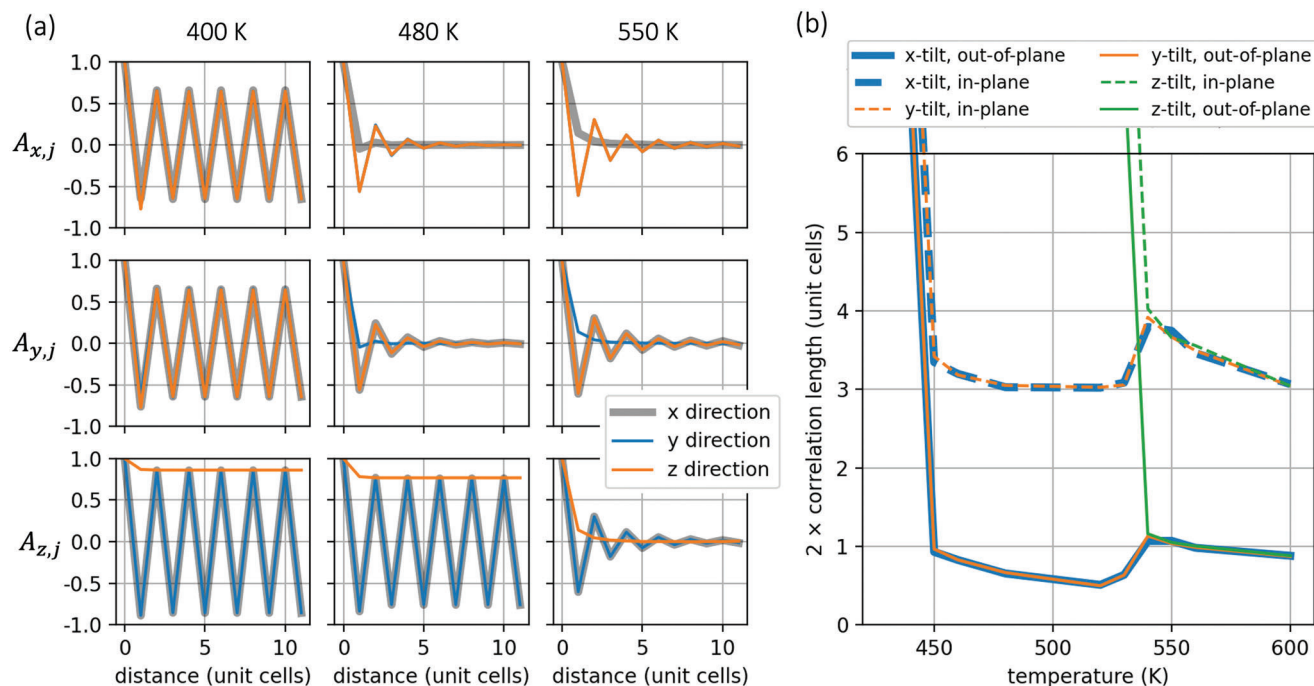
$$\tilde{\theta}_i(t; \mathbf{n}) = \frac{1}{T} \int_t^{t+T} \theta_i(t'; \mathbf{n}) dt' \qquad (3)$$

We have explored this further by examining the histogram of tilting angles after first performing rolling averages. Figure 3c shows the histogram of $\theta_x$ for both the raw data and after applying a rolling average of the tilting angles with a range of averaging timescales. For this analysis a square window was used to compute the rolling averages, however equivalent results are found by using a first-order Butterworth low pass filter[38] to remove the high-frequency components. In the instantaneous (without time averaging) picture, we find that the tilts are distributed in a unimodal distribution around zero. Naively, this would seem to indicate that the octahedra simply oscillate around their cubic, zero-tilt, reference positions, and that no local structure is present in the material. However, on averaging over a 1–10 ps timescale, a flat topped or double peaked distribution emerges, indicating that on this timescale the effective potential for octahedral rotation in fact has minima at non-zero tilt angles.

An explanation for this behavior is that the average tilt is hopping between two minima with a characteristic time on the order of 5 ps. Superposed on this average value are high-frequency oscillations around the current minimum. If the high-frequency oscillations give a Gaussian distribution about the current mean, then the total distribution would be the sum of these two such Gaussians at the two minima. For broad Gaussians, this sum would merge the two peaks, giving a unimodal instantaneous tilt distribution. If averages of the order of 4 ps are performed, however, these deviations from the mean are reduced, revealing the two minima. If an average is performed over 10 or more picoseconds, then the mean also jumps between both wells and the average converges to zero. Given that the double peak structure disappears when averaging over 10 ps, we can conclude that the characteristic hopping time is less than 10 ps. This can be compared to simulations performed by Fransson and co-workers.[10] They performed similar ML-MD simulations of $CsPbBr_3$, but instead decomposed atomic motions into supercell phonon modes. They reported that close to the upper transition in $CsPbBr_3$, the autocorrelation of the phonon mode component associated with in-phase tilting had a characteristic decay time of 5.22 ps. With this interpretation, the temporal autocorrelation functions for the tetragonal and cubic structures (at 480 and 550 K, respectively in Figure 3b) can be viewed as the superposition of the decorrelation of the high-frequency oscillatory behavior (as revealed in the orthorhombic at 400 K) and the longer timescale average tilt decorrelation.

