Dynamic Vacancy Levels in $CsPbCl_3$ Obey Equilibrium Defect Thermodynamics

Irea Mosquera-Lois and Aron Walsh*

Thomas Young Centre & Department of Materials, Imperial College London, London SW7 2AZ, United Kingdom

(Received 15 August 2025; revised 29 September 2025; accepted 20 October 2025; published 12 November 2025)

Halide vacancies are the dominant point defects in perovskites, with V_{Cl} identified as a detrimental trap for the optoelectronic performance of $CsPbCl_3$, which has applications ranging from photodetectors to solar cells. Understanding these defects under operating conditions is key since their electronic levels exhibit large thermal fluctuations that challenge the validity of static 0 K models. However, quantitative modeling of defect processes requires hybrid density functional theory with spin-orbit coupling, which is too expensive for direct molecular dynamic simulations. To address this, we train a multitask machine learning force field to study V_{Cl} in orthorhombic $CsPbCl_3$ at 300 K. While we observe strong oscillations in the optical transition level arising from the soft potential energy surface, neither the nonradiative capture barriers nor the thermodynamic charge transition levels are affected. Our results reveal that V_{Cl} is not responsible for the nonradiative losses previously assumed. Instead, its impact on performance arises from other mechanisms, such as limiting the open-circuit voltage and promoting ionic migration. Our findings demonstrate that, despite strong dynamical effects in halide perovskites, the conventional static formalism of defect theory remains valid for predicting thermodynamic behavior, providing a sound basis for the design of high-performance energy materials.

DOI: 10.1103/dxmb-8s96

I. INTRODUCTION

Point defects dictate the physical properties of most functional materials. Their dilute concentrations in crystals challenge experimental characterization, requiring a combination of experiment and theory to understand and optimize defect behavior. Typically, theoretical studies assume a static framework by considering the 0 K ground-state structures for all possible defect species and calculating their concentrations and optoelectronic properties [1]. However, recent studies have questioned the validity of this approximation for soft crystals like metal halide perovskites, which exhibit large-amplitude and anharmonic structural dynamics at room temperature [2–5]. It remains an open question whether this is linked to the empirical defect tolerance observed for perovskite materials and devices [6].

There have been multiple reports on the strong and slow fluctuations of the electronic levels of halide vacancies in $CsPbBr_3$ [7–9], $CsPbI_3$ [10–12], and $CH_3NH_3PbI_3$ [13–15] at device operating conditions of solar cells, resulting in the defect oscillating between a deep and shallow character on the picosecond timescale, and thus, drastically changing its behavior. However, the conclusions of these reports are limited by their level of theory (simpler approximations within density functional theory), which can lead to an incorrect description of the potential energy surface (PES) for the halide vacancies [16–23] and overestimates the defect level fluctuations [16]. Another limitation stems from their focus on the variation of electronic *eigenvalues* and optical levels, rather than the *thermodynamic* transition levels. While optical or vertical levels are important for radiative processes (e.g., defect absorption or photoluminescence), the impact of defects on device performance is often controlled by nonradiative processes like nonradiative carrier capture, which is determined by the thermodynamic transition levels [6,18].

In this study, we overcome these limitations by using an accurate level of theory to describe defect processes in lead halide perovskites that combines the hybrid exchange-correlation functional (HSE [24]) with spin-orbit coupling (SOC). Using this description, we quantify the fluctuations of the optical transition levels at 300 K and relate them to the corresponding thermodynamic charge transition levels, thereby revealing how thermal motion affects defect electronic behavior. Due to the high computational cost of HSE + SOC, we leverage multitask machine learning force fields (MLFFs) to run molecular dynamics over long

*Contact author: a.walsh@imperial.ac.uk

time and length scales that would be prohibitively expensive from first principles. We target $CsPbCl_3$ due to its relevance for applications in tandem solar cells [25–27] and radiation detectors [28]. Our focus is on the chloride vacancy ($V_{Cl}^{0,+}$) in orthorhombic [28–31] $CsPbCl_3$, since it has been suggested as the main defect promoting nonradiative carrier recombination and damaging device performance [32–41], thereby motivating a range of passivating strategies [32,38,42–48].

II. RESULTS

A. Machine learning force field

To accurately describe the PES of halide vacancies in lead-based perovskites, a high level of theory is required [16–22]. Relativistic effects are necessary to describe the Pb 6*p*-based conduction band, which results in an energy shift of almost 1 eV when spin-orbit coupling is included [49,50]. The use of a hybrid exchange-correlation functional is further necessary to remove large errors in the band gap and the alignment of defect levels with the band edges [21,22]. As illustrated in Fig. 1(a), a low level of theory like Perdew-Burke-Ernzerhof (PBE) with scalar-relativistic effects leads to significant errors in the relative energies of V_{Cl}^{0} when compared to a high-level reference (HSE + SOC), especially for high-energy configurations. While a hybrid functional (HSE) reduces these errors, it is still limited for configurations with a long separation between the vacancy nearest neighbors [Fig. 1(c)], thus resulting in incorrect dynamics and the need to include SOC. However, the computational cost of advanced methods like HSE + SOC makes them impractical for molecular dynamics simulations to probe defect behavior at finite temperatures.

To overcome this obstacle, we use a multihead message-passing atomic cluster expansion (MACE) [51] machine learning force field that integrates training data calculated with two levels of theory. This approach allows us to leverage a large dataset of low-cost low-fidelity calculations (PBE), supplemented by a much smaller set of expensive high-fidelity data (HSE + SOC). As shown in Fig. 1(b), the model shares most of its parameters for transforming atomic environments into features but uses separate readout layers to map these features to atomic energies. This architecture enables the model to learn general structure-energy relationships from low-fidelity data, while high-fidelity data corrects for the subtle differences between the two potential energy surfaces. The trained model can reproduce both levels of theory and accurately describes the PES of V_{Cl}^{0} as the distance between the Pb atoms neighboring the vacancy varies [Fig. 1(c)]. Finally, to describe the transition between charge states, we require an accurate description of the potential energy surface for both V_{Cl}^{0} and V_{Cl}^{+}. We therefore train and validate separate models for each charge state, with both models achieving good accuracies on their respective test sets (see Appendix C).

B. Optical defect levels

After validating the models, we apply them to investigate how temperature affects the optical defect level, which corresponds with an instantaneous transition between charge states where the defect geometry remains fixed. To evaluate the optical level $\epsilon_{opt}^{0\rightarrow+}$ at a typical operating temperature of 300 K, we perform *NPT* molecular dynamics for the neutral charge state using the model trained on V_{Cl}^{0}. For each frame R in the neutral trajectory, we calculate the energy difference between the neutral and positively charged states using separate machine learning models for each charge state:

$$\epsilon_{opt}^{0\rightarrow+}(R_0) = E^{+}(R_0) - E^{0}(R_0), \tag{1}$$

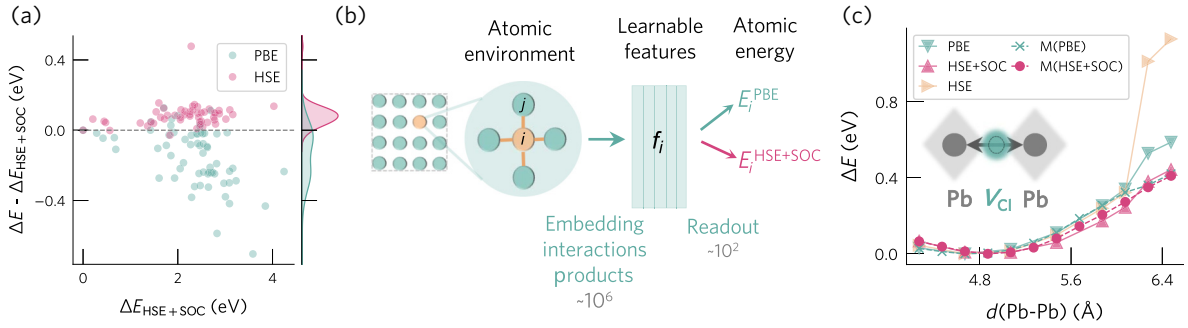

FIG. 1. (a) Errors in the relative energies of V_{Cl}^{0} for a range of sampled configurations from molecular dynamics between 0 and 300 K calculated with PBE (green) and HSE (pink) in comparison with HSE + SOC. Both PBE and HSE lead to significant differences and will thus affect the resulting dynamics. (b) Schematic of the architecture of the multihead model, illustrating the common operations and parameters to learn the atomic features and separate readout layers used for each level of theory. (c) Potential energy surface of V_{Cl}^{0} (at the site with C_s symmetry) in orthorhombic $CsPbCl_3$, mapping the energy versus the distance between the Pb atoms neighboring the vacancy. Reference density functional theory (DFT) data are shown as triangles, with different colors indicating different levels of theory. Model predictions [dashed lines with crosses, M(PBE), and circles, M(HSE + SOC)] accurately reproduce reference data.

