Journal Name

Mapping Temperature–Dependent Energy–Structure–Property Relationships for Solid Solutions of Inorganic Halide Perovskites

Jack Yang[*a]

Formation of solid solutions with complex compositions has been exhaustively adopted in material research for improving chemical and physical properties. This is also the true for halide perovskites, in the hope of further enhancing their stabilities and reducing the toxicities in lead–containing compounds. Replacement of lead with tin, even partially, is a route to achieve the latter goal. However, this has to be compromised with reduction in band gaps as well as structural stabilities. High–throughput statistical samplings over different configurations for random solid solutions have played pivotal roles in guiding the chemical designs of halide perovskite with better stabilities while retaining high photovoltaic efficiencies, but it remains challenging to intuitively and comprehensively understand the intriguing energy–structure–property (ESP) relationships in solid solutions encompassing multiple degrees–of–freedoms. In this work, first–principle dynamic and electronic structure calculations are performed across 51 different compositions of $Cs(Pb_xSn_{1-x})X_3$ (X=Cl, Br and I), to systematically reveal the compositional and temperature dependent stabilities, vibrational anharmonicities and band gaps in solid solutions of halide perovskites. This is enabled, in particular, by applying a recently proposed 'anharmonicity score' that provides a single numerical metric to characterise the structural dynamics in a multi–atomic system. Further combination with unsupervised machine–learning enable us to produce an ESP map to visually correlate the anharmonicity score with structural distortions and energies. However, temperature–dependent variations in band gap energies, which strongly depend on orbital interactions in metal–halide octrahedra, do not necessarily follow the same trend as anharmonicity scores. This work represents our latest developments in applying data–driven approach to establish ESP relationships for guiding the future designs of functional perovskites.

1 Introduction

Tremendous research interests have been drawn into halide perovskites for photovoltaic devices and applications.[1–4] The hybrid organic–inorganic perovskites, in particular, methylammonium lead halides [MAPbX₃, X=Cl, Br and I] with up to 25 % of reported photoconversion efficiency,[5] are surpassing those for commercialised thin–film solar cells, such as CdTe. Among the efforts to overcome the problem of 'jelly–like' soft structures for hybrid perovskites,[6] fully inorganic perovskites,[7,8] represented by CsPbX₃, are at the top of the candidate list. CsPbI₃ in its cubic α–phase owns the most desirable electronic band gap of 1.73 eV for photovoltaic applications,[7,9,10] but is prone to phase conversion into its tetragonal β–phase at room temperature with a reduction in band gap.

The stability bottlenecks in halide perovskites are commonly overcome via compositional engineerings. In fact, thanks to the solution processabilities and ion–exchange reactions occurred therein,[11] ion mixings have been widely explored as a chemical strategy to enhance the stabilities of halide perovskites.[7,9,10,12–15] For examples, mixing bromides with iodides at low concentration had been explored to enhance the phase stabilities while impeding the halide segregation in cubic perovskites.[16,17] Partial incorporation of Sr^{2+} can further enhance the air–stability of CsPbI₂Br, by preventing its transformation to a yellow phase.[18,19] The carrier life–time in CsPbBr₃ could be increased by partial incorporation of PbCl₂.[20] The stabilities of hybrid pervoskites could be enhanced via mixing Cs^+ with A–site organic cations,[13,14] and many others.

For domestic applications, replacing lead by other non–toxic elements, while maintaining comparable device performances, is critically needed for bringing halide perovskite solar cells to com-

[a] Materials and Manufacturing Futures Institute, School of Material Science and Engineering, University of New South Wales, Sydney, New South Wales 2052, Australia
E-mail: jianliang.yang1@unsw.edu.au

mercial products.[21] (Partial) replacement of Pb^{2+} metal cations in $APbX_3$ perovskites by a similar group IX metal cations, particularly Sn^{2+},[22–25] forming an inorganic solid solution with the compositions of $Cs(Pb_xSn_{1-x})X_3$, is one possible strategy to mitigate the toxicity problem, provided that the high thermodynamic tendency towards oxidative degradation of Sn–containing compounds can also be reduced.

At a more fundamental level, structural instabilities of perovskites in high symmetry phase are correlated with their electronic structures. The pseudo–Jahn–Teller effect[26] that couples the ground and excited states via a vibrational soft mode, is the dominant contribution to such an instability, whereby upon structural distortion, new bonding state arises in the low symmetry structure which contributes towards additional stabilities. This is more commonly known as the vibrational anharmonicity,[27] widely associated with many perovskite materials. Because these soft modes couple strongly with the electronic states, as a consequence, the electronic properties of halide perovskites are also significantly modulated by the dynamics of ions. Particularly, this affects the dynamics of photo–excited carriers,[28–31] as well as phase stabilities under photon irradiation,[32–34] which all play deterministic roles in its devices stabilities and photoelectric conversion efficiencies. Therefore, without taking into account the dynamic effects, our understandings on the critical structure–energy–property (ESP) relationships for halide perovskites will be incomplete.

For lead halide perovskites, more specifically, the stereochemical effects originated from the coupling between B–s^2 lone pairs and X–p states, the energetic of which determines the electronic band gap, plays a dominating role in the degree of structural instablilities of the materials.[35] Such effect is expected to be stronger when Pb^{2+} is substituted by Sn^{2+}, with the latter possessing more shallow (with respect to the Fermi level) lone–pair electronic states.[36,37] With this in mind, the knowledge of the interplay between different Pb^{2+}/Sn^{2+} mixing and halide compositions on the room–temperature structural and electronic dynamics[16,38–40] will be highly valuable in guiding the designs and optimisations of perovskite photovoltaics with complex chemical compositions. However, whilst the fundamental physics of vibrational anharmonicity is well–known, quantification of materials' degree–of–anharmonicity remains elusive, making it generally challenging for establishing the (quantitative) ESP relationships across many degree–of–freedoms critically needed for guiding material designs.

To tackle the aforementioned challenges, here we undertake a further explorative investigation on the finite–temperature stability, structural as well as electronic dynamics of $Cs(Pb_xSn_{1-x})X_3$ solid solutions, building on the recent work on the 0 K energy landscapes studied on the same system.[41] Our contributions in this work are as following: (a) Free energy calculations show that, while solid solutions of $Cs(Pb_xSn_{1-x})X_3$ are stablised by configurational entropy, the exact solubility limit is determined by the mismatch in electronegtivities for B-site cation and halide anions. (b) Using the recently proposed anharmonicity score,[42] the compositional–dependent enhancement effect of vibrational anharmonicity due to Sn–substitution of Pb, as well as

increase in ionic size of the halide anions, can be clearly revealed. Most importantly, through unsupervised machine–learning analysis of the molecular dynamic trajectory, we are able to provide a clear visual revelation on the geometric interpretation of the anhmaronic score. This enables us to construct an ESP map for halide perovskites that correlate their vibrational anharmonicities with structural change and energetic stabilities. Finally, (c) further electronic structure calculations reveal that temperature–dependent trends of band gap changes are highly specific to the nature of halide anions, but do not necessarily follow the same trends as the anharmonicity scores. This requires more detailed consideration on orbital interactions in thermally fluctuating structures beyond geometric information.

2 Methodologies

2.1 *Ab initio* Static and Dynamic Calculations

Crystal structures of the randomly substituted $Cs(Pb_xSn_{1-x})X_3$ are sampled in a $(4 \times 2 \times 2)$ supercell, giving rise to a total of 17 different B–site metal compositions for each one of the three halides (X=Cl, Br and I) investigated here. Both the atomic positions and lattice constants are fully relaxed at Perdew–Burke–Ernzerhof (PBE)[43] level of theory implemented in the VASP[44] code with standard projector–augmented–wave[45] (PAW) method at 350 eV pseudopotential cut–off. Optimizations are terminated when the total energy change is below 10^{-4} eV. Monkhorst–Pack k–point sampling is used at a uniform grid spacing of 0.04 Å^{-1}. The static energy landscapes for $Cs(Pb_xSn_{1-x})X_3$ have been discussed in a separate work,[41] which will not be repeated here.