The temporal correlation structure of the tilts in cubic $CsPbI_3$ indicates that multiple timescales are present. Performing temporal averaging of the tilts suggests that high frequency oscillations are masking an underlying double well effective potential. We can also conclude the characteristic hopping time of the mean value is between 4 and 10 ps. This effect has been shown in the cubic phase at 550 K, close to the cubic to tetragonal transition temperature of this potential of 533 K.

**ADVANCED**
**SCIENCE NEWS**
www.advancedsciencenews.com

NANO · MICRO
**small**
www.small-journal.com

**Figure 4.** Spatial correlation structure of $PbI_6$ octahedral tilts in $CsPbI_3$ with temperature. a) Spatial autocorrelation function for all tilting directions in the three perovskite phases. b) Trends in correlation length as a function of temperature.

## 2.3. Spatial Correlation Structure

The spatial structure of tilting dynamics can be studied using similar methods. In analogy to Equation (2), we can define the spatial autocorrelation function as:

$$A_{i,x}^{\text{spatial}}(d) = \frac{1}{\mathcal{N}} \mathbb{E}_{t,\mathbf{n}}[\, \theta_i(t; n_x, n_y, n_z)\, \theta_i(t; n_x + d, n_y, n_z)\,]$$

$$A_{i,y}^{\text{spatial}}(d) = \frac{1}{\mathcal{N}} \mathbb{E}_{t,\mathbf{n}}[\, \theta_i(t; n_x, n_y, n_z)\, \theta_i(t; n_x, n_y + d, n_z)\,] \qquad (4)$$

$$A_{i,z}^{\text{spatial}}(d) = \frac{1}{\mathcal{N}} \mathbb{E}_{t,\mathbf{n}}[\, \theta_i(t; n_x, n_y, n_z)\, \theta_i(t; n_x, n_y, n_z + d)\,]$$

The function $A_{i,j}^{\text{spatial}}$ has nine components since there are three tilting angles, and we can take the autocorrelation along three spatial directions. For example, $A_{x,y}(d)$ describes how the correlation of $\theta_x$ decays as one moves away in the $y$-direction. **Figure 4**a shows the spatial autocorrelation function for a range of temperatures spanning the three perovskite phases.

At 550 K, the spatial correlation function shows that each tilting direction exhibits anti-phase correlation along two spatial directions, which decays over the course of about five unit cells, as well as a weak in-phase correlation in the third direction that decays within 1 unit cell. Importantly, the spatial correlation length of $\theta_i$ is short in the $i$-direction, but long in the perpendicular directions. This can be interpreted as a two dimensional, planar correlation structure and is a result of the corner connectivity of the perovskite octahedral network.[37] All three octahedral tilting components exhibit their own planar correlation structure: $\theta_x$ is

correlated in the $y$–$z$ plane, $\theta_y$ is correlated in the $x$–$z$ plane and $\theta_z$ is correlated in the $x$–$y$ plane.

As the temperature decreases and the material undergoes a phase transition, certain octahedral tilts become "locked" as a result of symmetry breaking. This corresponds to the spatial autocorrelation function not approaching zero at large distances. According to the Glazer notation, the average structure of the tetragonal phase of $CsPbI_3$ can be denoted as $a^0 a^0 c^+$. This is captured in our simulations at 480 K. The correlations of $\theta_x$ and $\theta_y$ across all three spatial directions have an asymptote of zero ($a^0 a^0$) while the correlation of $\theta_z$ does not decay to zero and has in-phase correlation in the $z$-direction, corresponding to the $c^+$ in the Glazer notation. Due to the rigid coupling of the halide ions, $\theta_z$ always has anti-phase correlation in the in-plane directions ($x$ and $y$).[37]

The autocorrelation functions in Figure 4a reveal a surprising feature of the material. In the cubic phase, the out-of-plane correlation of $\theta_i$ is in-phase. This can be compared to the sign of the out-of-plane correlation of $\theta_z$ (the "locked" tilt) in the tetragonal phase, which is also in-phase. The dynamic correlations in the cubic phase therefore share the same 3D tilting pattern as the locked tilting angle in the tetragonal phase. A similar story holds for the dynamic correlations in the tetragonal phase. Here, the out-of-plane correlations for $\theta_x$ and $\theta_y$ are very small, but negative. This mirrors the pattern in the orthorhombic phase. The result is that the dynamic correlations in both phases have the same Glazer tilting pattern as the "locked" tilt in phase immediately below. The structure of the lower symmetry phases is preserved locally in the dynamic structure of higher symmetry phases.