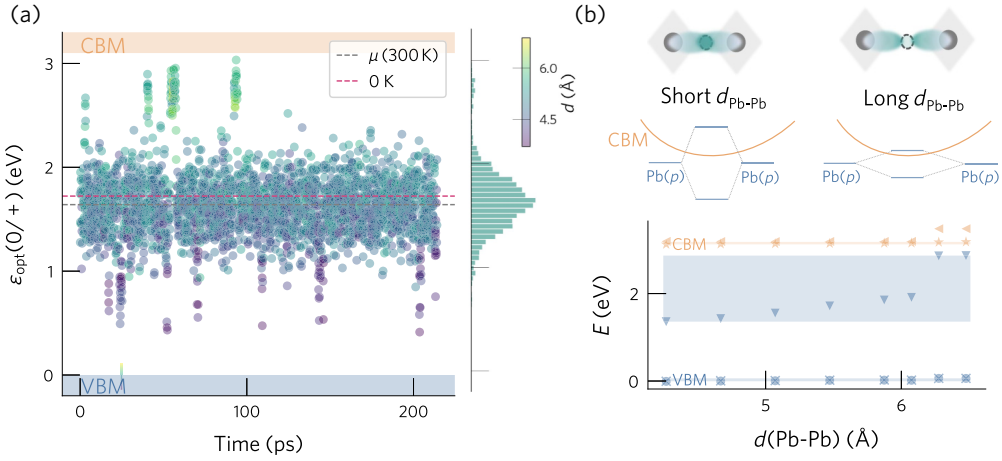

FIG. 2. (a) Variations of the optical defect level $\epsilon_{opt}(0 \rightarrow +)$ at 300 K and comparison with the 0 K value (pink) for a 640-atom supercell containing a single vacancy. Color map quantifies the separation between Pb neighboring the vacancy, with shorter distances resulting in stronger Pb(p) overlap, and thus, a level deeper in the gap. These level fluctuations are analysed in Fig. 10 through a Fourier transform to identify the vibrational frequencies driving the stronger fluctuations. (b) Electronic eigenvalues (calculated with HSE + SOC) of $V_{Cl}{}^0$ configurations in a 160-atom supercell with increasing Pb-Pb distance, illustrating that the defect level approaches the conduction band minimum as the separation between Pb atoms increases, as previously reported in the literature [7]. Occupied and unoccupied levels are shown in blue and orange, respectively, while the filled area highlights the variation of each eigenvalue.

where the energies $E^{q=0}$ and $E^{q=+}$ are evaluated with the models for the neutral and positively charged vacancy, respectively, and R_0 denotes the structure in the neutral state. The time evolution of $\epsilon_{opt}^{0\rightarrow+}$ exhibits strong fluctuations of more than 1 eV [Fig. 2(a)], which agrees with previous reports on halide vacancies in $CsPbBr_3$ [7] and $CsPbI_3$ [10–12] and observations of broad subgap emission around 1.65 eV in the photoluminescence spectra of $CH_3NH_3PbCl_3$ [43,52].

The fluctuations of $\epsilon_{opt}^{0\rightarrow+}$ are partly caused by changes in the separation between the two Pb atoms neighboring the vacancy. When a neutral Cl vacancy forms, the dangling p orbitals of the Pb atoms neighboring the vacancy interact and form bonding and antibonding combinations, with the bonding orbital appearing in the band gap and hosting one electron. At finite temperature, the Pb atoms oscillate around their equilibrium position, with shorter Pb-Pb distances resulting in stronger orbital overlap, and thus, lowering the energy of the bonding orbital. This relationship is illustrated in Fig. 2(b) by evaluating the electronic eigenvalues for $V_{Cl}{}^0$ configurations with increasing Pb-Pb separation with DFT (HSE + SOC).

Beyond changes in d(Pb-Pb), $\epsilon_{opt}^{0\rightarrow+}$ is also broadened by thermal fluctuations of the global structure, including volume variations and octahedral tilting and distortions, reflecting the softness of the perovskite lattice. Despite the significant ionic motion at 300 K and its impact on $\epsilon_{opt}^{0\rightarrow+}$, we find that its ensemble average is in good agreement with the static prediction based on the 0 K ground-state structure of $V_{Cl}{}^0$ (1.65 vs 1.58 eV). This agreement supports the validity of the static approximation for predicting the average position of optical levels, even in a soft host material with strong dynamics.

C. Effect on nonradiative carrier capture

It has been postulated that the slow and strong variations of $\epsilon_{opt}^{0\rightarrow+}$ can change the level character from shallow to deep, and thus, change how the defect traps free carriers and its effect on nonradiative carrier capture [7]. However, nonradiative carrier capture is an adiabatic process governed by the minimum-energy path between charge states on their respective PES. This process is illustrated in the configurational coordinate (CC) diagram of Fig. 3(b), where the PES for each charge state is plotted along a generalized coordinate, Q, which reflects structural changes during the transition—dominated here by the Pb-Pb distance adjacent to the vacancy, which decreases upon electron capture. After photoexcitation generates excess charge carriers, an electron capture event requires the system, initially at the equilibrium geometry of $V_{Cl}{}^+$, to overcome (or tunnel through) the energy barrier. This barrier originates from the intersection of the PES for the two charge states and is not directly related to the vertical optical transition, which corresponds to instantaneous (vertical) excitations. Thus, the variations of $\epsilon_{opt}^{0\rightarrow+}$ reflect possible optical transitions, as the defect samples the neutral state's PES and reflects the softness of that surface—as illustrated in Fig. 3(c), which shows the range of possible optical transitions during MD trajectories of the neutral and charged vacancies. This broad ϵ_{opt} is also a consequence of strong electron-phonon coupling [53], as quantified by the large Huang-Rhys factor of 24.4. However, these variations do not influence the recombination kinetics, which are largely determined by the capture barriers.

From the configurational coordinate diagram, we find that electron capture by $V_{Cl}{}^0$ is a fast process due to the small energy barrier ($E_n = 0.05$ eV). To complete the

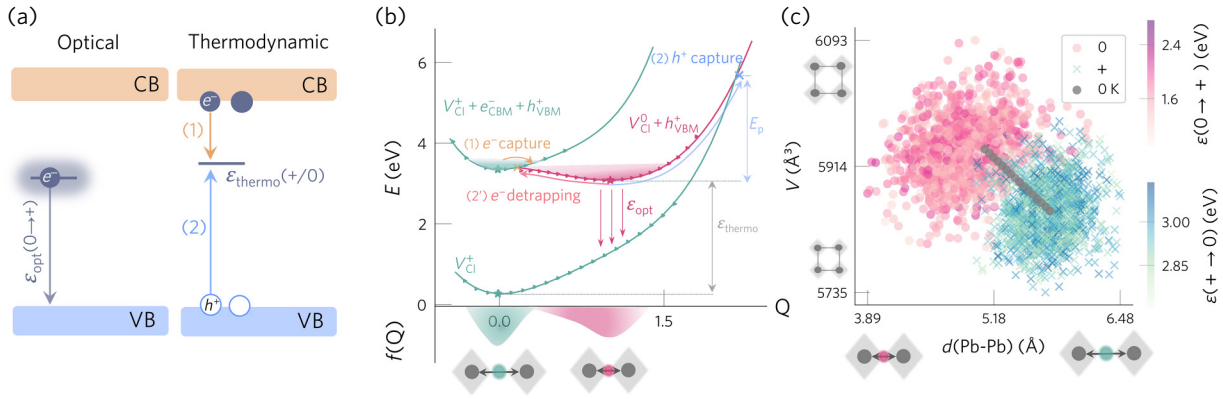

FIG. 3. (a) Schematic energy-level diagram illustrating two types of charge transitions. Optical transition level (ϵ_{opt}) corresponds to an instantaneous (vertical) change in charge state. In contrast, the thermodynamic transition level (ϵ_{thermo}) reflects a thermally activated process, in which the defect relaxes between the equilibrium geometries of the two charge states. (b) CC diagram showing the transition between charge states $V_{Cl}{}^{+}$ and $V_{Cl}{}^{0}$ along a generalized reaction coordinate, Q, that tracks structural relaxation. Green and pink parabolas represent the PESs of $V_{Cl}{}^{+}$ (before and after photoexcitation) and $V_{Cl}{}^{0}$, respectively. Stars mark the equilibrium configurations of each charge state. Vertical pink arrows indicate possible optical transitions $\epsilon_{opt}(0 \rightarrow +)$, while orange and blue arrows depict the adiabatic energy barriers for nonradiative carrier capture. Shaded distributions $f(Q)$ visualize the population of lattice configurations at finite temperature, with their widths reflecting the softness of each PES and the amplitude of vibrational motion around equilibrium. (c) Two-dimensional (2D) map illustrating the structural variation of the defect in both charge states at 300 K (from *NPT* trajectories of the neutral and positively charged states, represented with pink dots and blue crosses, respectively) and comparison with the one-dimensional interpolated path connecting their equilibrium structures (gray dots). Colored bars illustrate the variation of optical levels for each configuration in the trajectories.

nonradiative recombination cycle, however, $V_{Cl}{}^{0}$ would also need to capture a hole. This process is hindered by a high barrier for nonradiative hole capture ($E_p = 2.53$ eV), suggesting that V_{Cl} primarily acts as a reversible electron trap. This is consistent with experimental observations of fast electronic trapping in $CH_3NH_3PbCl_3$ and assigned to V_{Cl} [52].