For each chemical composition, the lowest–energy structure is extracted on which *ab initio* molecular dynamics (MD) is performed at both 100 and 300 K. All MD calculations are carried out at the `Normal` accuracy setting for VASP with a single k–point. The GW version of the pseudopotential is used to achieve better electronic convergence. Each system is first equilibrated for 500 fs to the target temperature with the velocity scaling algorithm. Production run for a 9 ps trajectory with 1 fs step size is subsequently performed under the *NVT* ensemble using an Andersen thermostat with a collision probability of 0.5. MD snapshots are collected every 1 fs for structural analysis, whereas electronic structure calculations are performed on MD frames extracted every 10 fs along the trajectories. To expedite the calculations, density–of–states (DOS) calculations are performed at the Γ point only, from which the electronic band gaps are extracted using the `pymatgen` Python interface.[46]

2.2 Quantifying Vibrational Anharmonicity

The "degree of vibrational anharmonicity" is an important metric for characterising the structural dynamics of perovskites, which may be applied to establish the structural correlation with the thermal, electronic and optical properties of the materials, particularly for those exhibiting strong electron–phonon couplings. However, works aiming at quantifying the degree of anharmonicity in solid remain sparse. In our previous work,[16] we examined the distributions of atomic displacements sampled from *ab initio* MD, from which we fitted a Gaussian profile to the small–

amplitude part of the distribution, and used the Kullback–Leibler (KL) divergence between the true and fitted distributions as a measurement of anharmonicity in $CsPb(Br_xI_{1-x})_3$. The rationale behind this is that for strongly anharmonic systems, atoms will undergo large–amplitude vibrations, thus exhibiting a distribution that is significantly different from the Gaussian profile with higher probabilities shown the tailing part. However, the range of atomic displacements within which the vibrations are considered as 'harmonic' was, to some extent, arbitrarily defined.

Recently, Knoop *et al.*[42] proposed a non–parametric score of anharmonicity based on the underlying shape of the potential energy surface $\mathscr{V}(\mathbf{R})$, which will be adopted in this work. In brief, the anharmonic score at a given temperature T is defined as:

$$\sigma(T) = \sqrt{\frac{\sum_{j,\alpha} \left\langle \left(F_{j,\alpha}^A\right)^2 \right\rangle_T}{\sum_{j,\alpha} \left\langle \left(F_{j,\alpha}\right)^2 \right\rangle_T}}, \tag{1}$$

in which $F_{j,\alpha}$ is the force obtained from *ab initio* MD on atom j along α Cartesian direction. $\langle \cdot \rangle_T$ represents thermodynamic averaging over a selected ensemble. $F_{j,\alpha}^A$ is the *anharmonic component* of the atomic force, which is given by the difference between $F_{j,\alpha}$ and its *harmonic component* $F_{j,\alpha}^{(2)}$ at MD sampled displacement $\Delta\mathbf{R}_{j,\alpha}$. Under the harmonic approximation, $F_{j,\alpha}^{(2)}$ can be calculated with the knowledge of the force constants in real space $\Phi_{\alpha\beta}^{jk} = \left.\frac{\partial^2 \mathscr{V}(\mathbf{R})}{\partial R_j^\alpha \partial R_k^\beta}\right|_{\mathbf{R}_0}$ as $F_{j,\alpha}^{(2)} = -\sum_{k,\beta} \Phi_{\alpha\beta}^{jk} \Delta\mathbf{R}_{j,\alpha}$. Here, the harmonic force constants are calculated for each optimised structure at 0 K using the PHONOPY[47] code. Because all doped systems remain in a perovskite-type structure only with small deviations from the idealized lattice positions, we relaxed the symmetry tolerance in PHONOPY to 10^{-2} Å and used 3×10^{-2} Å step size to generate displaced structures in the $(4 \times 2 \times 2)$ supercell. This reduces the total number of single point calculations in VASP that are needed for harmonic force constants across 51 compositions. The k–points and pseudopotential parameters remain the same as those in the MD calculations detailed above. Phonon DOS and vibrational free–energy calculations are performed using a denser $(20 \times 20 \times 10)$ mesh grid.

2.3 Quantifying Structural Similarities

Based on the mathematical form of σ, it can be anticipated that this merit may be directly correlated to the 'structural similarity' between an MD frame and the 0 K equilibrium structure (via $\Delta\mathbf{R}_{j,\alpha}$). The concept of a quantifiable 'structural similarity' has been the cornerstone for establishing workable structural–property relationships via (kernel–based) machine–learning in computational material discovery.

Here, the atomic environment for every atom in a given chemical structure is described by the Smoothed Overlap of Atomic Environment (SOAP) structural descriptor.[48] The atomic environment is referred to the radial and angular positions of neighbouring atoms within a cut–off distance r_c around a given atom (also called the centre). The descriptor treats each of the neighbouring

atoms as an atom–centered Gaussian of width σ_0, which can then be expanded with a set of radial and angular basis functions up to order n_{max} and ℓ_{max}, respectively. The SOAP descriptor for a centre is obtained by collecting these expansion coefficients for that specific center, whereas collecting the descriptors for all centres considered in a chemical structure leads to the global structural descriptor.

From here, the global similarity between two multiatomic structures, represented by descriptors $\mathscr{A}$ and $\mathscr{B}$, will be measured via the Regularised Entropy Match (REMatch) kernel[49] $\mathbf{K}_\gamma = \mathrm{Tr}\mathbf{P}_\gamma C(\mathscr{A}, \mathscr{B})$, where $\mathbf{P}_\gamma$ is a doubly stochastic matrix that aims to find matching pairs of atomic environments in two structures by minimising the overall dissimilarities.

Computations of the kernel similarity is performed with the DScribe[50] Python package. We have set $\sigma_0 = 0.3$ Å, $r_c = 6$ Å, $\ell_{max} = n_{max} = 9$ and the regularisation parameter $\gamma = 1$. The scikit-learn[51] package is used for dimensionality reduction analysis with Kernel Principal Component Analysis.

3 Results and Discussions

3.1 Entropic Contributions to the Mixing Thermodynamics

Previously, the compositional–dependent demixing energies of $Cs(Pb_xSn_{1-x})X_3$ solid solutions were systematically investigated,[41] which showed that $Cs(Pb_xSn_{1-x})I_3$ are stable against phase segregation into cubic $CsPbI_3$ and $CsSnI_3$, whereas only Pb–rich $Cs(Pb_xSn_{1-x})Cl_3$ and $Cs(Pb_xSn_{1-x})Br_3$ may exist as single–phased solid solutions. However, these phase stabilities were calculated from DFT–derived total energies at 0 K, which can only captured the effects of chemical bonding on their stabilities.

Here, by taking into account both the configurational and vibrational entropic contributions, we further investigate the temperature–dependent stabilities of $Cs(Pb_xSn_{1-x})X_3$ solid solutions against phase segregation close to realistic operating conditions. More specifically, the phase stability is judged based on the compositional and temperature dependent *mixing* free energies $\Delta G_{mix}(x,T)$[52] for the reaction

$$x\mathrm{CsPbX}_3 + (1-x)\mathrm{CsSnX}_3 \rightarrow \mathrm{Cs(Pb}_x\mathrm{Sn}_{1-x})\mathrm{X}_3,$$
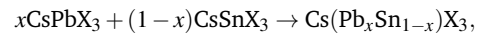
such that

$$\Delta G_{mix} = G[\mathrm{Cs(Pb}_x\mathrm{Sn}_{1-x})\mathrm{X}_3] - xG[\mathrm{CsPbX}_3] - (1-x)G[\mathrm{CsSnX}_3], \tag{2}$$

in which $G[\mathrm{A}]$ represents the the free energy for compound A calculated as following.