Finally, one can also look at the variation in correlation length with temperature, which is shown in Figure 4b. The spatial correlation length was calculated by fitting an exponential decay
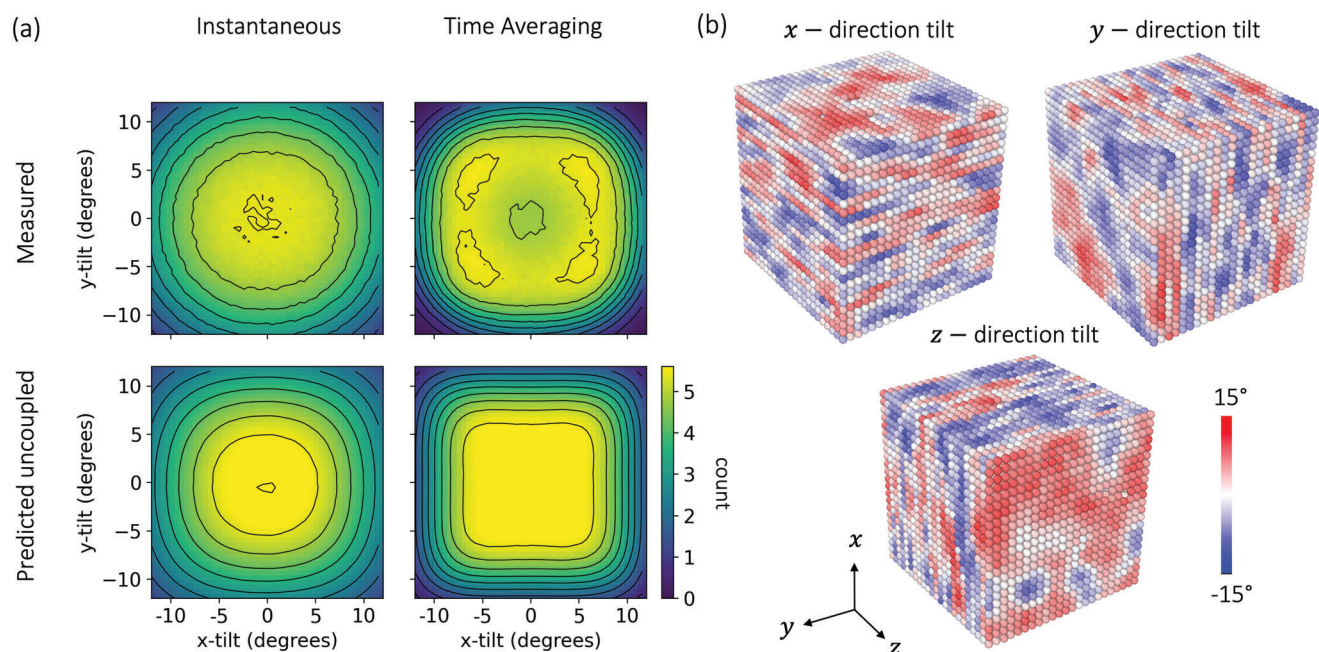
**Figure 5.** Nature of the cubic perovskite phase of CsPbI$_3$. a) Heat maps of $\theta_x$ and $\theta_y$ in the cubic phase at 550 K. The top row is the observed data, and the bottom row is the predicted distribution assuming that the two tilts are independent random variables. We have shown the results for both instantaneous tilts and for a rolling average of 2 ps. The scale (count) is arbitrary, and has been re-scaled for ease of interpretation. b) Visualization of the time averaged tilting structure. In each plot, the lead atoms are drawn with the color corresponding to one of the three tilting angles. Visualiations were produced using Ovito.

($e^{-d/\lambda}$) to the absolute values of the autocorrelation function. In Figure 4b, we have plotted $2\lambda$ which corresponds to the thickness or diameter of the correlated regions. One can see that the characteristic length scales in the tetragonal and cubic phases increase slightly as the temperature lowers toward each phase transition. In the context of instantaneous and time-averaged structure, one finds the same qualitative results for the spatial autocorrelation function when first performing time averaging of the tilting angles over 2 ps.

**2.4. Onset of the Cubic Phase**

So far, we have shown that both the tetragonal and orthorhombic phases have a planar structure of tilting correlations. Furthermore, in both of these phases, the local transient tilting correlation is the same as the global pattern in the next lower symmetry phase. The natural question is whether these correlations translate to local domains of tetragonal material within the cubic phase, or orthorhombic material in the tetragonal. If this is the case, one can also ask how these domains are arranged spatially. To properly characterise the cubic phase, we have assessed the correlation between different tilts at a single site, and visualized the three dimensional structure of the correlated tilts in the material.

First, one can generalize the histograms in Figure 3c to show the joint distribution of multiple tilting directions. **Figure 5**a shows a heat map of $\theta_x$ and $\theta_y$ for all octahedra at 550 K, close to the tetragonal transition temperature of this machine learned potential of 533 K. Since the cubic phase is isotropic, all three

tilting angles have the same distribution. Similarly, only the heat map of $\theta_x$ and $\theta_y$ is needed, since the other pairings will be identical. Heat maps of both instantaneous tilts, and the same quantity after applying a 2 ps rolling average, are shown. One can see that instantaneously the joint distribution is unimodal. Furthermore, there is no correlation between the different tilting angles—they are independent random variables. We can see this by comparing to the predicted heat map assuming that the two tilting angles are independent variables. This was computed as the product of the marginal distributions of each tilt, as discussed in the supplementary information. The two heat maps are very similar, implying that excitations of these tilting modes do not interact.