This behavior aligns with trends observed for halide vacancies in other perovskites. For example, V_I in $CsBI_3$ (B = Pb, Sn, Ge) also exhibits a negligible contribution to nonradiative recombination due to high capture barriers [18]. While halide vacancies seem inactive as nonradiative recombination centers, they could still limit device performance through other mechanisms, like trap-assisted Auger recombination, which could be important for wide band gap semiconductors and/or under high photogenerated carrier densities (e.g., in concentrator photovoltaics) [54,55]. Beyond carrier recombination, in the presence of a high density of (+/0) trap states below the conduction band, the electron quasi-Fermi level will become pinned under illumination. As the concentration of halide vacancies is expected to exceed the photogenerated free carrier density ($\approx 10^{16}$ cm^{-3} under 1 sun conditions), these trap states will act as an electron reservoir. This will limit the build-up of conduction band electrons, suppressing the quasi-Fermi level splitting and capping the achievable open-circuit voltage in a solar cell, which supports observations of voltage improvements under Cl-rich conditions in $CH_3NH_3PbCl_3$ [56,57] and $CsPbCl_3$ [43]. Beyond pinning the Fermi level, the presence of V_{Cl} also promotes ion migration and enhances local lattice distortions (Fig. 11). These factors may underlie the observed performance improvements when the concentration of V_{Cl} is reduced via (post)synthesis treatments.

D. Thermodynamic defect level

Predicting the position of the thermodynamic charge transition level (ϵ_{thermo}) is critical to evaluating the relative stability of defect charge states and how they affect device performance. Within the static configurational coordinate framework, ϵ_{thermo} corresponds to the energy difference between the minima of the potential energy surfaces for different charge states, and it determines the equilibrium occupation of defect levels. Accurate predictions of ϵ_{thermo} are thus not only essential for linking spectroscopic measurements to specific defects, but also for estimating free carrier concentrations and linking synthesis conditions to device performance.

The transition level is often predicted under a static 0 K approximation, where the level is given by the difference in internal energies of the defect at the relaxed geometry of each charge state, e.g., $\epsilon_{thermo} = E^0(R^0) - E^+(R^+) - \epsilon_{VBM}$. However, photovoltaic devices operate at ambient temperatures, and thus, in reality, the level should be calculated from the difference in *Gibbs free energies*, $\epsilon_{thermo}(T) = G^0(R^0) - G^+(R^+) - \epsilon_{VBM}(T)$[58]. To assess the limitations of the 0 K approximation, we compare

both frameworks. In our finite-temperature approach, we evaluate the free energies by averaging the total internal energies sampled from *NPT* molecular dynamics simulations [59] at 300 K and decomposing the entropy into spin and vibrational terms:

$$\epsilon_{\text{thermo}}(+/0) = (\langle E^0\rangle - TS^0) - (\langle E^+\rangle - TS^+) - \epsilon_{\text{VBM}}, \tag{2}$$

where the spin component is determined from the total spin angular momentum, S, $S_{\text{spin}} = k_B \ln(2S+1)$, and the vibrational term is evaluated using the harmonic approximation. Notably, here we neglect the temperature dependence of the bulk band edges [58], since we are focusing on the intrinsic defect effects [e.g., how defect dynamics alone affect $\epsilon_{\text{thermo}}(+/0)$].

To understand the contribution of different factors, we calculate a series of $\epsilon_{\text{thermo}}(+/0)$. In order of increasing sophistication: (i) at 0 K using the constant volume approach (i.e., the lattice volume fixed to its bulk equilibrium values), (ii) at 0 K and constant pressure (i.e., allowing lattice volume relaxation for each charge state), (iii) at 300 K but neglecting entropies, and (iv) at 300 K and including entropies. The first three approaches are in good agreement, with ϵ_{thermo} values of 2.71, 2.76, and 2.79 eV above the VBM, which demonstrates that lattice relaxation and structural motion at 300 K only change ϵ_{thermo} by 50 and 30 meV, respectively. The differences in spin and vibrational entropy are $-T\Delta s_{\text{spin}}(+\to 0) = -18$ meV and $-T\Delta s_{\text{vib}}(+\to 0) = -109$ meV at 300 K, respectively. This results in a cancellation of effects between volume relaxation and vibrational entropy, resulting in good agreement between the static 0 K and dynamic 300 K values (with transition levels of 2.71 and 2.66 eV, respectively). While this cancellation of effects is specific to this defect, the overall small magnitudes of the entropic terms at 300 K validate the accuracy of the static 0 K approximation for predicting charge transition levels for vacancies in halide perovskites at room temperature.

III. DISCUSSION AND CONCLUSIONS

A longstanding question in the defect physics of halide perovskites has been whether their softness and dynamic disorder undermine the conventional static framework used to describe defects in semiconductors. Our results show that this is not the case. We have investigated the dynamics of defect levels in halide perovskites, addressing directly whether the static picture of defect energetics remains valid for these materials. By performing molecular dynamics simulations, we have confirmed previous reports of large fluctuations in the optical level of halide vacancies, which persist even at the HSE + SOC level of theory. These fluctuations originate from large-amplitude thermal vibrations, a consequence of the anharmonic potential energy surfaces of halide vacancies. However, our analysis clarifies that these dynamic variations do not alter thermodynamic transition levels or the potential energy landscape associated with nonradiative recombination kinetics. In fact, such dynamics are a natural manifestation of the defect motion that traditional configurational coordinate diagrams are designed to represent. These thermally activated processes involve slow adiabatic transitions and are governed by energy barriers defined by the shape and crossing of potential energy surfaces of different charge states.

Beyond the impact of defect level oscillations, we have also validated the accuracy of the static 0 K approximation for predicting optical and thermodynamic transition levels for soft and dynamic crystals like halide perovskites. By comparing the properties calculated at 0 and 300 K, we have found the static approximation to give accurate predictions, thereby reducing computational costs without sacrificing accuracy. However, we note that larger deviations could be expected for cases where the room temperature phase is more distinct from the 0 K phase.

By studying the potential energy surface of V_{Cl}, we have found it to be benign for nonradiative recombination, following similar trends to other halide vacancies in $CsBI_3$ (B=Pb, Sn, Ge) and reflecting the defect tolerance of these compounds. While the kinetics are too slow for nonradiative recombination, V_{Cl} could still limit the performance by pinning the electron quasi-Fermi level, promoting ionic migration, or assisting Auger recombination; this explains why Cl-rich conditions result in improved device performance [43,56,57].

Building on the growing application of machine learning for point defects [53,60–62], we have outlined a procedure to calculate optical and thermodynamic defect levels at finite temperatures from molecular dynamics simulations using machine learning force fields. This approach is particularly well-suited to materials that adopt a dynamically stable phase under the device operating conditions (e.g., many perovskites), where the defect structure and levels may vary with the bulk phase. Furthermore, we have demonstrated that a multitask machine learning approach can reduce the cost of generating training data, thereby enabling long molecular dynamics simulations at a high level of theory that would be prohibitive from direct first principles or using a standard MLFF. While the multitask approach reduces overheads, it remains prohibitively expensive for application to all possible point defects and charge states. A pragmatic approach could combine static calculations to characterize the overall defect chemistry, and then apply the dynamic description for the dominant point defects if thermal effects are expected to be important (e.g., high operating temperatures or materials with dynamically stabilized phases). Beyond point defects, the multitask framework can also be applied to study systems with complex configurational landscapes or cases

where simple exchange-correlation functionals fail to provide a reliable description. This includes materials with heavy elements or strong electron correlation (e.g., transition metals), catalytic surfaces where electron localization is key, or where phase stabilities can be influenced by defect formation.

IV. METHODS

A. Density functional theory calculations

All reference total energy and force calculations were performed with density functional theory using the projector-augmented wave method [63], as implemented in the Vienna *ab initio* simulation package (VASP) [64,65]. We used the recommended PBE projector-augmented wave potentials (version 64) for Cs ($5s^2 5p^6 6s^1$), Pb ($5d^{10} 6s^2 6p^2$), and Cl ($3s^2 3p^5$). We used two different exchange-correlation functionals to generate two separate datasets with PBE [66] and HSE + SOC [24,67]. The mixing and screening parameters in the HSE functional were set to $\alpha = 0.375$ and 0.1 $Å^{-1}$ to reproduce the band gap of $CsPbCl_3$ measured at 300 K (3.08 and 2.93 eV [68], respectively), as previously adopted in the literature and found to reproduce experimental lattice parameters and band gaps [69–71] (see Table III). We used an energy cutoff for the plane-wave basis set of 400 eV and a reciprocal space grid of $1 \times 1 \times 2$ for the 80-atom supercell (which is a $2 \times 2 \times 1$ expansion of the orthorhombic unit cell, see Table II).

B. Training of machine learning force fields

We used the structure similarity kernel in VASP to generate the PBE training sets of configurations using its on-the-fly molecular dynamics approach [72–75]. This involved heating runs performed under the *NPT* ensemble with a pressure of 1 atm and from an initial temperature of 100 K up to 30% above our target temperature of 300 K. This procedure was repeated for bulk $CsPbCl_3$ and the two charge states of V_{Cl}. We then used the DIRECT method to select the 10% most diverse structures in each dataset (bulk, V_{Cl}^{0} and V_{Cl}^{+}) and evaluate them at the HSE + SOC level through static self-consistent calculations, resulting in the initial high-fidelity datasets.

After generating the initial training sets with VASP, we trained two separate MACE [51] force fields for V_{Cl}^{0} and V_{Cl}^{+} on the bulk+defect datasets to obtain models with higher accuracy and speed. 10% of the configurations in these datasets were used as validation sets to monitor the loss during training. We used a ScaleShiftMACE model with Ziegler-Biersack-Littmark pair repulsion [76], 2 message passing layers, 128 equivariant messages, correlation order of 3, angular resolution of 3, and cutoff radius of 5.5 Å. The batch size was set to 2 and the Huber loss function was used, with weights of 1, 100, and 100 for the mean square errors in the energies, forces, and stresses, respectively. For the last 20% of the training epochs, the weights were updated to values of 1000, 10, and 100 for energy, force, and stress, respectively—following the recommended strategy of increasing the weight on the energy errors during the final training epochs. The models were trained until the validation loss converged, which required around 150 epochs (see Fig. 8). The reference elemental energies for training the models were defined as the potential energies of isolated Cs, Pb, and Cl atoms, as recommended in the MACE documentation. We used these trained models to run molecular dynamics at 300 K, sample the 10 most diverse configurations, compute their errors via DFT, add the resulting data to our training sets, and retrain the models—repeating this loop until the models achieved our target accuracy.