The (constant–volume or Helmholtz) configurational free energies for solid solutions with Pb concentration x at temperature T is calculated from the bridging relationship

$$G_{conf}(x,T) = -k_B \ln Z(x,T), \tag{3}$$

where k_B is the Boltzmann constant with partition function[53]

$$Z(x,T) = \sum_{n=1}^{N} g_n \exp\left(-E_n/k_B T\right). \tag{4}$$
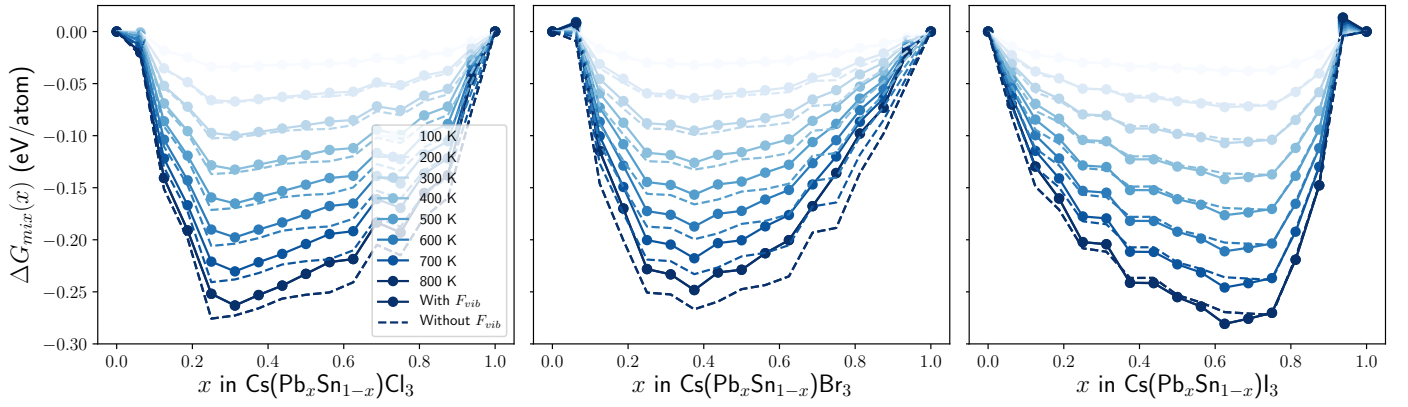
Here g_n is the degeneracy factor for the n–th configuration and E_n

Fig. 1 Mixing free energies $\Delta G_{mix}(x,T)$ as a function of Sn concentrations and temperatures for $Cs(Pb_xSn_{1-x})X_3$. Dashed lines representing $\Delta G_{mix}(x,T)$ solely based on 0 K DFT energies for different dopant configurations, whereas solid lines are calculated $\Delta G_{mix}(x,T)$ with additional contributions from vibrational free energies $F_{vib}(x,T)$ taken into account.

is the 0 K total energy calculated with DFT. Further improvement in the model[52] can be achieved by replacing E_n with vibrational Helmholtz free energy

$$A_n(x,T) = E_n + F_{vib}^n(x,T)$$

in Eq. 4. From the bridge relation (3), $F_{vib}^n(x,T)$ can be obtained from the vibrational partition function $Z_{vib}(x,T)$ for a solid, given by

$$Z_{vib}(x,T) = \prod_{\mathbf{q},\nu} \frac{\exp[-\hbar\omega(\mathbf{q},\nu)/2k_BT]}{1-\exp[-\hbar\omega(\mathbf{q},\nu)/2k_BT]},$$

in which $\omega(\mathbf{q},\nu)$ is the phonon eigenfrequency for the ν–th phonon mode of wavevector $\mathbf{q}$. We note that this approach of calculating $F_{vib}^n(x,T)$ ignores the contributions from phonons of imaginary frequencies,[54–56] and is reasonable when the temperatures of interest (typically the room temperature) are above the Debye temperatures[57] for the materials, which are around 200 K for inorganic halide perovskites.[58]

In principle, for each composition x, $F_{vib}^n(x,T)$ should be calculated for every atomic configuration sampled at this x. However, given the fact that more than 900 configurations have been sampled for $Cs(Pb_xSn_{1-x})X_3$ across all three halides, it is practically not very feasible to perform lattice dynamic calculations for individual structure with 80 atoms each (mostly) at $P1$ symmetry. As such, the vibrational free energy for the lowest energy configuration at a given x, dubbed as $F_{vib}^{\min}(x,T)$, shall be used for approximating $F_{vib}^n(x,T)$ of the higher energy configurations at that concentration. Our analysis [see Supporting Information] shows that the error in estimating ΔG_{mix} from such an approximation is less than 10 %, partially because the free energy curves $F_{vib}^n(x,T)$ do not exhibit significant variations across different configurations at a given x, even though the vibrational signatures of solids are in principle more sensitive to detailed atomistic structures. As such, the approximation adopted here should be able to provide a reasonable comparison on the compositional–dependent entropic stabilisation effects in $Cs(Pb_xSn_{1-x})X_3$ solid solutions.

Fig. 1 summarises the mixing free energy landscapes for all three halides. In stark contrast to the 0 K results discussed above,

Fig. 1 shows that both configurational and vibrational entropies lead to stabilisation of the $Cs(Pb_xSn_{1-x})X_3$ solid solutions. The lowest $\Delta G_{mix}(x,T)$ values at a specific T shows little variance across the three halides, with $Cs(Pb_xSn_{1-x})Br_3$ tend to be the least stable solid solutions. Interestingly, it is the critical concentration x_c, at which $\Delta G_{mix}(x,T)$ become most negative, exhibit strong dependency upon the chemical nature of the halogen anion, which may correspondingly determine the true solubility limit of Sn^{2+} in $Cs(Pb_xSn_{1-x})X_3$ under realistic experimental conditions. Fig. 1 shows that x_c increases from ~ 0.3 (or Sn–rich) for chlorides to ~ 0.6 (or Pb–rich) for iodidies, accompanied by an unsymmetric shape in $\Delta G_{mix}(x,T)$ for all three halides. This indicates that $Cs(Pb_xSn_{1-x})X_3$ do not behave as ideal solid solutions. Since it is the cohesive energies for solids that play an important role in the partition function [Eq. 3] for the configurational entropy, such behaviour should be understood from the view of chemical bonding. This can be intuitively rationalised based on the model of electronegativity mismatches:[15] with the electronegativities decrease for halides as $Cl^->Br^->I^-$, but increase for B–site cation as $Sn^{2+}<Pb^{2+}$, there is a stronger thermodynamic tendency for $Cl^-(I^-)$ to bond with $Sn^{2+}(Pb^{2+})$ in order to strengthen the charge transfer between B–site cation and halogen anion, further enhancing the thermodynamic stabilities of the solid solutions.

Vibrational entropy could also play an important contribution in stabilising solid solutions, particularly for the accurate determinations of phase transition temperatures.[59] However, vibrational entropy could either stabilise or destablise the mixing of solid solutions, which depends critically on the subtle balance between G_{conf} and F_{vib} (for a nice illustration of such effect, see the work by Manzoor *et al.*[60]). From Fig. 1, it can be seen that, for chloride and bromide, F_{vib} generally destabilises the solid solutions across all compositions. The largest increase in $\Delta G_{mix}(x)$ is around 40 meV/atom (3.9 kJ/mol/atom) for $Cs(Pb_xSn_{1-x})Br_3$ at $x = 0.625$ at 800 K, which is comparable to the effects observed in, for instance, $Sn(S_{1-x}Se_x)$ solid solutions.[37,52] However, for $Cs(Pb_xSn_{1-x})I_3$, we see an almost identical (even minor stabilising) $\Delta G_{mix}(x)$ profile with the inclusion of $F_{vib}(x,T)$. Overall,
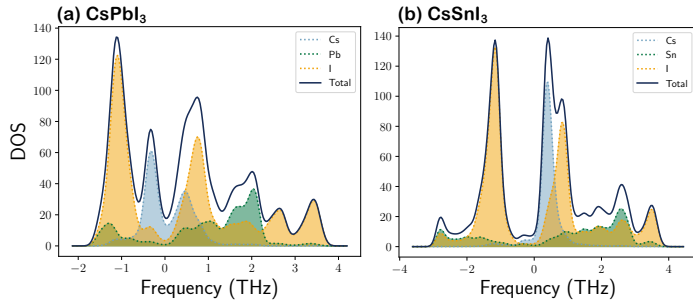
Fig. 2 Harmonic phonon density–of–states (DOS) for (a) CsPbI$_3$ and (b) CsSnI$_3$.