On the 2 ps timescale, the story is qualitatively different. Instead of a unimodal distribution, we find a multimodal distribution with a significant drop in intensity at the center and four faint peaks. The modes of the measured distribution are not on the axes, but instead in the four quadrants. This suggests that most octahedra have multiple large non-zero tilts. The interpretation—in terms of local domains of tilted octahedra—is that if planar regions of correlated tilting angle form, then different regions with different orientations typically overlap, rather than being mutually exclusive. Furthermore, when comparing to the predicted distribution given that the tilts are independent, it is clear that on this time scale the two variables are not independent. While the two plots share many features, there is a substantial decrease in intensity at the center indicating a strong tendency to avoid having both tilting angles equal to zero. In summary, Figure 5a shows that on the 2 ps timescale, most octahedra have multiple large nonzero tilts, and the excitations of different tilting mode directions are coupled.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**NANO · MICRO
small**

www.small-journal.com

This behavior has been presented close to the tetragonal transition temperature. Similar analysis at higher temperatures revealed that multimodality in the 2D and 3D joint tilt distributions persist for many tens of kelvin above the transition temperature, as shown in the Supporting Information. In particular, even at 600K (77 K above the transition temperature of this potential) there is still a strong tendency to avoid having all three tilting angles equal to zero, clearly demonstrating that local structure of lower symmetry is present even deep into the cubic phase.

Finally, one can visualize this behavior. This has been done by taking frames from the simulations used in the above analysis and coloring the octahedra according to a particular angle of tilt. To visually appreciate the size and shape of homogeneous regions, it is necessary to apply a mask to the tilting angles. This is because a homogeneous region of correlated $\theta_x$ in the cubic phase exhibits anti-phase correlation in two axes, and in-phase correlation in the third. More precisely, we can see from Figure 4a that the correlations have the same parity pattern as the average structure in the tetragonal phase. In Glazer notation, this is $a^0a^0c^+$. Therefore, we can apply the following operation to $\theta_x$ such that homogeneous regions with the tilting pattern of the tetragonal phase are rendered the same color:

$$\theta_x(t; n_x, n_y, n_z) \quad \rightarrow \quad (-1)^{n_y}(-1)^{n_z}\theta_x(t; n_x, n_y, n_z) \tag{5}$$

A similar operation can be done to the other tilting directions:

$$\theta_y(t; n_x, n_y, n_z) \quad \rightarrow \quad (-1)^{n_x}(-1)^{n_z}\theta_y(t; n_x, n_y, n_z)$$
$$\theta_z(t; n_x, n_y, n_z) \quad \rightarrow \quad (-1)^{n_x}(-1)^{n_y}\theta_z(t; n_x, n_y, n_z) \tag{6}$$

A single frame of the trajectory, after performing time averaging over 2 ps, has been processed in this way and is shown in Figure 5b. The three subplots correspond to coloring the octahedra according to the $x$-, $y$- and $z$-tilts respectively.

There is a clear planar structure in the patterns for each tilting direction. The normal of the planes is aligned with the tilting direction, as suggested by the spatial correlation functions. One can also see that areas which are largely homogeneous with large tilting angle when looking at one tilting direction intersect with equivalent regions in other plots. The tilted regions form and overlap with one another—in the same place, at the same time. Supplementary video 1 shows these visualizations over 400 ps of simulation time. When making the same visualizations without the time averaging, a similar planar structure is observed.

On this basis of these results, interpreting the cubic phase as a dynamic average of the tetragonal phase is difficult. The planar structures do share the correct tilting pattern, however, they form together, with different orientations, overlapping one another. In fact, the formation of these regions is coupled and they are slightly more likely to form together, rather than being mutually exclusive. This leads to a local structure which is not characteristic of a specific orientation of the tetragonal phase. Instead, most regions appear to be at the intersection of multiple planes of correlated tilts, and thus have an even lower local symmetry. In summary, the answer to the question "is the cubic phase a dynamic average of the lower symmetry phases?" is more nuanced. The local tilting patterns in the cubic phase share the pattern of

the tetragonal, but different orientations are superimposed at the same point in space and time.

## 2.5. Nature of the Tetragonal Phase

The same analysis can be applied to the tetragonal phase, as shown in **Figure 6**. Here, the $z$-direction tilt is globally locked. Therefore even though the structure is not isotropic, only the heat map of the two disordered tilting angles is shown. Qualitatively different behavior is found, in that the tilting heat maps for both the instantaneous and time-averaged data show excellent agreement between the measured and the predicted data given independent variables. Correlated planes in the tetragonal phase can robustly be viewed as uncoupled, non-interacting excitations on both timescales. Furthermore, the time averaged picture shows almost no splitting of the maximum. The same results are found at higher temperatures within the tetragonal range. It can be noted that the unimodal distributions of $\theta_x$ and $\theta_y$ are compatible with the non-zero anti-phase spatial correlations seen in the top two rows of figure 4a at 480 K. This is because even if the modal tilting value is zero, one can still ask whether neighboring tilts are typically the same direction or opposite. The octahedra with zero tilt simply contribute zero to this expectation.