C. Point defect calculations

The reference DFT defect calculations were set up and analyzed using DOPED [77]. To account for spurious finite-size supercell effects, the Kumagai-Oba [78,79] charge correction scheme was used to calculate E_{corr}, as automated in DOPED. This correction was calculated for the lowest energy (static) defect configuration and applied *a posteriori* to the machine learning force field predictions for the total energies of the charged defects. We note that this is an approximation, as the correction (effective screening) should be configuration dependent; however, there is no general solution for charged defects. The low defect charge and large dielectric screening ensure that this term is small (60 meV for the 80-atom supercell). To identify the defect ground-state structures at 0 K, we used the ShakeNBreak method [80–82]. Carrier capture calculations were performed with the CarrierCapture code [83] using a 160-atom supercell of the orthorhombic structure and reciprocal space grid of $1 \times 1 \times 2$.

D. Validation of machine learning force fields

To generate the test sets, we performed *NPT* molecular dynamics simulations with the models trained at 300 K, running three independent 40 ps runs. We then sampled 30 diverse configurations from these trajectories and performed DFT calculations on them, which were used to calculate the mean absolute errors (MAE) and root mean square error (RMSE) on the predicted properties (energy, forces, and stresses) of each model (see errors in Figs. 4–6 of the Appendix A). The MAE and RMSE for the forces and stresses were calculated component wise, as defined in Ref. [84]. The predicted properties exhibited small absolute errors with no outliers, confirming that the models accurately described the potential energy surfaces of each defect at 300 K. We also validated the models on the transition between charge states by performing a linear interpolation between the equilibrium structures in

each charge and comparing the predicted and reference energies.

E. Molecular dynamics

To model the behavior of the defects at room temperature, we performed *NPT* molecular dynamics with LAMMPS [85] using both 160-atom ($l = 16, 16, 22$ Å) and 640-atom supercells ($l = 31, 31, 22$ Å). The Nosé-Hoover thermostat and barostat were used (1 atm, 300 K, and timestep of 3 fs). The 300 K optical level was evaluated after equilibration and production times of 50 and 200 ps, respectively. The vibrational entropies of V_{Cl}^{0} and V_{Cl}^{+} were calculated using the finite displacement method with PHONOPY using a 160-atom supercell. Given the small magnitude of the change in vibrational harmonic free energies between charge states, we expect anharmonic contributions to lead to even smaller changes, thereby not justifying the additional cost of an anharmonic treatment. These trajectories were analyzed with PYTHON using tools from the ASE [86], PYMATGEN [87–89], DSCRIBE [90,91], UMAP [92], DIRECT [93], MATPLOTLIB [94], and SEABORN [95] packages, and visualized with OVITO [96] and CrystalMaker [97].

ACKNOWLEDGMENTS

We thank Xinwei Wang for initial calculations on the electronic structure and carrier capture behavior of V_{Cl} in $CsPbCl_3$. I.M.-L. thanks Imperial College London (ICL) for funding a President's PhD scholarship. We are grateful to the UK Materials and Molecular Modelling Hub for computational resources, which is partially funded by the EPSRC (Grants No. EP/P020194/1 and No. EP/T022213/1). This work used the ARCHER2 UK National Supercomputing Service [98] via our membership of the UK's HEC Materials Chemistry Consortium, funded by the EPSRC (Grant No. EP/L000202). We acknowledge the ICL High Performance Computing services for computational resources.

I.M.-L. contributed to the investigation, methodology, formal analysis, writing—original draft, review, and editing. A.W. contributed to conceptualization, project administration, supervision, and writing—review and editing. These author contributions are defined according to the CRediT contributor roles taxonomy.

The authors declare no competing interests.

DATA AVAILABILITY

The data that support the findings of this article are openly available [99].

APPENDIX A: VALIDATION OF MACHINE LEARNING FORCE FIELDS

We validated the performance of our machine learning force fields by comparing their predicted energies, forces, and stresses on a test set of configurations with the ground-truth DFT values. As shown in Figs. 4 and 5, the predicted properties exhibit small absolute errors with no outliers, demonstrating that the models accurately describe the potential energy surfaces of each system at 300 K. Beyond validating the models on configurations sampled from MD, we also validated them on the 0 K potential energy surfaces, which we generated by performing a linear interpolation between the ground-state structures of each charge state (Fig. 6).

APPENDIX B: TRAINING OF MACHINE LEARNING FORCE FIELDS

The number of configurations for training the MLFFs is shown in Table I. We further analysed the learning efficiency of the multi-task models with a learning curve (Fig. 7), where we plot the validation errors for models trained

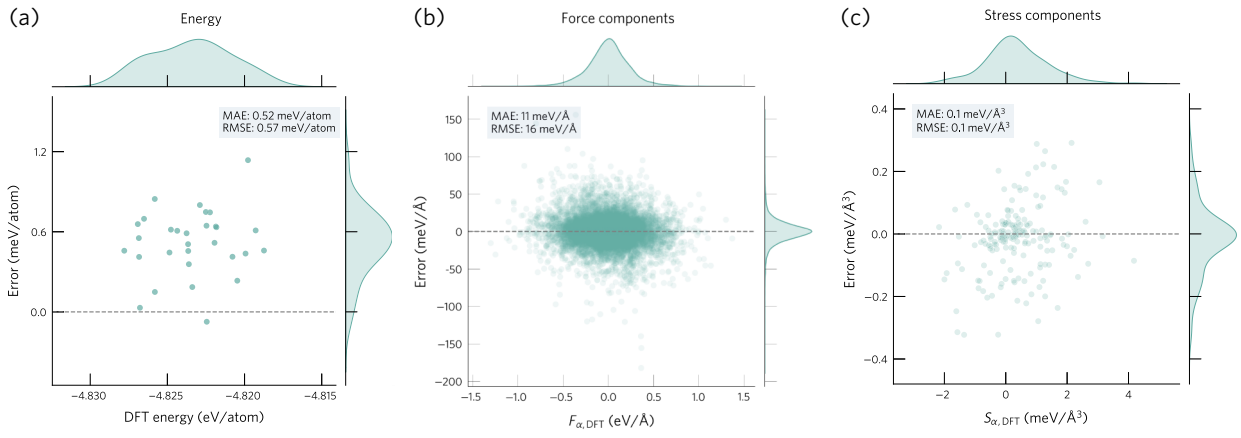

FIG. 4. Distribution of absolute errors in the predicted energies, forces, and stresses for the test set of V_{Cl}^{0}. These defect structures were based on 160-atom supercells.

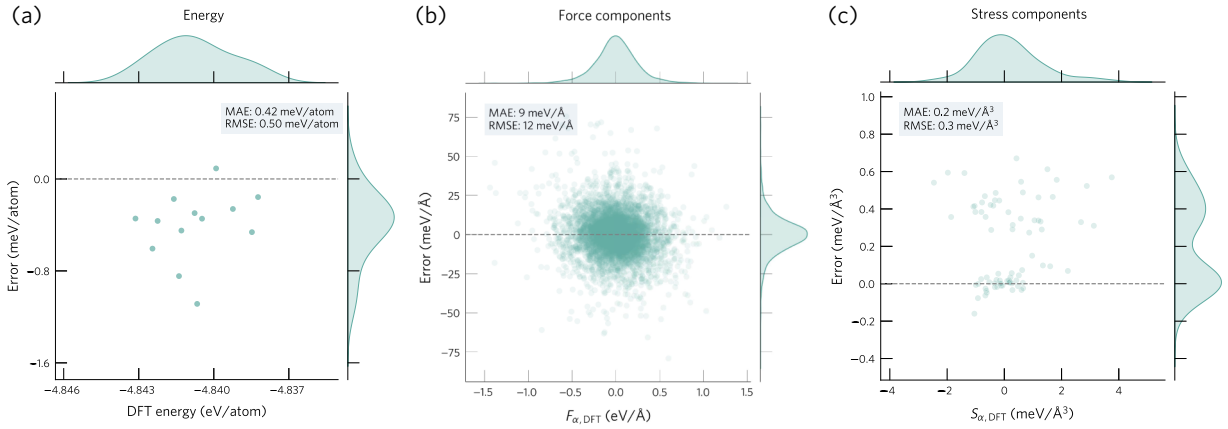

FIG. 5. Distribution of absolute errors in the predicted energies, forces, and stresses for the test set of V_{Cl}^{+}.