Fig. 3 Joint probability distributions between F and F^A from MD sampling at 300 K for different systems. All forces have been renormalised by the standard deviation of DFT forces [denoted as $\sigma(F)$, bounded between the horizontal dashed lines]. (a) Si, a harmonic system acting as benchmark, with MD performed on a 64–atom supercell, which resemble the result presented by Knoop *et al.*[42] All atoms in (b) CsPbI$_3$ and (c) CsSnI$_3$. (d-f) Joint probability distributions presented separately for forces on Cs, Sn and I atoms in CsSnI$_3$.

$\Delta G_{mix}(x)$ shown in Fig. 1 provided a clear evidence of configurational entropy stabilisation of Cs(Pb$_x$Sn$_{1-x}$)X$_3$ solid solutions [except for the very diluted cases of Cs(Pb$_{1/16}$Sn$_{15/16}$)Br$_3$ and Cs(Pb$_{15/16}$Sn$_{1/16}$)I$_3$, both contain only 1 configuration].

3.2 Vibrational Behaviours of Cubic CsPbI$_3$ and CsSnI$_3$

Before we discuss the compositional dependent structural dynamic of Cs(Pb$_x$Sn$_{1-x}$)X$_3$ at finite temperature, we first examine this more specifically in cubic CsPbI$_3$ and CsSnI$_3$, in order to facilitate the subsequent discussions. It has been well established that the pseudo–Jahn–Teller effect (PJTE)[26] plays dominant contributions to the instability of high–symmetric cubic phase of perovskites. Typically in the high–symmetry structure, the phases of interacting atomic orbitals in BX$_6$ octahedron, which contribute to a ground (valence band, VB) and an excited (conduction band, CB) state, cancel out exactly, leading to a destabilising anti–bonding molecular orbital. Upon symmetry–lowering structural distortion along a soft phonon mode, such phase cancellation is removed, which resulted in the formation of new chemical bonds between B–site cation and X–site anion in ABX$_3$. Analysis of PJTE based on perturbation theory shows that, PJTE leads to an adiabatic potential energy surface of the form $E(\Delta x) = \alpha\Delta x^2 + \beta\Delta x^4$ (where Δx is the magnitude of atomic displacement from the equilibrium position) without invoking the third–order term. Materials that exhibit strong anhmarmonicity are those that possess a new minima at $\Delta x \neq 0$ on this potential energy surface. This, in term, is determined by β, which is directly proportional to the coupling strength between the ground and excited electronic states, while inversely related to the energy gap separating them. In halide perovskites, the VB edges are dominated by the Pb/Sn–ns^2 lone–pair states, whereas the CB edges are contributed by the halogen p–states. A shallower Sn–5s^2 states compared to the Pb–6s^2 states results in a generally smaller electronic band gap for CsSnX$_3$ compared to CsPbX$_3$, consequentially, this enhances the PJTE and thus the vibrational anharmonicity in CsSnX$_3$. The anharmonic vibrational behaviours are typically reflected as the presence of negative (or imaginary) harmonic phonon frequencies (see phonon DOS shown Fig. 2). For both CsPbI$_3$ and CsSnI$_3$, it can be seen that I contributes strongest to the phonon DOS at around 1.5i THz. However, when Pb is replaced by Sn, an even softer phonon peak at around 3i THz for the vibration of Sn–sublattice appears, signifying the enhanced
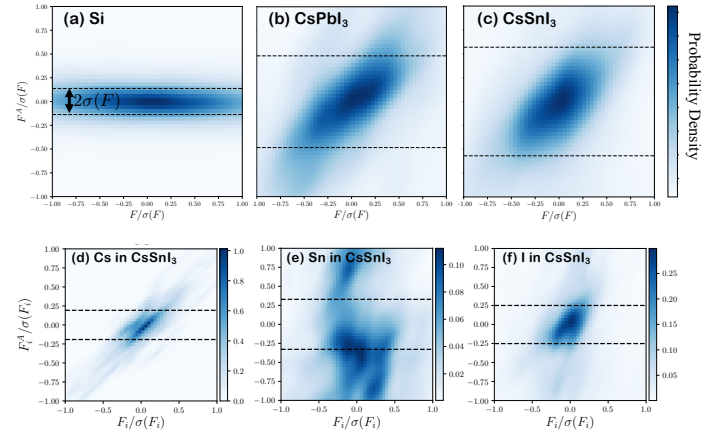
anharmonic nature for the latter case.

Based on this understanding, a gradual change of vibrational anharmonicity (at finite temperature) with respect to the concentration gradient of B–site cations in Cs(Pb$_x$Sn$_{1-x}$)X$_3$ should be expected, but the challenge in computational material science so far has been on how to quantify the degree of vibrational anharmonicities across many degree of freedoms in complex materials.

The anharmonic score [Eq. (1)] is proposed to solve this challenging by tracing back to the joint probability distribution functions (PDF) between total (F) and anharmonic (F^A) forces sampled along an MD trajectory. In a harmonic system, such as Si [Fig. 3(a)], the normalised anharmonic components of the atomic forces are well–confined within the range of $\pm\sigma(F)$, the standard deviation in total forces, irrespective of the magnitudes of the total atomic forces F. However, the PDFs for CsPbI$_3$ [Fig. 3(b)] and CsSnI$_3$ [Fig. 3(c)] appears to be in the shape of a multivariant Gaussian distribution that is fundamentally different to the PDF for Si, due to their intrinsic anharmonic characters. It indicates that, overall, the magnitudes of the anharmonic atomic forces are directly proportional to the total forces. The larger variant in F for CsSnI$_3$ [dashed lines in Fig. 3(c)] compared to CsPbI$_3$ provides a robust statistical measure on the (as expected) enhanced vibrational anharmonicity in the former case, even though the PDF for CsPbI$_3$ clearly shows higher relative probabilities at the high $|F^A|$ tails.

The way that the anharmonic score was formulated allows us to further reveal the contributions from individual atoms to the vibrational anharmonicity in the entire structure, which Fig. 3(d–f) shows the breakdown for each element in CsSnI$_3$. It can be seen here that the large Cs$^+$ cations remains relatively immobile. The Sn^{2+} cations, on the other hand, shows a more complex and strong anharmonic character at 300 K. Most notably, F^A on the Sn sub–lattice does not increase monotonically with F, but fluctuate
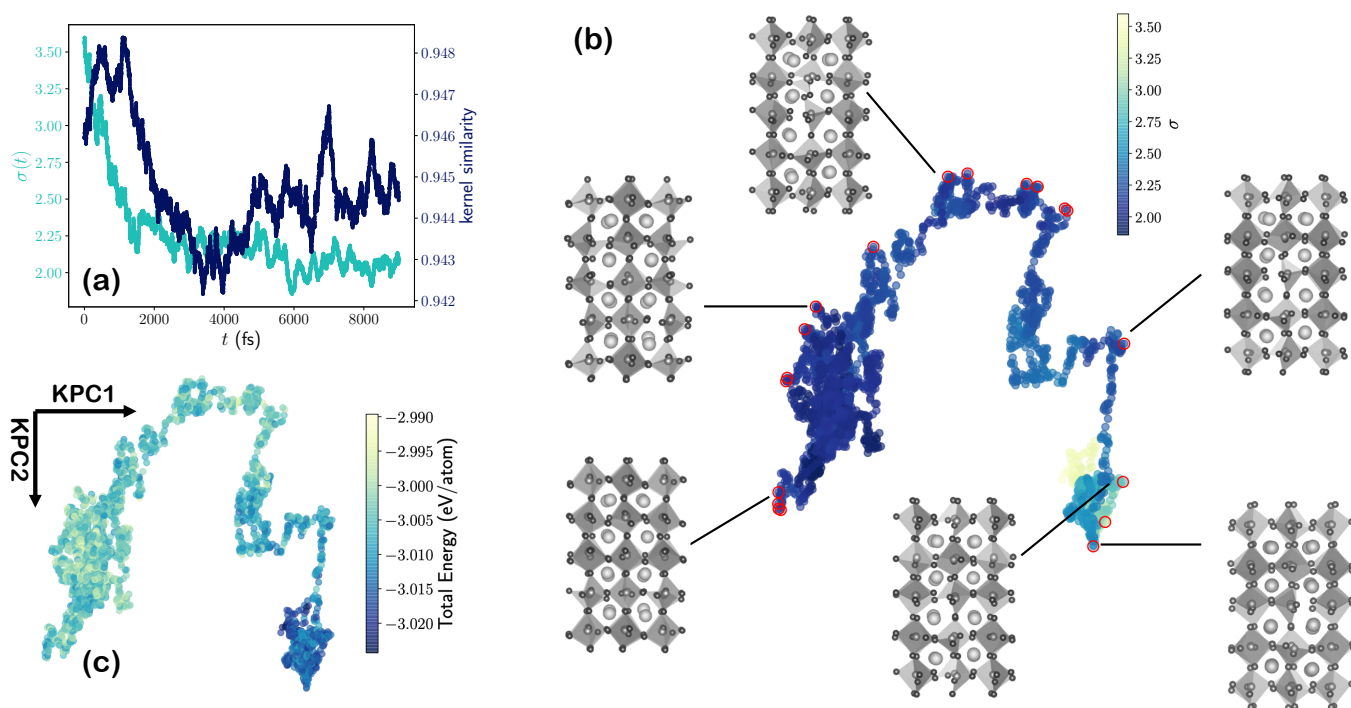
Fig. 4 (a) Time dependent anharmonic scores and structural similarities (measured by the SOAP–REMatch kernel with respect to the 0 K static structure) for the MD trajectory of $CsSnI_3$ at 300 K. (b-c) Energy–structure–property maps for 3000 equally spaced MD frames for $CsSnI_3$ at 300 K, in which the structural similarities among the frames are measured with the SOAP–REMatch kernel. In (b) each point on the map is colour–coded according to its anharmonicity score (σ), with a few representative structures on the vertices of the convex hull for the 2D maps (red open circles) shown. The same map is shown again in (c) with each point colour coded according to its total energy.