An equivalent visualization is also shown. A different filter is applied to the tilting angles, since the dynamic correlations in this phase share the tilting pattern of the orthorhombic material, with anti-phase out-of-plane correlation (as in Figure 4). Homogeneous regions are therefore rendered the same colour by the following mapping:

$$\theta_x(t; n_x, n_y, n_z) \quad \rightarrow \quad (-1)^{n_x}(-1)^{n_y}(-1)^{n_z}\theta_x(t; n_x, n_y, n_z)$$
$$\theta_y(t; n_x, n_y, n_z) \quad \rightarrow \quad (-1)^{n_x}(-1)^{n_y}(-1)^{n_z}\theta_y(t; n_x, n_y, n_z) \tag{7}$$

Again, one can see the planar correlation in local tilting in the two disordered tilts. Video S2 (Supporting Information) shows these visualizations over the full trajectory.

The local structure of the tetragonal phase is, again, difficult to interpret in terms of local symmetry broken domains. While each tilting angle exhibits planar correlated structures, with the 3D tilting pattern of the orthorhombic material, the dynamics of the two disordered tilting angles are independent. The local structure is a superposition of these separate excitations.

## 3. Conclusions

Extensive large simulations using an accurate machine-learned interatomic potential have been used to study the local structure of $CsPbI_3$. Our study confirms that the local structure in the halide perovskite is dynamic in both space and time. We have shown through temporal averaging of the octahedral tilts that a double well behavior can still be found in the cubic phase. The timescale associated with the hopping motion is in the 5–10 ps range.

Analysis of the spatial autocorrelation function has revealed a planar correlation structure in the tetragonal and cubic phases. Furthermore, the local transient regions in these phases show a 3D tilting pattern which matches that of the Glazer tilting pattern

ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

NANO · MICRO
small
www.small-journal.com
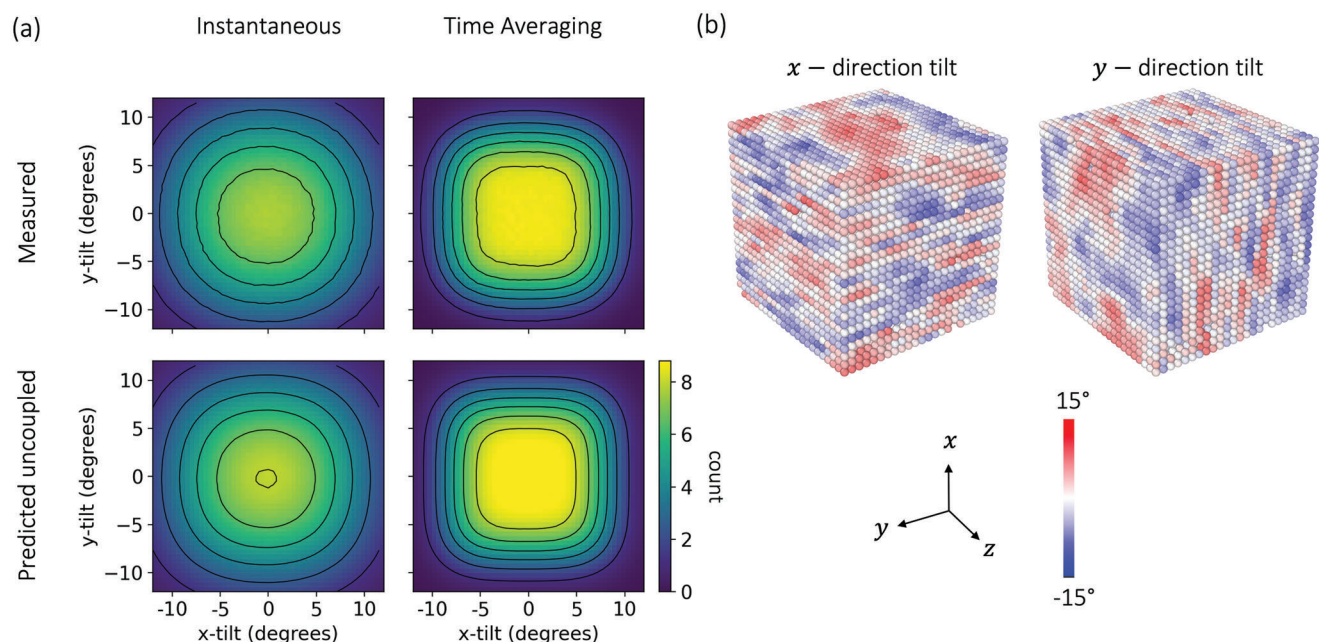
**Figure 6.** Nature of the tetragonal phase of $CsPbI_3$. a) Heat maps of the x- and y-direction tilts in the tetragonal phase at 450 K. The scale (count) is arbitrary and has been re-scaled for ease of interpretation. b) Visualization of the time averaged tilting structure. In each plot, the lead atoms are drawn with the color corresponding to one of the three tilting angles. Visualizations were produced using Ovito.