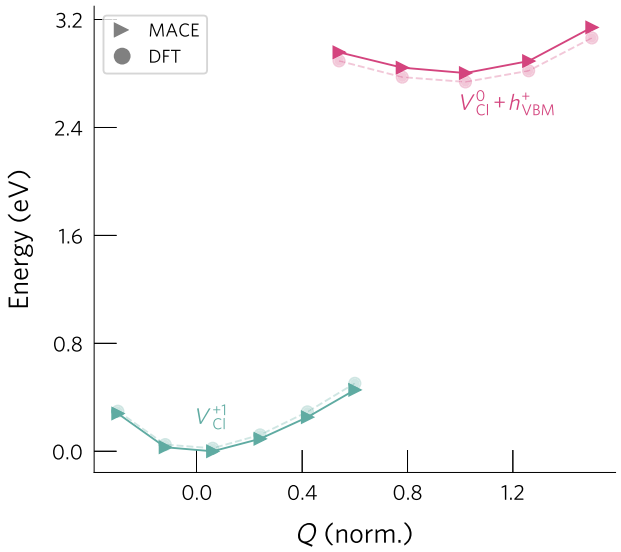

FIG. 6. Comparison of energies predicted by DFT (HSE + SOC) and the MACE model for the 0 K potential energy surfaces of V_{Cl}^{+} and V_{Cl}^{0}. Structures were generated through a linear interpolation between the ground-state configurations of each charge state.

on a dataset of increasing size. This analysis shows that accurate models can be obtained with moderate training datasets comprising 270 and 25 structures for the PBE and HSE+SOC fidelity levels. The training curves plotting the training and validation errors for the predicted energies, forces and stresses are depicted in Fig. 8.

TABLE I. Number of configurations in the training and validation datasets for the multihead force field.

Charge	Functional	Train	Validation
V_{Cl}^{0}	PBE	1576	174
	HSE + SOC	221	31
V_{Cl}^{+}	PBE	1448	158
	HSE + SOC	171	21

APPENDIX C: ADDITIONAL COMPUTATIONAL DETAILS

1. Supercells

Different supercell expansions were used for training the model and production simulations (Tables II and III). Initially, we started generating the training set using a

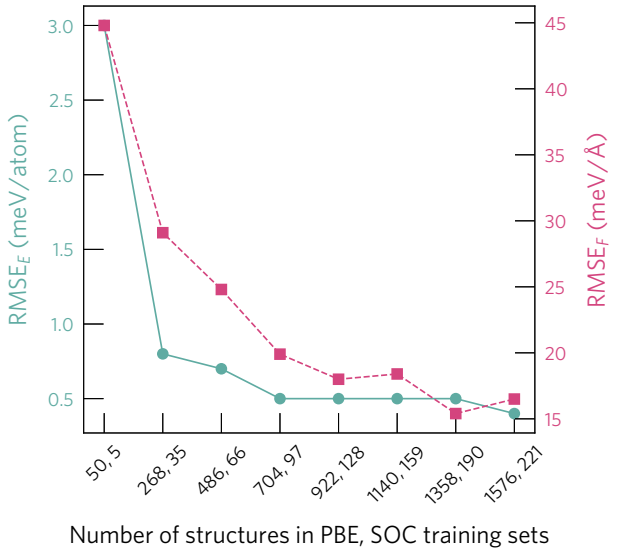

FIG. 7. Learning curve for V_{Cl}^{0}, plotting the validation root mean square errors for the energy and forces of models trained on datasets of increasing size. Same validation dataset was used for monitoring the loss during training for all models. For each data split, the configurations were sampled using the DIRECT method [93] to maximize diversity. These results illustrate that high accuracies ($\mathrm{RMSE}_E < 1$ meV/atom) can be obtained with datasets of (700, 100) configurations for the PBE and HSE + SOC fidelity levels.

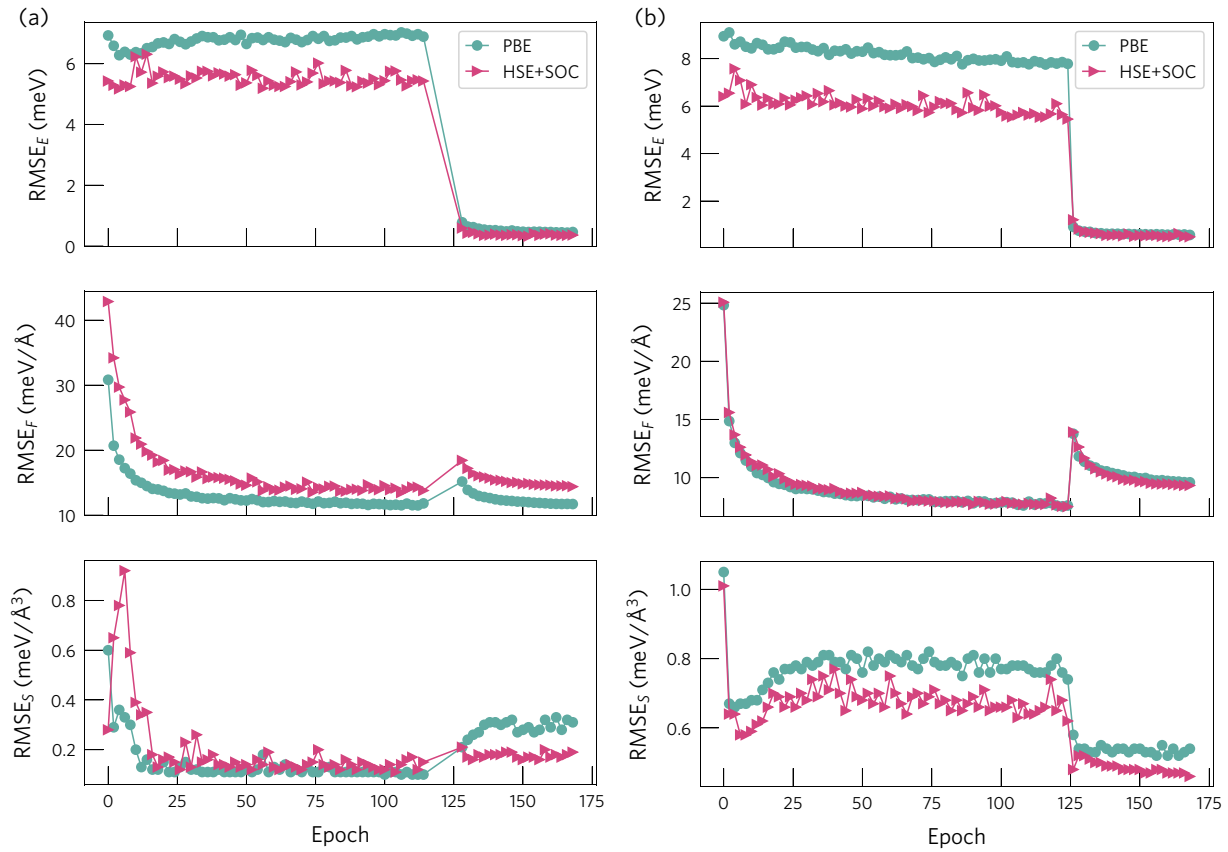

FIG. 8. Training curves for $V_{Cl}{}^0$ (a) and $V_{Cl}{}^+$ (b), showing the validation RMSEs in the energies, forces, and stresses during training for the PBE and HSE + SOC heads. Discontinuity that occurs in the last 20% of epochs arises from the adaptive loss function, which increases the weight of the energy term.

$2 \times 2 \times 1$ supercell expansion of the orthorhombic unit cell (80 atoms). However, we found that when Pb-V_{Cl}-Pb was aligned along the shortest direction ($c = 11.2$ Å), the length of the Pb-Pb interaction was constrained by the supercell size (as illustrated in Fig. 9). To solve this, we expanded the training set using a $2 \times 2 \times 2$ expansion ($c = 22.5$ Å, 160 atoms), which did not limit the Pb-Pb distance. To validate the MLFFs, we used the 160-atom supercell, which allowed us to validate the scalability of the model while remaining tractable for HSE + SOC calculations. For similar reasons, we used a 160-atom supercell for the analysis of the configurational coordinate diagram, which required HSE + SOC calculations to validate the model predictions. For consistency, the 2D map illustrating the structural variation of the vacancy in both charge states [Fig. 3(c)] was also performed using the 160-atom supercell. Similarly, we evaluated the optical and thermodynamic levels using the same supercell. However, to ensure that finite size effects did not overestimate the fluctuations of ϵ_{opt}, we also performed molecular dynamics simulations using a 640-atom supercell ($4 \times 4 \times 2$ expansion). The fluctuations of the optical level were further analysed with a Fourier transform to identify the vibrational frequencies that lead to the strongest fluctuations (Fig. 10). Figure 11 compares the tilt angles in a pristine and defect containing supercell, illustrating that higher tilt angles are observed when the vacancy is present. The variation of the pseudo-cubic lattice parameters for this supercell size ($4 \times 4 \times 2$ expansion) at 300 K is shown in Fig. 12.

TABLE II. Supercell expansions of the orthorhombic unit cell used for training and validating the models and for production simulations (e.g., CC diagram, calculation of ϵ_{thermo}, and ϵ_{opt}).