over a wide range $[> 2\sigma(F)]$ even at vanishing total forces. This reflects the situation where the harmonic well is extremely steep along certain directions of atomic vibrations, such that even with a small amplitude of atomic movement, the atom will encounter a large anharmonic force to counter–balance the harmonic component. This could arise from the large exchange–repulsion between the charge density from shallow $Sn–5s^2$ lone pair and the diffuse $I–p$ states, such that a small out–of–equilibrium movement of Sn atoms will encounter a sharp–rising exchange–repulsion wall with surrounding I atoms that tries to restore the Sn^{2+} back to the equilibrium position.[36]

3.3 Energy–Structure–Property Map Helps Enhancing Understanding of the Anharmonicity Score

Molecular dynamics simulations can provide us with rich information for helping us to understand how chemical bonding, structures and their dynamics affect their physical properties. However, the information is also complex and usually of high–dimensionality that is practically impossible for us to retrieve this information easily. To overcome this challenge, and to obtain a more insightful understanding on the meaning of anharmonicity scores, we further apply unsupervised machine–learning method, based on a robust kernel (the SOAP–REMatch kernel) for structural similarity comparison, to discover the correlation between energy, anharmonic scores, and the dynamic fluctuation of the atomistic structures in cubic $CsSnI_3$ at 300 K.

In Fig. 4(a), for each MD snapshot of $CsPbI_3$ at 300 K, we first plot its corresponding σ, against its *overall* structural similarity to the 0 K static structure of cubic $CsPbI_3$, as measured by the SOAP–REMatch similarity. The anti–correlation between σ and structural similarity is immediately apparent, with large σ generally translates to a low structural similarity. This is understandable as harmonic approximations are only valid in the vicinity of zero atomic displacements on the PES. Therefore, as the atoms move further away from their equilibrium positions, it lowers their 'structural similarity' with respect to the equilibrium structure, and correspondingly, increases the degree of vibrational anharmonicity.

However, before proceeding, it is critical to keep in mind that this correlation only holds when comparing the structures sampled *within an MD trajectory at a given temperature*, but not trajectories across two different temperatures [see SI for the same plots at 100 and 300 K]. The simplest way to understand this is to realise that the harmonic forces $F^{(2)}$ depends not only on the atomic displacements but also the force constants. So if we consider the effective temperature–dependent potential energy surface $\mathscr{V}(\mathbf{R})$ as a thermally weighted combinations of harmonic vibrational energy surfaces contributed by phonons at different frequencies, it means that the shapes of $\mathscr{V}(\mathbf{R})$ effectively changes at a different temperature even $\Delta\mathbf{R}$ remains unchanged, which consequentially lead to different anharmonic scores.

Instead of using our knowledge to select a 'referencing landmark' for structural comparisons, a solely 'data–driven' approach can be applied to discover the geometric meaning of anharmonic
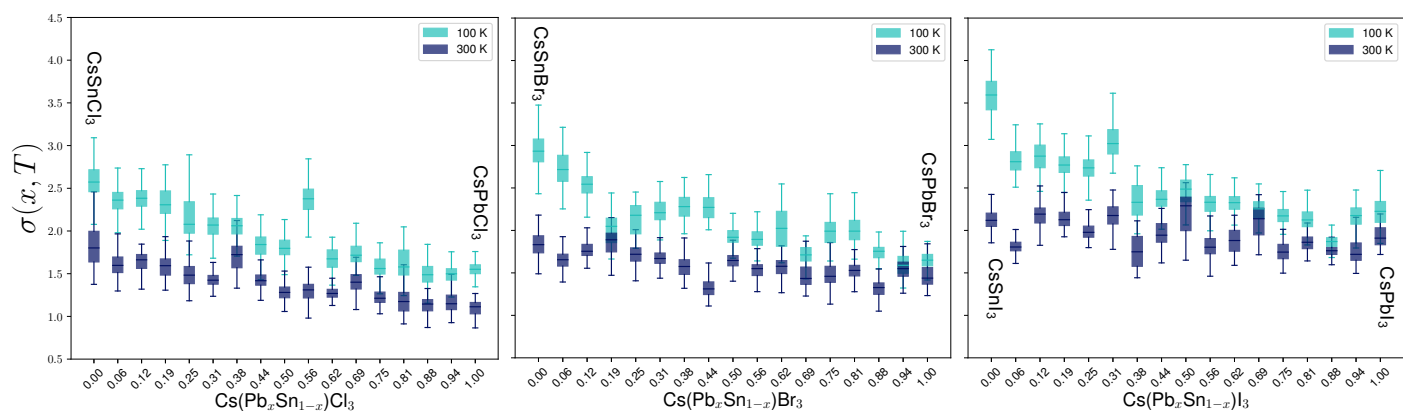
Fig. 5 Composition–dependent anharmonicity scores σ for $Cs(Pb_xSn_{1-x})X_3$ (X=Cl, Br and I) at 100 and 300 K. Box plots show the distributions of σ calculated for MD frames extracted every 1 fs in the last 7 ps of the MD trajectories. In this case, the ensemble averaging in Eq. (1) is performed over the distribution of all atomic forces extracted within an MD frame.

scores directly within an MD trajectory.[16,49,61] For this purpose, the SOAP–REMatch similarities are calculated for all pairs of frames in the trajectory to build the SOAP–REMatch similarity kernel. Dimensionality reduction (a sub–category for unsupervised machine learning) can be subsequently applied to build a two–dimensional map to help discovering (dis)similar atomistic configurations sampled by MD. Here, the kernel principal component analysis (KPCA)[62] is adapted to achieve this goal. The first two KPCs identified represents directions in the high–dimensional space that encodes largest structural variations. By projecting each structure (represented by the high–dimensional SOAP structural descriptor) along these two KPCs, with further decoration of each point by its corresponding property (such as energy or anharmonic score), 'energy–structure–property' (ESP) maps[61,63] can be constructed to uncover intriguing structure–property relationships in the dataset.