next lower temperature phase. Visualizations reveal that these correlations do indeed correspond to large planar regions of similarly tilted octahedra. Analysis of the joint distribution of tilts then reveals that these planar regions instantaneously appear to be uncoupled excitations. On the other hand, when removing higher frequency oscillations through the same time averaging procedure, one finds some correlation in these features. In the cubic phase, the structure is composed of dynamic, preferentially overlapping 2D tilted regions. In the tetragonal phase, the behavior is qualitatively different, with the planar tilting excitations being independent both instantaneously and on longer timescales. Questions remain about how specific this behavior is to $CsPbI_3$ and we hope this will be investigated in the future.

The accuracy and efficiency of modern machine learning potentials have made this work possible. Due to the large correlation lengths and long timescales, these simulations would have been impossible using first-principles methods. At the same time, empirical force fields that perform well for organic systems are unable to quantitatively describe the nature of systems such as this inorganic perovskite. Systematically convergable machine learning frameworks such as the atomic cluster expansion are therefore creating new avenues of potential research in materials science.

## 4. Experimental Section

*Atomic Cluster Expansion*: The machine learning model in this work was based on the Atomic Cluster Expansion (ACE) framework.[33,39,40] A julia implementation of the descriptor and fitting process was used and production simulations were performed using the Performant ACE evaluator[41] in the Large-scale Atomic/Molecular Massively Parallel Sim-

ulator (LAMMPS).[42] The ACE1 julia code is under development at https://github.com/ACEsuit/ACE1.jl , and the exact version of ACE1 used in this project, including the modifications described below, can be found at https://github.com/WillBaldwin0/ACE1.jl/tree/ortho-specdep-pairpot.

ACE models allowed for a body-ordered expansion of the potential energy surface (PES) and the model presented in this work used four body descriptors.

*Two-Body Radial Basis Functions*: In this project, additional work was put into designing ACE models that robustly predicted strong interatomic repulsion at close approaches. In an ACE model, the distance between a pair of atoms was described by a set of radial basis functions.[39] In the ACE1 Julia implementation, this radial basis was defined as set of polynomials $p_n(x)$ that were orthogonal with respect to a weighting function $w(x)$:

$$\delta_{nm} = \int_0^{r_{cut}} p_n(r) p_m(r) w(r) dr \tag{8}$$

It was also possible to first transform the radial coordinate by a function f, so that $p(r) \rightarrow p(f(r))$, in order to provide more radial resolution at certain distances.[43] For this project, a special choice was made for the weighting function $w(r)$ to encourage strong interatomic repulsion between atoms, even at distances much smaller than those seen in the training set. The weighting function $w(r)$ was constant beyond the typical nearest neighbor distance, but below this, it was shaped to approximately match the species dependent radial distribution function of the training set. The intuition behind this choice was that the weighting function is related to the regularization of the potential energy landscape through the $L^2$ regularization of the fitting procedure. If $w(r)$ is relatively large in some interval $r \in [a, b]$, the resulting potential energy landscape is heavily regularized in this region. By choosing $w(r)$ to match the radial distribution function, the potential energy landscape gradually becomes less regularized as the data becomes less frequent at small interatomic separations. This allows PES to be larger and have high curvature at small separations. This procedure was useful to ensure stable dynamics, but the shape of the weighting function did not affect the accuracy of the model.

*Bayesian Regression and Active Learning*: As originally described by Drautz,[39] the atomic cluster expansion was used to construct a linear model of the PES. The coefficients were subsequently fitted using Bayesian linear regression. The Bayesian linear regression algorithm returned the fitted parameters—which can be interpreted as the mean of the posterior distribution over the parameters—and a covariance matrix encoding the uncertainty in the solution. This covariance matrix could be used to infer an uncertainty in a future prediction. In practice, this was done by drawing a small number samples from the posterior parameter distribution, and constructing an ensemble of models.

The training procedure in this study involved two steps. First, the Hyper Active Learning (HAL) Bayesian active learning framework[44] was used to produce a model that was able to simulate the material over long time scales, with usable accuracy. A summary of the HAL procedure is as follows:

1) Generate a small initial database of training configurations labeled with DFT, for instance by perturbing an equilibrium structure.
2) Fit an ACE model to this database using Bayesian linear regression.
3) Run a molecular dynamics simulation using the ACE model, tracking the predicted uncertainty at each simulation step (see above). The dynamics is propagated based on a weighted sum of the potential energy, and a term representing the uncertainty of the models prediction. Hence, the dynamics is driven to areas in configuration space with a higher model uncertainty.[44]
4) If the predicted uncertainty exceeds some predetermined threshold, stop the simulation and perform the reference calculation on the current configuration.
5) Add this labeled configuration to the dataset, retrain the model, and restart the simulation.