Expansion	Lattice (Å)	No. atoms	Usage
$2 \times 2 \times 1$	16.1, 15.7, 11.2	80	Training
$2 \times 2 \times 2$	16.1, 15.7, 22.5	160	Training, Validation, CC, MD, ϵ_{thermo}
$4 \times 4 \times 2$	32.2, 31.5, 22.5	640	MD (Variation of ϵ_{opt})

TABLE III. Comparison of computed and measured values (in parentheses) for the lattice parameters [28] and the band gap [68] of orthorhombic $CsPbCl_3$.

a	*b*	*c*	Band gap (eV)
7.88 (7.90)	8.05 (7.90)	11.28 (11.25)	3.08 (2.93)

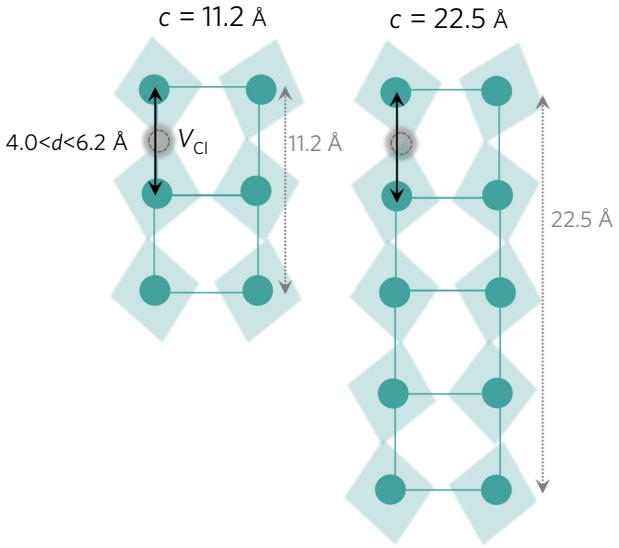

FIG. 9. Effect of supercell size on the Pb-V_{Cl}-Pb distance. $2 \times 2 \times 1$ expansion of the orthorhombic cell (16.1, 15.7, 11.2 Å) constrains the Pb-Pb distance for configurations with a high Pb-Pb separation along the *c* direction (e.g., $d > 5.5$ Å) due to limited supercell dimensions along *c*.

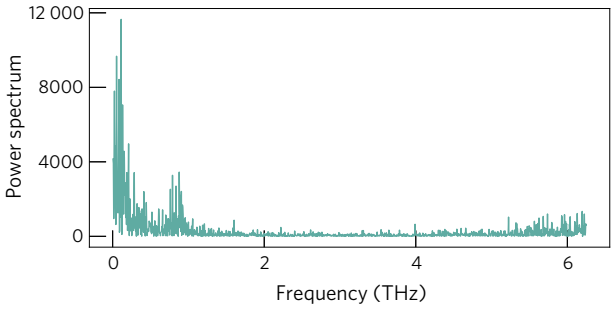

FIG. 10. Fourier transform of the time-dependent optical level during the molecular dynamics trajectory of V_{Cl}^0, illustrating that the strongest peaks occur for low frequencies (≈0.12 THz).

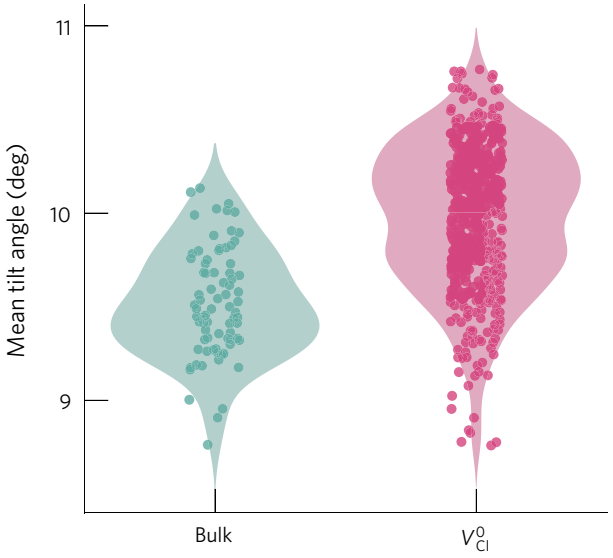

FIG. 11. Mean tilt angles of bulk $CsPbCl_3$ and V_{Cl}^0, in green and pink, respectively, during an *NPT* trajectory at 300 K, illustrating the higher tilt angles of the supercell containing the Cl vacancy.

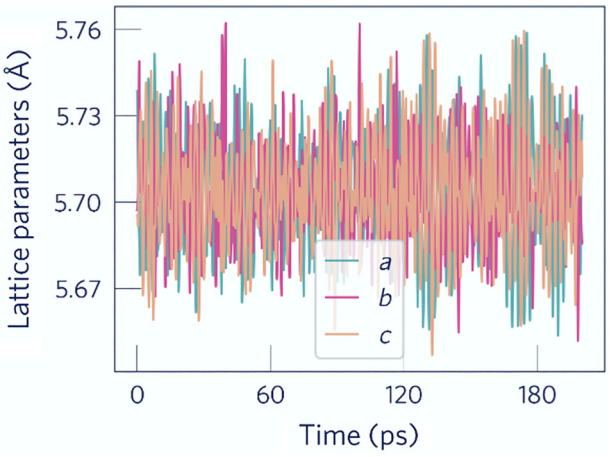

FIG. 12. Variation of pseudocubic lattice parameters during a MD simulation (*NPT*, 300 K, 1 atm, 200 ps) of V_{Cl}^0 in a 640-atom supercell ($4 \times 4 \times 2$ expansion of the orthorhombic unit cell).

[1] I. Mosquera-Lois, S. R. Kavanagh, J. Klarbring, K. Tolborg, and A. Walsh, Imperfections are not 0 K: Free energy of point defects in crystals, Chem. Soc. Rev. **52**, 5812 (2023).

[2] D. A. Egger, A. Bera, D. Cahen, G. Hodes, T. Kirchartz, L. Kronik, R. Lovrincic, A. M. Rappe, D. R. Reichman, and O. Yaffe, What remains unexplained about the properties of halide perovskites? Adv. Mater. **30**, 1800691 (2018).

[3] M. Zacharias, G. Volonakis, F. Giustino, and J. Even, Anharmonic electron-phonon coupling in ultrasoft and

locally disordered perovskites, npj Comput. Mater. **9**, 153 (2023).
[4] M. Dubajic, J. R. Neilson, J. Klarbring, X. Liang, S. A. Bird, K. C. Rule, J. E. Auckett, T. A. Selby, G. Tumen-Ulzii, Y. Lu, *et al.*, Dynamic nanodomains dictate macroscopic properties in lead halide perovskites, Nat. Nanotech. **20**, 755 (2025).
[5] S. A. Seidl, X. Zhu, G. Reuveni, S. Aharon, C. Gehrmann, S. Caicedo-Dávila, O. Yaffe, and D. A. Egger, Anharmonic fluctuations govern the band gap of halide perovskites, Phys. Rev. Mater. **7**, L092401 (2023).
[6] I. Mosquera-Lois, Y.-T. Huang, H. Lohan, J. Ye, A. Walsh, and R. L. Z. Hoye, Multifaceted nature of defect tolerance in halide perovskites and emerging semiconductors, Nat. Rev. Chem. **9**, 287 (2025).
[7] A. V. Cohen, D. A. Egger, A. M. Rappe, and L. Kronik, Breakdown of the static picture of defect energetics in halide perovskites: The case of the Br vacancy in $CsPbBr_3$, J. Phys. Chem. Lett. **10**, 4490 (2019).
[8] J. Ran, B. Wang, Y. Wu, D. Liu, C. Mora Perez, A. S. Vasenko, and O. V. Prezhdo, Halide vacancies create no charge traps on lead halide perovskite surfaces but can generate deep traps in the bulk, J. Phys. Chem. Lett. **14**, 6028 (2023).
[9] L. Deng, J. Ran, B. Wang, A. Boziki, A. Tkatchenko, J. Jiang, and O. V. Prezhdo, Strong dependence of point defect properties in metal halide perovskites on description of van der Waals interaction, J. Phys. Chem. Lett. **15**, 10465 (2024).
[10] 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).
[11] Y. Liu, W.-H. Fang, and R. Long, Significant impact of defect fluctuation on charge dynamics in $CsPbI_3$: A study combining machine learning with quantum dynamics, J. Phys. Chem. Lett. **15**, 3764 (2024).
[12] Y. Shi, W. Chu, L. Zhang, B. Wang, W. A. Saidi, J. Zhao, and O. V. Prezhdo, Band gap narrowing in lead-halide perovskites by dynamic defect self-doping for enhanced light absorption and energy upconversion, Chem. Mater. **37**, 655 (2025).
[13] W. Chu, Q. Zheng, O. V. Prezhdo, J. Zhao, and W. A. Saidi, Low-frequency lattice phonons in halide perovskites explain high defect tolerance toward electron-hole recombination, Sci. Adv. **6**, eaaw7453 (2020).
[14] W. Li, Y.-Y. Sun, L. Li, Z. Zhou, J. Tang, and O. V. Prezhdo, Control of charge recombination in perovskites by oxidation state of halide vacancy, J. Am. Chem. Soc. **140**, 15753 (2018).
[15] B. Wang, W. Chu, Y. Wu, D. Casanova, W. A. Saidi, and O. V. Prezhdo, Electron-volt fluctuation of defect levels in metal halide perovskites on a 100 ps time scale, J. Phys. Chem. Lett. **13**, 5946 (2022).
[16] J. Kang, Effects of band edge positions on defect structure in lead halide perovskites: A case study on the Br vacancy in $CsPbBr_3$, Phys. Rev. Mater. **4**, 085405 (2020).
[17] C. Ming, H. Wang, D. West, S. Zhang, and Y.-Y. Sun, Defect tolerance in $CsPbI_3$: Reconstruction of the potential energy landscape and band degeneracy in spin–orbit coupling, J. Mater. Chem. A **10**, 3018 (2022).
[18] J. Zhang, X. Zhang, M. E. Turiansky, and C. G. Van de Walle, Iodine vacancies do not cause nonradiative recombination in halide perovskites, PRX Energy **2**, 013008 (2023).
[19] J. Kang, J. Li, and S.-H. Wei, Atomic-scale understanding on the physics and control of intrinsic point defects in lead halide perovskites, Appl. Phys. Rev. **8**, 031302 (2021).
[20] D. Meggiolaro, S. G. Motti, E. Mosconi, A. J. Barker, J. Ball, C. A. R. Perini, F. Deschler, A. Petrozza, and F. D. Angelis, Iodine chemistry determines the defect tolerance of lead-halide perovskites, Energy Environ. Sci. **11**, 702 (2018).
[21] D. Meggiolaro and F. De Angelis, First-principles modeling of defects in lead halide perovskites: Best practices and open issues, ACS Energy Lett. **3**, 2206 (2018).
[22] M.-H. Du, Density functional calculations of native defects in $CH_3NH_3PbI_3$: Effects of spin–orbit coupling and self-interaction error, J. Phys. Chem. Lett. **6**, 1461 (2015).
[23] X. Zhang, M. E. Turiansky, J.-X. Shen, and C. G. Van de Walle, Defect tolerance in halide perovskites: A first-principles perspective, J. Appl. Phys. **131**, 090901 (2022).
[24] J. Heyd, G. E. Scuseria, and M. Ernzerhof, Hybrid functionals based on a screened coulomb potential, J. Chem. Phys. **118**, 8207 (2003).
[25] D. Chen, Y. Ba, M. Deng, W. Zhu, W. Chai, H. Xi, D. Chen, J. Zhang, C. Zhang, and Y. Hao, High-purity, thick $CsPbCl_3$ films toward selective ultraviolet-harvesting visibly transparent photovoltaics, ACS Appl. Energy Mater. **4**, 12121 (2021).
[26] J. Tong, Q. Jiang, F. Zhang, S. B. Kang, D. H. Kim, and K. Zhu, Wide-bandgap metal halide perovskites for tandem solar cells, ACS Energy Lett. **6**, 232 (2021).
[27] R. Li, B. Chen, N. Ren, P. Wang, B. Shi, Q. Xu, H. Zhao, W. Han, Z. Zhu, J. Liu, Q. Huang, D. Zhang, Y. Zhao, and X. Zhang, $CsPbCl_3$-cluster-widened bandgap and inhibited phase segregation in a wide-bandgap perovskite and its application to niox-based perovskite/silicon tandem solar cells, Adv. Mater. **34**, 2201451 (2022).
[28] Y. He, C. C. Stoumpos, I. Hadar, Z. Luo, K. M. McCall, Z. Liu, D. Y. Chung, B. W. Wessels, and M. G. Kanatzidis, Demonstration of energy-resolved γ-ray detection at room temperature by the $CsPbCl_3$ perovskite semiconductor, J. Am. Chem. Soc. **143**, 2068 (2021).
[29] S. Hirotsu, Experimental studies of structural phase transitions in $CsPbCl_3$, J. Phys. Soc. Jpn. **31**, 552 (1971).
[30] M. I. Cohen, K. F. Young, T. Chang, and W. S. Brower Jr., Phase transitions in $CsPbCl_3$, J. Appl. Phys. **42**, 5267 (1971).
[31] T. Haeger, M. Ketterer, J. Bahr, N. Pourdavoud, M. Runkel, R. Heiderhoff, and T. Riedl, Thermal properties of $CsPbCl_3$ thin films across phase transitions, J. Phys. Mater. **3**, 024004 (2020).
[32] S. Dongre S, E. E. Siddharthan, R. Thapa, S. Ramu, and R. G. Balakrishna, Dual vacancy passivation in $CsPbCl_3$ perovskite nanocrystals: Implications on optoelectronic applications, ACS Appl. Nano Mater. **6**, 13227 (2023).
[33] D. Chen, J. Ding, Q. Tan, P. Yang, Y. Liu, and Q. Wang, Roles of defects in perovskite $CsPbX_3$ (X=I, Br, Cl): A first-principles investigation, Phys. Scr. **99**, 115911 (2024).