The ESP maps for the 300 K MD trajectory of cubic $CsPbI_3$, with each point color–coded according to the anharmonicity scores [Fig. 4(b)] and total energies [Fig. 4(c)] reveal that: (a) Configurations that are of the highest and lowest anharmonic scores are separated furthest along the first KPC. The close spatial distances amongst points of low/high anharmonicity scores means configurations with similar anharmonicit scores are also structurally very similar. Hence this provides us with direct visual evidence that vibrational anharmonicity is correlated with the amplitudes of atomic motions in the material. (b) Configurations with high anharmonicity scores also appears to be those with the lowest energies on the ESP landscape, which is in line with how we understand the stabilising effects associated with soft–mode vibrations.

Structures located on the peripherals of the ESP maps corresponds to set of most distinctly different atomistic configurations sampled by MD, and in terms of KPCA, these points can be identified as the vertices of the convex hull[62] that encloses all the points on the ESP map. A few of these configurations are highlighted in Fig. 4(c). Because at finite temperature, more than one vibrational modes will be excited and contribute to the atomic motions captured by each MD snapshot, it is not trivial to visually

identify the subtle differences in these MD frames exemplified, not to mention how their properties might be different. Our finding thus suggests that using similarity kernel and anharmonicity scores together is particularly useful to circumvent limitations in visual heuristics for analysing and rationalising the results of molecular dynamic simulations for solids.

3.4 Composition and Temperature Dependent Vibrational Anharmonicities

With the introduction of the anharmonic score, we are now able to systematically examine the trends of vibrational aharmonicities in solid solutions of halide perovskites across a compositional gradient at different temperatures. The result from this investigation is shown in Fig. 5, where we plotted the distributions of σ at 100 and 300 K for $Cs(Pb_xSn_{1-x})X_3$ (X=Cl, Br and I), across 17 different Pb/Sn ratios for each halides. The boxes in Fig. 5 reflects the spread of σ calculated for each MD frame across the last 7 ps for each MD trajectories [which appears to be more stable as shown in Fig. 4(a)].

Fig. 5 reveals that: (a) Within a specific halide, σ increases as the concentration of Sn^{2+} is increased, clearly reflecting the shallow $5s^2$–lone pair effects from Sn^{2+} cations.[41] (b) It is also evident that the vibrational anharmonicities gradually enhance from chloride to iodide, because of the stronger Pauli repulsions between larger halide anions with the B–site lone–pair electrons. Moreover, the change in σ across the range of Pb/Sn ratios is seemingly steeper for chloride compared to iodide, indicating that iodide, being the largest anion in the three halides investigated here, starts to overtake the B–site cations in dictating the vibrational behaviours of $Cs(Pb_xSn_{1-x})I_3$. Examining the compositional trends more closely at the diluted ends of the compositional gradients further reveals the asymmetric effects from B–site intermixing, whereby small concentration ($x = 0.06$) of Pb^{2+} into $CsSnX_3$ can significantly quench the vibrational anharmonicity in $CsSnX_3$ (especially for iodide), but on the opposite end, the change in the vibrational properties is very minor when low–concentration of Sn^{2+} is mixed into $CsPbX_3$. (c) The systematic decrease of vibrational anharmonicity upon rising temperature
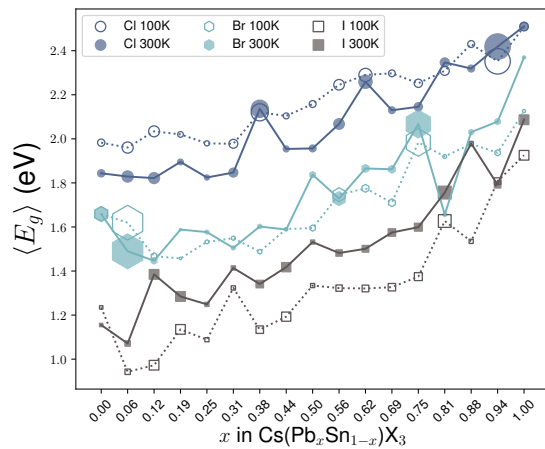
Fig. 6 Composition–dependent expectation values of electronic band gap energies for cubic $Cs(Pb_xSn_{1-x})X_3$ at 100 and 300 K. The sizes of the markers are proportional to the standard deviations in the calculated band gaps across the MD trajectory.

over the entire compositional range is also well revealed by σ. The high σ reflects the predominant excitation of soft phonon modes at low temperature, which is consistent with our previous theoretical analysis on the temperature–dependent vibrational anharmonicity in $CsPb(I_{1-x}Br_x)_3$ based on the atomic displacement histogram using the Kullbeck–Leibler divergences.[16]

It should be note that, Fig. 5 shows all systems investigated here have σ values greater than 1, which are cases that had not been exemplified in the original paper by Knoop *et al.*[42] We have doubled checked our own implementation of the algorithm with Si as a reference (see Fig. 3) with MD performed in the same protocol as Knoop *et al.*[42] and confirmed its reproducibility. Hence, we believe the high anharmonic scores for $Cs(Pb_xSn_{1-x})X_3$ is a true reflection of the soft structural character for this family of materials.

3.5 Composition and Temperature Dependent Electronic Dynamics

As discussed above, vibrational anharmonicty are strongly coupled with the electronic structures of $Cs(Pb_xSn_{1-x})X_3$, hence the vibrations of atoms will also dynamically modulate their electronic structures. Such a coupling between the electronic and structural degrees–of–freedoms has direct impact in the photophysics of perovskites, including the lifetime of photoinduced carriers and photoelectric conversion efficiencies. Therefore, it is important to further examine the fluctuations of electronic properties for $Cs(Pb_xSn_{1-x})X_3$ at elevated temperatures and seek for possible correlations with their structural dynamics.

With this in mind, we calculated the (density–of–states) electronic band gap energies for the MD frames of $Cs(Pb_xSn_{1-x})X_3$ (X=Cl, Br and I) at both 100 and 300 K. Before we discuss our findings in details, it is important to acknowledge that the band gap energies for $Cs(Pb_xSn_{1-x})X_3$ are underestimated when calculated with semi–local PBE functional at a single K–point. Including the relativistic effects for heavy Pb atoms tends to signifi-

cantly close up the band gap, whereas hybrid functionals (such as HSE06), although are generally believed to provide better band gap estimations, open up the band gap too much. Taking into account higher–order vertex correction and thermal fluctuations, Wiktor *et al.*[64] demonstrated that more satisfactory estimations for the band gaps of halide perovskites are attainable. Nevertheless, the PBE gaps are reasonably close to the true values for qualitative assessments[41] and this is critical for expediting the property screenings for 30,600 MD frames across two different temperature, three different halogens and 17 different Pb/Sn ratios, which, to our knowledge, are not routine exercises to perform in order to thoroughly examine and understand the structure–property relationships in complex materials.

Firstly, we examine the overall composition and temperature dependent trends in the band gap energies by using the expectation value of band gap energy at a given temperature for a specific composition, $\langle E_g \rangle = \frac{1}{N}\sum_i^N E_{g,i}$, where $E_{g,i}$ is the band gap energy for the i–th MD frame. The results are shown in Fig. 6. The enlargement of $\langle E_g \rangle$ with the increase of Pb contents, as well as from heavier I to lighter Cl anion can clearly be reflected in Fig. 6. Such a behaviour is temperature independent. Qualitatively, $\langle E_g \rangle(x)$ exhibits a 'bowing–effect' which has been widely documented for halide perovskites,[65] but the extent of bowing effects shows no strong variance across three halides. The standard deviations in $E_g(x,T)$, which can be directly related to the strength of electron–phonon couplings,[66] did not reveal clear compositional–dependency in this case, which are indicated by the sizes of the markers in Fig. 6.

Nevertheless, the most interesting and systematic trend revealed in Fig. 6 is the differences in the temperature–dependencies of $\langle E_g \rangle$ across three halides. Upon temperature increase from 100 to 300 K, there is a clear overall increase (decrease) in $\langle E_g \rangle$ for iodides (chlorides) across the entire range of chemical compositions, whereas for bromides, the band gap energies are largely unchanged. This trend is clearly more complex compared to the temperature–dependent vibrational anharmonicities revealed in Fig. 5, and must be understood with a closer look at the electronic structures. Since an increase in the electronic band gaps is generally expected at elevated temperature, we focus here more specifically on the opposite case for chlorides.