Once HAL had been used to create a stable model for the system, the resulting model was used to sample 214 configurations from a range of temperatures representative of the four phases of interest, which made up the final dataset. This included configurations of up to 160 atoms, sampled from both constant pressure and constant volume simulations.

*Reference Data*: All density functional theory calculations were performed using the FHI-aims DFT code at the LDA level of theory.[45] A k-point density of 16 Å$^{-1}$ was used, and "tight" basis set and integration grids were used throughout.

*Simulation Parameters*: Unless otherwise stated, all production simulations were performed in the constant pressure, constant temperature ensemble using a simulation cell of 69 120 atoms, corresponding to $24^3$ pseudo cubic unit cells. The timestep was 4 fs. Statistics were collected by starting from the expected structure at the target temperature (obtained by sampling from slow cooling simulations) and equilibrating for 1 ns. A further nanosecond was then sampled every 200 fs.

*MYP Force Field Parameters*: $CsPbI_3$ parameters correspond to Ref. [23] but for a 7% increase of the Cs-I Buckingham prefactor, here set to 150 000 kcal mol$^{-1}$, in order to better reproduce the experimental transition temperatures.

## Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

## Acknowledgements

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The machine learning potential, training and test datasets and example scripts on how to run the potential in LAMMPS are available at https://github.com/WillBaldwin0/cspbi3_mlip_docker. A docker image with a working LAMMPS installation and example scripts is also available through the same link. The trajectory files of all simulations that we have analysed are available in a zenodo repository at 10.5281/zenodo.8252919.

## Keywords

[1] J. Breternitz, M. Tovar, S. Schorr, *Sci. Rep.* **2020**, *10*, 16613.
[2] E. Strelcov, Q. Dong, T. Li, J. Chae, Y. Shao, Y. Deng, A. Gruverman, J. Huang, A. Centrone, *Sci. Adv.* **2017**, *3*, 1602165.
[3] E. M. Tennyson, T. A. Doherty, S. D. Stranks, *Nat. Rev. Mater.* **2019**, *4*, 573.
[4] T. W. Jones, A. Osherov, M. Alsari, M. Sponseller, B. C. Duck, Y.-K. Jung, C. Settens, F. Niroui, R. Brenes, C. V. Stan, Y. Li, M. Abdi-Jalebi, N. Tamura, J. E. Macdonald, M. Burghammer, R. H. Friend, V. Bulović, A. Walsh, G. J. Wilson, S. Lilliu, S. D. Stranks, *Energy Environ. Sci.* **2019**, *12*, 596.
[5] T.-C. Wei, H.-P. Wang, T.-Y. Li, C.-H. Lin, Y.-H. Hsieh, Y.-H. Chu, J.-H. He, *Adv. Mater.* **2017**, *29*, 1701789.
[6] R. X. Yang, J. M. Skelton, E. L. da Silva, J. M. Frost, A. Walsh, *J. Phys. Chem. Lett.* **2017**, *8*, 4720.
[7] J. S. Bechtel, A. Van der Ven, *Phys. Rev. Mater.* **2018**, *2*, 025401.
[8] J. Klarbring, *Phys. Rev. B* **2019**, *99*, 104105.
[9] M. Songvilay, N. Giles-Donovan, M. Bari, Z.-G. Ye, J. L. Minns, M. A. Green, G. Xu, P. M. Gehring, K. Schmalzl, W. D. Ratcliff, C. M. Brown, D. Chernyshov, W. van Beek, S. Cochran, C. Stock, *Phys. Rev. Mater.* **2019**, *3*, 093602.
[10] E. Fransson, P. Rosander, F. Eriksson, J. M. Rahm, T. Tadano, P. Erhart, *Communications Physics* **2023**, *6*, 173.
[11] A. Cohen, T. M. Brenner, J. Klarbring, R. Sharma, D. H. Fabini, R. Korobko, P. K. Nayak, O. Hellman, O. Yaffe, *Adv. Mater.* **2022**, *34*, 2107932.
[12] M. A. Carignano, S. A. Aravindh, I. S. Roqan, J. Even, C. Katan, *J. Phys. Chem. C* **2017**, *121*, 20729.
[13] T. Lanigan-Atkins, X. He, M. J. Krogstad, D. M. Pajerowski, D. L. Abernathy, G. N. Xu, Z. Xu, D. Y. Chung, M. G. Kanatzidis, S. Rosenkranz, R. Osborn, O. Delaire, *Nat. Mater.* **2021**, *20*, 977.

[14] N. J. Weadock, T. C. Sterling, J. A. Vigil, A. Gold-Parker, I. C. Smith, B. Ahammed, M. J. Krogstad, F. Ye, D. Voneshen, P. M. Gehring, A. M. Rappe, H.-G. Steinrück, E. Ertekin, H. I. Karunadasa, D. Reznik, M. F. Toney, *Joule* **2023**, *7*, 1051.