[34] D. E. Sommer, D. R. Gamelin, and S. T. Dunham, Defect formation in Yb-doped $CsPbCl_3$ from first principles with implications for quantum cutting, Phys. Rev. Mater. **6**, 025404 (2022).

[35] J.-P. Ma, Y.-M. Chen, L.-M. Zhang, S.-Q. Guo, J.-D. Liu, H. Li, B.-J. Ye, Z.-Y. Li, Y. Zhou, B.-B. Zhang, O. M. Bakr, J.-Y. Zhang, and H.-T. Sun, Insights into the local structure of dopants, doping efficiency, and luminescence properties of lanthanide-doped $CsPbCl_3$ perovskite nanocrystals, J. Mater. Chem. C **7**, 3037 (2019).

[36] X. Zhu, L. Ge, Y. Wang, M. Li, R. Zhang, M. Xu, Z. Zhao, W. Lv, and R. Chen, Recent advances in enhancing and enriching the optical properties of Cl-based $CsPbX_3$ nanocrystals, Adv. Opt. Mater. **9**, 2100058 (2021).

[37] S. Seth, T. Ahmed, A. De, and A. Samanta, Tackling the defects, stability, and photoluminescence of $CsPbX_3$ perovskite nanocrystals, ACS Energy Lett. **4**, 1610 (2019).

[38] N. Fiuza-Maneiro, J. Ye, S. K. Sharma, S. Chakraborty, S. Gómez-Graña, R. L. Z. Hoye, and L. Polavarapu, Unlocking brightness in $CsPbCl_3$ perovskite nanocrystals: Screening ligands and metal halides for effective deep trap passivation, ACS Energy Lett. **10**, 1623 (2025).

[39] X. Zheng, Y. Hou, H.-T. Sun, O. F. Mohammed, E. H. Sargent, and O. M. Bakr, Reducing defects in halide perovskite nanocrystals for light-emitting applications, J. Phys. Chem. Lett. **10**, 2629 (2019).

[40] L. Fei, L. Yang, P. Li, and J. Ma, Unlocking near 100% PLQY: The impact of post-treatment on emission efficiency of Mn^{2+}: $CsPbCl_3$ nanocrystals, Adv. Opt. Mater. **13**, 2402826 (2025).

[41] J. A. Peters, Z. Liu, M. C. De Siena, M. G. Kanatzidis, and B. W. Wessels, Defect levels in $CsPbCl_3$ single crystals determined by thermally stimulated current spectroscopy, J. Appl. Phys. **132**, 035101 (2022).

[42] J. Ye, M. M. Byranvand, C. O. Martínez, R. L. Z. Hoye, M. Saliba, and L. Polavarapu, Defect passivation in lead-halide perovskite nanocrystals and thin films: Toward efficient LEDs and solar cells, Angew. Chem. Int. Ed. **60**, 21636 (2021).

[43] S. Cheng, X. Zheng, Z. Hou, R. Hu, S. Jiang, S. Xi, G. Wen, and X. Liu, Passivating the vacancy defects of $CsPbCl_3$ polycrystalline films by a Cl-containing ionic liquid for self-powered, charge-transport-layer-free UV photodetectors, J. Mater. Chem. C **10**, 5693 (2022).

[44] Y. Zhang, X. Cheng, D. Tu, Z. Gong, R. Li, Y. Yang, W. Zheng, J. Xu, S. Deng, and X. Chen, Engineering the bandgap and surface structure of $CsPbCl_3$ nanocrystals to achieve efficient ultraviolet luminescence, Angew. Chem. Int. Ed. **60**, 9693 (2021).

[45] G. H. Ahmed, J. K. El-Demellawi, J. Yin, J. Pan, D. B. Velusamy, M. N. Hedhili, E. Alarousu, O. M. Bakr, H. N. Alshareef, and O. F. Mohammed, Giant photoluminescence enhancement in $CsPbCl_3$ perovskite nanocrystals by simultaneous dual-surface passivation, ACS Energy Lett. **3**, 2301 (2018).

[46] Z.-J. Yong, S.-Q. Guo, J.-P. Ma, J.-Y. Zhang, Z.-Y. Li, Y.-M. Chen, B.-B. Zhang, Y. Zhou, J. Shu, J.-L. Gu, L.-R. Zheng, O. M. Bakr, and H.-T. Sun, Doping-enhanced short-range order of perovskite nanocrystals for near-unity violet luminescence quantum yield, J. Am. Chem. Soc. **140**, 9942 (2018).

[47] Y. Li, Q. Liu, X. Liu, J. Feng, L. He, H. Li, C. Li, and H. Zhang, Simultaneous enhancement of photoluminescence and stability of $CsPbCl_3$ perovskite enabled by titanium ion dopant, J. Phys. Chem. Lett. **12**, 10746 (2021).

[48] S. Jiang, Z. Hou, X. Zheng, Q. Wu, X. Yang, W. Cai, Y. Yi, S. Cheng, G. Wen, and X. Liu, Ultrafast self-powered $CsPbCl_3$ ultraviolet photodetectors with choline chloride for surface passivation and charge transport regulation, J. Mater. Chem. C **11**, 5667 (2023).

[49] F. Brivio, K. T. Butler, A. Walsh, and M. van Schilfgaarde, Relativistic quasiparticle self-consistent electronic structure of hybrid halide perovskite photovoltaic absorbers, Phys. Rev. B **89**, 155204 (2014).

[50] P. Umari, E. Mosconi, and F. De Angelis, Relativistic *GW* calculations on $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ perovskites for solar cell applications, Sci. Rep. **4**, 4467 (2014).