In Fig. 7, the fluctuations in the atom–resolved electronic density–of–states (DOS) closer to the band edges for $Cs(Pb_{0.5}Sn_{0.5})Cl_3$ and $Cs(Pb_{0.5}Sn_{0.5})Br_3$ at 100 and 300 K are compared. For the chloride, it can be seen that, upon temperature elevation, the overall DOS peak shapes remain largely unchanged. The lighter chloride anions, which are expected to vibrate faster, gives rise to a much larger fluctuations in DOS for $Cs(Pb_xSn_{1-x})Cl_3$ close to the valence band edge at 300 K. As a result, the chemical bonding in the BX_6 octahedra are expected to be weakened causing a gap closure at 300 K. Such an effect could also exist for bromides and iodides, but as revealed in Fig. 7 (right), the DOS peak for Sn and Br closer to the valence band edge at 300 K are significantly enhanced, which strengthened the orbital interactions that, in principle, open up the electronic band gaps. As a result, one sees the band gap for bromides remains
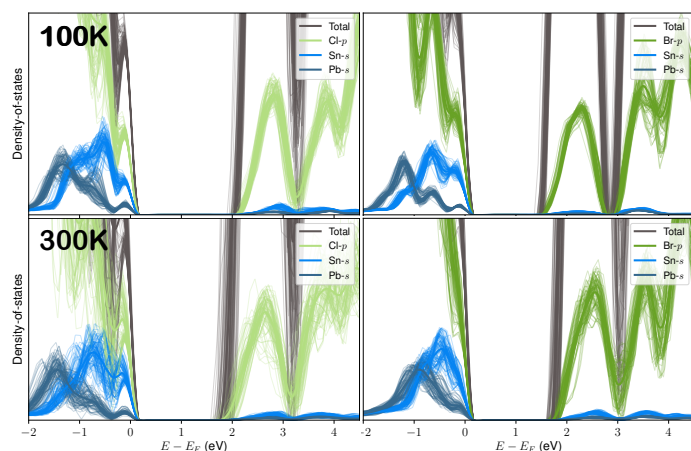
Fig. 7 Fluctuations of electronic density–of–states for 100 random MD snapshots for $Cs(Pb_{0.5}Sn_{0.5})Cl_3$ (left) and $Cs(Pb_{0.5}Sn_{0.5})Br_3$ (right) at 100 and 300 K.

relatively unchanged with respect to temperature.

4   Conclusions

To gain a deeper understanding on the physical properties of inorganic halide perovskite $Cs(Pb_xSn_{1-x})X_3$ (X=Cl, Br and I) for future material designs of lead–free perovskite photovoltaics, we performed DFT–based dynamic and electronic structure calculations at both zero and elevated temperatures, augmented by unsupervised machine–learning techniques to reveal the energy–structure–property relationships in these materials. Our findings are summarised as following.

Across all three halides, configurational entropies are shown to play a significant role in stabilising the $Cs(Pb_xSn_{1-x})X_3$ solutions, whereas vibrational entropies generally destabilise them, but the effects are minor. The mismatch in electronegativities between the B^{2+} cations and X^- anions are found to play a dominant role in determining the critical mixing concentrations at which the mixing free energies are the lowest on the energy landscape.

By using the recently proposed anharmonic scores for materials,[42] we are able to more clearly and systematically show the enhancement of vibrational anharmonicities in $Cs(Pb_xSn_{1-x})X_3$ with respect to the increase of Sn^{2+} concentrations, changes of halide from lighter Cl^- to heavier I^-, as well as decrease in temperature, compared to our previous approach based on Kullback–Leibler divergence of the histogram for atomic displacements.[16] With further application of the SOAP–REMatch similarity kernel[49] and kernel principal component analysis, we built a two–dimensional energy–structure–property map based on the MD trajectory for cubic $CsSnI_3$ at 300 K. This provides us with a clear visual interpretation on the geometric meaning of the anharmonic scores, the values of which can be directly correlated with the amount of structural fluctuations in the solid with respect to its equilibrium atomistic positions.

Finally, we investigated the fluctuations of electronic band gaps at 100 and 300 K for $Cs(Pb_xSn_{1-x})X_3$ to examine how electron–phonon couplings are affected by both chemical compositions and thermal vibrations. Surprisingly, it is found that band gap ener-

gies for bromides remain largely unchanged upon temperature rise, whereas those for iodides (chlorides) increased (decreased) across all Sn concentrations. This is believed to be the results of the differences in the modulation of bonding/antibonding orbital interactions in the BX_6 octahedra lattice vibrations.

We hope that this work will not only provide a deeper physical insight into the interplay between chemical bonding and structural/electronic dynamics for inorganic halide perovskites, but more importantly, a systematic methodological framework to discover and understand the energy–structure–function relationships in perovksite (opto)electronic materials.

Acknowledgements

References

1 H. Zhou, Q. Chen, G. Li, S. Luo, T.-b. Song, H.-S. Duan, Z. Hong, J. You, Y. Liu and Y. Yang, *Science*, 2014, **345**, 542.

2 F. Hao, C. C. Stoumpos, D. H. Cao, R. P. Chang and M. G. Kanatzidis, *Nature Photonics*, 2014, **8**, 489.

3 L. Dou, A. B. Wong, Y. Yu, M. Lai, N. Kornienko, S. W. Eaton, A. Fu, C. G. Bischak, J. Ma, T. Ding *et al.*, *Science*, 2015, **349**, 1518.

4 Y. Zhao and K. Zhu, *Chem. Soc. Rev.*, 2016, **45**, 655.

5 *Best Research-Cell Efficiency Chart*, 2020 (accessed July 1st, 2020).

6 J. M. Frost and A. Walsh, *Acc. Chem. Res.*, 2016, **49**, 528.

7 R. J. Sutton, G. E. Eperon, L. Miranda, E. S. Parrott, B. A. Kamino, J. B. Patel, M. T. Hörantner, M. B. Johnston, A. A. Haghighirad, D. T. Moore *et al.*, *Adv. Energy Mater.*, 2016, **6**, 1502458.

8 J. Liang, C. Wang, Y. Wang, Z. Xu, Z. Lu, Y. Ma, H. Zhu, Y. Hu, C. Xiao, X. Yi *et al.*, *J. Am. Chem. Soc.*, 2016, **138**, 15829.

9 Q. A. Akkerman, V. D'Innocenzo, S. Accornero, A. Scarpellini, A. Petrozza, M. Prato and L. Manna, *J. Am. Chem. Soc.*, 2015, **137**, 10276.

10 S. Dastidar, D. A. Egger, L. Z. Tan, S. B. Cromer, A. D. Dillon, S. Liu, L. Kronik, A. M. Rappe and A. T. Fafarman, *Nano Lett.*, 2016, **16**, 3563.

11 Y. Fu, F. Meng, M. B. Rowley, B. J. Thompson, M. J. Shearer, D. Ma, R. J. Hamers, J. C. Wright and S. Jin, *J. Am. Chem. Soc.*, 2015, **137**, 5810.

12 D. Bai, H. Bian, Z. Jin, H. Wang, L. Meng, Q. Wang and S. F. Liu, *Nano Energy*, 2018, **52**, 408.

13 Z. Li, M. Yang, J.-S. Park, S.-H. Wei, J. J. Berry and K. Zhu, *Chem. Mater.*, 2016, **28**, 284–292.

14 C. Yi, J. Luo, S. Meloni, A. Boziki, N. Ashari-Astani, C. Grätzel, S. M. Zakeeruddin, U. Röthlisberger and M. Grätzel, *Energy Environ. Sci.*, 2016, **9**, 656–662.

15 W.-J. Yin, Y. Yan and S.-H. Wei, *J. Phys. Chem. Lett.*, 2014, **5**, 3625.

16 J. Yang, Y. Wang, T. Wu and S. Li, *Chem. Mater.*, 2020, **32**, 2470.

17 Y. Wang, X. Guan, W. Chen, J. Yang, L. Hu, J. Yang, S. Li, K. Kalantar-Zadeh, X. Wen and T. Wu, *ACS Appl. Mater. Interf.*, 2020, **34**, 38376.