[15] X.-G. Zhao, G. M. Dalpian, Z. Wang, A. Zunger, *Phys. Rev. B* **2020**, *101*, 155137.

[16] R. E. Beal, D. J. Slotcavage, T. Leijtens, A. R. Bowring, R. A. Belisle, W. H. Nguyen, G. F. Burkhard, E. T. Hoke, M. D. McGehee, *J. Phys. Chem. Lett.* **2016**, *7*, 746.

[17] Z. Yao, W. Zhao, S. Liu, *J. Mater. Chem. A* **2021**, *9*, 11124.

[18] L. Chen, B. Xu, Y. Yang, L. Bellaiche, *Adv. Funct. Mater.* **2020**, *30*, 1909496.

[19] F. Bertolotti, L. Protesescu, M. V. Kovalenko, S. Yakunin, A. Cervellino, S. J. Billinge, M. W. Terban, J. S. Pedersen, N. Masciocchi, A. Guagliardi, *ACS Nano* **2017**, *11*, 3819.

[20] A. Mattoni, A. Filippetti, M. I. Saba, P. Delugas, *J. Phys. Chem. C* **2015**, *119*, 17421.

[21] T. Hata, G. Giorgi, K. Yamashita, C. Caddeo, A. Mattoni, *J. Phys. Chem. C* **2017**, *121*, 3724.

[22] A. Mattoni, C. Caddeo, *J. Chem. Phys.* **2020**, *152*, 104705.

[23] A. Giri, S. Thakur, A. Mattoni, *Chem. Mater.* **2022**, *34*, 9569.

[24] M. Pols, J. M. Vicent-Luna, I. Filot, A. C. V. Duin, S. Tao, *J. Phys. Chem. Lett.* **2021**, *12*, 5519.

[25] P. V. G. M. Rathnayake, S. Bernardi, A. Widmer-Cooper, *J. Chem. Phys.* **2020**, *152*, 024117.

[26] A. P. Bartók, S. De, C. Poelking, N. Bernstein, J. R. Kermode, G. Csányi, M. Ceriotti, *Sci. Adv.* **2017**, *3*, 1701816.

[27] F. Musil, A. Grisafi, A. P. Bartók, C. Ortner, G. Csányi, M. Ceriotti, *Chem. Rev.* **2021**, *121*, 9759.

[28] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, A. Walsh, *Nature* **2018**, *559*, 547.

[29] V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, G. Csányi, *Chem. Rev.* **2021**, *121*, 10073.

[30] V. L. Deringer, N. Bernstein, G. Csányi, C. B. Mahmoud, M. Ceriotti, M. Wilson, D. A. Drabold, S. R. Elliott, *Nature* **2021**, *589*, 59.

[31] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, M. Bokdam, *Phys. Rev. Lett.* **2019**, *122*, 225701.

[32] E. Fransson, J. Wiktor, P. Erhart, *J. Phys. Chem. C* **2023**, *127*, 13773.

[33] R. Drautz, *Phys. Rev. B* **2020**, *102*, 024104.

[34] A. Marronnier, G. Roma, S. Boyer-Richard, L. Pedesseau, J. M. Jancu, Y. Bonnassieux, C. Katan, C. C. Stoumpos, M. G. Kanatzidis, J. Even, *ACS Nano* **2018**, *12*, 3477.

[35] 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, M. Kepenekian, *ACS Nano* **2016**, *10*, 9776.

[36] C. J. Howard, *Acta Crystallographica. Section A, Foundations of Crystallography* **2005**, *61*, 93.

[37] A. M. Glazer, *Acta Crystallogr. B* **1972**, *28*, 3384.

[38] S. Butterworth, *Experimental Wireless and the Wireless Engineer* **1930**, *7*, 536.

[39] R. Drautz, *Phys. Rev. B* **2019**, *99*, 014104.

[40] G. Dusson, M. Bachmayr, G. Csányi, R. Drautz, S. Etter, C. van der Oord, C. Ortner, *J. Comput. Phys.* **2022**, *454*, 110946.

[41] Y. Lysogorskiy, C. v. d. Oord, A. Bochkarev, S. Menon, M. Rinaldi, T. Hammerschmidt, M. Mrovec, A. Thompson, G. Csányi, C. Ortner, R. Drautz, *npj Comput. Mater.* **2021**, *7*, 97.

[42] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, S. J. Plimpton, *Comput. Phys. Commun.* **2022**, *271*, 108171.

[43] D. P. Kovács, C. v. d. Oord, J. Kucera, A. E. A. Allen, D. J. Cole, C. Ortner, G. Csányi, *J. Chem. Theory Comput.* **2021**, *17*, 7696.

[44] C. van der Oord, M. Sachs, D. P. Kovács, C. Ortner, G. Csányi, Hyperactive learning (hal) for data-driven interatomic potentials, **2022**, https://arxiv.org/abs/2210.04225, (Preprint).

[45] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, M. Scheffler, *Comput. Phys. Commun.* **2009**, *180*, 2175.

**2303565 (10 of 10)**