[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, Adv. Neural. Inf. Process. Syst. **35**, 11423 (2022).

[52] J. Wang, Y.-Q. Sun, and D. Shi, Fast electronic trapping and de-trapping by mid-gap states in $CH_3NH_3PbCl_3$ single crystal, Appl. Phys. Lett. **122**, 041101 (2023).

[53] M. E. Turiansky, J. L. Lyons, and N. Bernstein, Machine learning phonon spectra for fast and accurate optical lineshapes of defects, ArXiv:2508.09113.

[54] T. Kirchartz, High open-circuit voltages in lead-halide perovskite solar cells: Experiment, theory and open questions, Phil. Trans. R. Soc. A **377**, 20180286 (2019).

[55] J.-X. Shen, X. Zhang, S. Das, E. Kioupakis, and C. G. Van de Walle, Unexpectedly strong auger recombination in halide perovskites, Adv. Funct. Mater. **8**, 1801027 (2018).

[56] W. Zia, M. Malekshahi Byranvand, T. Rudolph, M. Rai, M. Kot, C. Das, M. Kedia, M. Zohdi, W. Zuo, V. Yeddu, M. I. Saidaminov, J. I. Flege, T. Kirchartz, and M. Saliba, MAPbCl3 light absorber for highest voltage perovskite solar cells, ACS Energy Lett. **9**, 1017 (2024).

[57] Y. Zhao, C. Wang, T. Ma, L. Zhou, Z. Wu, H. Wang, C. Chen, Z. Yu, W. Sun, A. Wang, H. Huang, B. Zou, D. Zhao, and X. Li, Reduced 0.418 V V_{OC}-deficit of 1.73 eV wide-bandgap perovskite solar cells assisted by dual chlorides for efficient all-perovskite tandems, Energy Environ. Sci. **16**, 2080 (2023).

[58] S. Qiao, Y.-N. Wu, X. Yan, B. Monserrat, S.-H. Wei, and B. Huang, Temperature effect on charge-state transition levels of defects in semiconductors, Phys. Rev. B **105**, 115201 (2022).

[59] X. Wu, C. Ming, J. Shi, H. Wang, D. West, S. Zhang, and Y.-Y. Sun, Defects in statically unstable solids: The case for cubic perovskite α-$CsPbI_3$, Chin. Phys. Lett. **39**, 046101 (2022).

[60] I. Mosquera-Lois, J. Klarbring, and A. Walsh, Point defect formation at finite temperatures with machine-learning force fields, Chem. Sci. **16**, 8878 (2025).

[61] I. Mosquera-Lois, S. R. Kavanagh, A. M. Ganose, and A. Walsh, Machine-learning structural reconstructions for accelerated point defect calculations, npj Comput. Mater. **10**, 1 (2024).

[62] S. R. Kavanagh, Identifying split vacancy defects with machine-learned foundation models and electrostatics, J. Phys. Energy **7**, 045002 (2025).

[63] G. Kresse and J. Furthmüller, Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set, Comput. Mater. Sci. **6**, 15 (1996).

[64] G. Kresse and J. Hafner, *Ab initio* molecular dynamics for liquid metals, Phys. Rev. B **47**, 558 (1993).

[65] G. Kresse and J. Hafner, *Ab initio* molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium, Phys. Rev. B **49**, 14251 (1994).

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

[67] S. Steiner, S. Khmelevskyi, M. Marsmann, and G. Kresse, Calculation of the magnetic anisotropy with projected-augmented-wave methodology and the case study of disordered $Fe_{1-x}Co_x$ alloys, Phys. Rev. B **93**, 224425 (2016).

[68] K. O. Mastej, B. Batnaran, A.-P. M. Reponen, Z. A. VanOrman, K. Banger, M. A. Hayward, V. L. Deringer, and S. Feldmann, An experimental data library for the full $CsPb(Cl_xBr_{1-x})_3$ compositional series, Chem. Commun. **61**, 6146 (2025).

[69] J. L. Lyons and M. W. Swift, Trends for acceptor dopants in lead halide perovskites, J. Phys. Chem. C **127**, 12735 (2023).

[70] J. L. Lyons, Effective donor dopants for lead halide perovskites, Chem. Mater. **33**, 6200 (2021).

[71] M. W. Swift and J. L. Lyons, Deep levels in cesium lead bromide from native defects and hydrogen, J. Mater. Chem. A **9**, 7491 (2021).

[72] 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).

[73] R. Jinnouchi, F. Karsai, and G. Kresse, On-the-fly machine learning force field generation: Application to melting points, Phys. Rev. B **100**, 014105 (2019).

[74] R. Jinnouchi, K. Miwa, F. Karsai, G. Kresse, and R. Asahi, On-the-fly active learning of interatomic potentials for large-scale atomistic simulations, J. Phys. Chem. Lett. **11**, 6946 (2020).

[75] P. Liu, C. Verdi, F. Karsai, and G. Kresse, Phase transitions of zirconia: Machine-learned force fields beyond density functional theory, Phys. Rev. B **105**, L060102 (2022).

[76] J. F. Ziegler and J. P. Biersack, in *Treatise on Heavy-Ion Science: Volume 6: Astrophysics, Chemistry, and Condensed Matter*, edited by D. A. Bromley (Springer US, Boston, MA, 1985), p. 93.

[77] S. R. Kavanagh, A. G. Squires, A. Nicolson, I. Mosquera-Lois, A. M. Ganose, B. Zhu, K. Brlec, A. Walsh, and D. O. Scanlon, Doped: PYTHON toolkit for robust and repeatable charged defect supercell calculations, J. Open Source Softw. **9**, 6433 (2024).

[78] Y. Kumagai and F. Oba, Electrostatics-based finite-size corrections for first-principles point defect calculations, Phys. Rev. B **89**, 195205 (2014).

[79] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, Fully *ab initio* finite-size corrections for charged-defect supercell calculations, Phys. Rev. Lett. **102**, 016402 (2009).

[80] I. Mosquera-Lois and S. R. Kavanagh, In search of hidden defects, Matter **4**, 2602 (2021).

[81] I. Mosquera-Lois, S. R. Kavanagh, A. Walsh, and D. O. Scanlon, Identifying the ground state structures of point defects in solids, npj Comput. Mater. **9**, 1 (2023).

[82] I. Mosquera-Lois, S. R. Kavanagh, A. Walsh, and D. O. Scanlon, ShakeNBreak: Navigating the defect configurational landscape, J. Open Source Softw. **7**, 4817 (2022).

[83] S. Kim, S. N. Hood, P. v. Gerwen, L. D. Whalley, and A. Walsh, CarrierCapture.jl: Anharmonic carrier capture, J. Open Source Softw. **5**, 2102 (2020).

[84] J. D. Morrow, J. L. A. Gardner, and V. L. Deringer, How to validate machine-learned interatomic potentials, J. Chem. Phys. **158**, 121501 (2023).

[85] 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, and S. J. Plimpton, LAMMPS—a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, Comput. Phys. Commun. **271**, 108171 (2022).

[86] A. H. Larsen, *et al.*, The atomic simulation environment—a PYTHON library for working with atoms, J. Condens. Matter Phys. **29**, 273002 (2017).

[87] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, PYTHON materials genomics (pymatgen): A robust, open-source PYTHON library for materials analysis, Comput. Mater. Sci. **68**, 314 (2013).

[88] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, Commentary: The materials project: A materials genome approach to accelerating materials innovation, APL Mater. **1**, 011002 (2013).

[89] S. P. Ong, S. Cholia, A. Jain, M. Brafman, D. Gunter, G. Ceder, and K. A. Persson, The materials application programming interface (API): A simple, flexible and efficient API for materials data based on representational state transfer (REST) principles, Comput. Mater. Sci. **97**, 209 (2015).

[90] L. Himanen, M. O. J. Jäger, E. V. Morooka, F. Federici Canova, Y. S. Ranawat, D. Z. Gao, P. Rinke, and A. S. Foster, DScribe: Library of descriptors for machine learning in materials science, Comput. Phys. Commun. **247**, 106949 (2020).

[91] J. Laakso, L. Himanen, H. Homm, E. V. Morooka, M. O. Jäger, M. Todorović, and P. Rinke, Updates to the dscribe library: New descriptors and derivatives, J. Chem. Phys. **158**, 234802 (2023).

[92] L. McInnes, J. Healy, N. Saul, and L. Großberger, UMAP: Uniform manifold approximation and projection, J. Open Source Softw. **3**, 861 (2018).

[93] J. Qi, T. W. Ko, B. C. Wood, T. A. Pham, and S. P. Ong, Robust training of machine learning interatomic potentials with dimensionality reduction and stratified sampling, npj Comput. Mater. **10**, 43 (2024).

[94] J. D. Hunter, MATPLOTLIB: A 2D graphics environment, Comput. Sci. Eng. **9**, 90 (2007).
[95] M. L. Waskom, SEABORN: Statistical data visualization, J. Open Source Softw. **6**, 3021 (2021).
[96] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO–the open visualization tool, Model. Simul. Mater. Sci. Eng. **18**, 015012 (2010).
[97] CrystalMaker®: A crystal and molecular structures program for Mac and Windows. CrystalMaker Software Ltd, Oxford, England (www.crystalmaker.com).
[98] https://www.archer2.ac.uk
[99] I. Mosquera-Lois and A. Walsh, Dynamic vacancy levels in $CsPbCl_3$ obey equilibrium defect thermodynamics, https://doi.org/10.5281/zenodo.17208040, 2025.