18 B. Parida, S. Yoon, S. M. Jeong, J. S. Cho, J.-K. Kim and D.-W. Kang, *Solar Energy Mater. Solar Cells*, 2020, **204**, 110212.

19 C. F. J. Lau, M. Zhang, X. Deng, J. Zheng, J. Bing, Q. Ma, J. Kim, L. Hu, M. A. Green, S. Huang *et al.*, *ACS Energy Lett.*, 2017, **2**, 2319.

20 B. Li, Y. Zhang, L. Zhang and L. Yin, *J. Power Sources*, 2017, **360**, 11.

21 M. Lyu, J.-H. Yun, P. Chen, M. Hao and L. Wang, *Adv. Energy Mater.*, 2017, **7**, 1602512.

22 N. K. Noel, S. D. Stranks, A. Abate, C. Wehrenfennig, S. Guarnera, A.-A. Haghighirad, A. Sadhanala, G. E. Eperon, S. K. Pathak, M. B. Johnston *et al.*, *Energy Environ. Sci.*, 2014, **7**, 3061.

23 Y. Liao, H. Liu, W. Zhou, D. Yang, Y. Shang, Z. Shi, B. Li, X. Jiang, L. Zhang, L. N. Quan *et al.*, *J. Am. Chem. Soc.*, 2017, **139**, 6693.

24 L. Lanzetta, J. M. Marin-Beloqui, I. Sanchez-Molina, D. Ding and S. A. Haque, *ACS Energy Lett.*, 2017, **2**, 1662.

25 P. V. Kamat, J. Bisquert and J. Buriak, *ACS Energy Lett.*, 2017, **2**, 904.

26 I. B. Bersuker, *Chem. Rev.*, 2013, **113**, 1351.

27 C. Katan, A. D. Mohite and J. Even, *Nature Mater.*, 2018, **17**, 377.

28 M. Z. Mayers, L. Z. Tan, D. A. Egger, A. M. Rappe and D. R. Reichman, *Nano Lett.*, 2018, **18**, 8041.

29 T. J. Jacobsson, J.-P. Correa-Baena, M. Pazoki, M. Saliba, K. Schenk, M. Grätzel and A. Hagfeldt, *Energy Environ. Sci.*, 2016, **9**, 1706.

30 J. M. Frost, *Phys. Rev. B*, 2017, **96**, 195202.

31 K. Miyata, D. Meggiolaro, M. T. Trinh, P. P. Joshi, E. Mosconi, S. C. Jones, F. De Angelis and X.-Y. Zhu, *Science Adv.*, 2017, **3**, e1701217.

32 Y. Yuan and J. Huang, *Acc. Chem. Res.*, 2016, **49**, 286.

33 E. T. Hoke, D. J. Slotcavage, E. R. Dohner, A. R. Bowring, H. I. Karunadasa and M. D. McGehee, *Chemi. Sci.*, 2015, **6**, 613.

34 P. Cottingham and R. L. Brutchey, *Chem. Mater.*, 2016, **28**, 7574.

35 M. D. Nielsen, V. Ozolins and J. P. Heremans, *Energy Environ. Sci.*, 2013, **6**, 570.

36 D. H. Fabini, G. Laurita, J. S. Bechtel, C. C. Stoumpos, H. A. Evans, A. G. Kontos, Y. S. Raptis, P. Falaras, A. Van der Ven, M. G. Kanatzidis *et al.*, *J. Am. Chem. Soc.*, 2016, **138**, 11820.

37 J. M. Skelton, L. A. Burton, A. J. Jackson, F. Oba, S. C. Parker and A. Walsh, *Phys. Chem. Chem. Phys.*, 2017, **19**, 12452.

38 W. Li, J. Tang, D. Casanova and O. V. Prezhdo, *ACS Energy Lett.*, 2018, **3**, 2713.

39 W. Li, A. S. Vasenko, J. Tang and O. V. Prezhdo, *J. Phys. Chem. lett.*, 2019, **10**, 6219.

40 G. Zhou, W. Chu and O. V. Prezhdo, *ACS Energy Lett.*, 2020.

41 J. Yang, *Phys. Chem. Chem. Phys.*, 2020, DOI:10.1039/D0CP03170D.

42 F. Knoop, T. A. Purcell, M. Scheffler and C. Carbogno, *arXiv:2006.14672*, 2020.

43 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.

44 G. Kresse and J. Furthmüller, *Phys. Rev. B*, 1996, **54**, 11169.

45 G. Kresse and D. Joubert, *Phys. Rev. B*, 1999, **59**, 1758.

46 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, *Comput. Mater. Sci.*, 2013, **68**, 314.

47 A. Togo and I. Tanaka, *Scr. Mater.*, 2015, **108**, 1.

48 A. P. Bartók, R. Kondor and G. Csányi, *Phys. Rev. B*, 2013, **87**, 184115.

49 S. De, A. P. Bartók, G. Csányi and M. Ceriotti, *Phys. Chem. Chem. Phys.*, 2016, **18**, 13754.

50 L. Himanen, M. O. J. Jäger, E. V. Morooka, F. Federici Canova, Y. S. Ranawat, D. Z. Gao, P. Rinke and A. S. Foster, *Comput. Phys. Commun.*, 2020, **247**, 106949.

51 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825.

52 J. M. Skelton, *J. Phys. Energy*, 2020, **2**, 025006.

53 R. Grau-Crespo, S. Hamad, C. R. A. Catlow and N. De Leeuw, *J. Phys. Cond. Matter*, 2007, **19**, 256201.

54 U.-G. Jong, C.-J. Yu, Y.-H. Kye, Y.-S. Kim, C.-H. Kim and S.-G. Ri, *J. Mater. Chem. A*, 2018, **6**, 17994.

55 U.-G. Jong, C.-J. Yu, Y.-H. Kye, S.-H. Choe, J.-S. Kim and Y.-G. Choe, *Phys. Rev. B*, 2019, **99**, 184105.

56 E. L. Da Silva, J. M. Skelton, S. C. Parker and A. Walsh, *Phys. Rev. B*, 2015, **91**, 144107.

57 S. G. Moustafa, A. J. Schultz, E. Zurek and D. A. Kofke, *Phys. Rev. B*, 2017, **96**, 014117.

58 A. Pisoni, J. Jacimovic, O. S. Barisic, M. Spina, R. Gaál, L. Forró and E. Horváth, *J. Phys. Chem. Lett.*, 2014, **5**, 2488.

59 A. Van De Walle and G. Ceder, *Rev. Mod. Phys.*, 2002, **74**, 11.

60 A. Manzoor, S. Pandey, D. Chakraborty, S. R. Phillpot and D. S. Aidhy, *npj Comput. Mater.*, 2018, **4**, 1.

61 F. Musil, S. De, J. Yang, J. E. Campbell, G. M. Day and M. Ceriotti, *Chem. Sci.*, 2018, **9**, 1289.

62 A. Anelli, E. A. Engel, C. J. Pickard and M. Ceriotti, *Phys. Rev. Mater.*, 2018, **2**, 103804.

63 A. Pulido, L. Chen, T. Kaczorowski, D. Holden, M. A. Little, S. Y. Chong, B. J. Slater, D. P. McMahon, B. Bonillo, C. J. Stackhouse, A. Stephenson, C. M. Kane, R. Clowes, T. Hasell, A. I. Cooper and G. M. Day, *Nature*, 2017, **543**, 657.

64 J. Wiktor, U. Rothlisberger and A. Pasquarello, *J. Chem. Phys. Lett.*, 2017, **8**, 5507.

65 G. M. Dalpian, X.-G. Zhao, L. Kazmerski and A. Zunger, *Chem.*

Mater., 2019, **31**, 2497.

66 X. Gong, O. Voznyy, A. Jain, W. Liu, R. Sabatini, Z. Pio-ntkowski, G. Walters, G. Bappi, S. Nokhrin, O. Bushuyev *et al.*, *Nature Mater.*, 2018, **17**, 550.