Design rules for optimizing quaternary mixed-metal chalcohalides

Pascal Henkel
Department of Applied Physics, Aalto University, P.O. Box 11000, FI-00076 AALTO, Finland

Jingrui Li
State Key Laboratory for Manufacturing Systems Engineering; Electronic Materials Research Laboratory, Key Laboratory of the Ministry of Education; and School of Electronic Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China

Patrick Rinke*
Physics Department, Technical University of Munich, 85748 Garching, Germany; Atomistic Modelling Center, Munich Data Science Institute, Technical University of Munich, 85748 Garching, Germany; Munich Data Science Institute, Technical University of Munich, 80333 Munich, Germany; and Department of Applied Physics, Aalto University, P.O. Box 11000, FI-00076 AALTO, Finland

(Received 9 May 2025; accepted 29 September 2025; published 13 November 2025)

Quaternary mixed-metal $M(\mathrm{II})_2M(\mathrm{III})Ch_2X_3$ chalcohalides are an emerging material class for photovoltaic absorbers that combines the beneficial optoelectronic properties of lead-based halide perovskites with the stability of metal chalcogenides. Inspired by the recent discovery of lead-free mixed-metal chalcohalides (MMCHs) materials, we utilized a combination of density functional theory and machine learning to determine compositional trends and chemical design rules in the lead-free and lead-based materials spaces. We explored a total of 54 $M(\mathrm{II})_2M(\mathrm{III})Ch_2X_3$ materials with $M(\mathrm{II})$ = Sn, Pb; $M(\mathrm{III})$ = In, Sb, Bi; Ch = S, Se, Te; and X = Cl, Br, I per phase (*Cmcm*, $Cmc2_1$, and $P2_1/c$). The $P2_1/c$ phase is the equilibrium phase at low temperatures, followed by $Cmc2_1$ and *Cmcm*. The fundamental band gaps in *Cmcm* and $Cmc2_1$ are smaller than those in $P2_1/c$, but direct band gaps are more common in *Cmcm* and $Cmc2_1$. The effective electron masses in $P2_1/c$ are significantly larger than in *Cmcm* and $Cmc2_1$, while the effective hole masses are nearly the same across all three phases. Using random forest regression, we found that the two electron acceptor sites [Ch and X] are crucial in shaping the properties of MMCH compounds. Furthermore, the electron donor sites [M(II) and M(III)] can be used to fine-tune the material properties to desired applications. These design rules enable precise tailoring of MMCH compounds for a variety of applications.

DOI: 10.1103/qjwt-29w9

I. INTRODUCTION

Photovoltaic (PV) technologies are ideal for providing clean, affordable, and secure energy, facilitating a shift toward sustainable energy production. Continuous development of PV materials is essential to enhance power conversion efficiency (PCE), extend device longevity, and reduce costs, thereby promoting the widespread adoption of solar cells. Metal halide perovskites have emerged as promising candidates [1–3]. Specifically, lead halide perovskites (LHPs) offer cost-effective fabrication, advantageous optoelectronic properties, and high defect tolerance [4–6]. However, the commercial viability of LHPs is challenged by lead toxicity and their moderate long-term stability in air [5,7–9]. Low toxicity alternatives to LHPs include Sn^{2+}-, Sb^{3+}-, or Bi^{3+}-based perovskites. However, these materials face challenges such as high defect densities and oxidation in air (Sn^{2+} to Sn^{4+}) [10]. Metal chalcogenides based on Pb^{2+}, Cd^{2+}, or Sb^{3+} [11–14] have emerged as promising LHP alternatives due to their high absorption cross-sections and tunable band gaps. Stable solar cells have already been demonstrated for these materials, albeit with modest efficiencies [3,11,13,15].

Perovskite-inspired quaternary mixed-metal chalcohalides (MMCHs) that combine halide perovskite (ABX_3) and metal chalcogenide (MCh_2) building blocks are a promising semiconductor material class for PV applications [16–18]. Their stoichiometry is $M(\mathrm{II})_2M(\mathrm{III})Ch_2X_3$ (or $A_2BCh_2X_3$ in short) [17], where the M(II) and M(III) sites are occupied by metals with ns^2 lone-pair electron(s), the Ch site by chalcogens, and the X site by halogens. The properties of both building blocks may combine synergistically in MMCHs as evidenced by a high defect tolerance owing to the strong dielectric screening (due to ns^2 lone-pair electrons), the resultant low capture cross-sections of defects, and dispersive valence and conduction band edges [7,8,16]. Also, the strong metal-chalcogen bond, which is formed in MMCH materials between the bivalent chalcogenide anions and both M(II) and M(III) metal cations, can overcome the known stability challenges of LHPs [5,16]. In addition, the strength of the Sn-Ch bond should

*Contact author: patrick.rinke@tum.de

prevent the oxidation from Sn^{2+} to Sn^{4+} in lead-free MMCH-based compounds. As a result, $Sn_2M(\mathrm{III})Ch_2X_3$ films can be stable in humid conditions when synthesized under reduction conditions, as verified by experimental x-ray photoelectron spectroscopy [16,17]. Recently, Nie *et al.* [16] achieved a PCE of 4.04% for a $Sn_2SbS_2I_3$-based single-junction solar cell, which highlights the PV potential of MMCHs. That 4.04% is a promising start, considering that perovskite solar cells initially showed a PCE of 3.8% [19] and now exceed 26% [3].

In this work, we focus on the MMCH materials space and apply density functional theory (DFT) and machine learning to assess its potential for PV applications. Recently, we explored solely the lead-free MMCH materials with DFT and discovered 24 materials [20], bringing the total up to 27 (including the already known $Sn_2SbS_2I_3$ [16,17,21–23], $Sn_2SbSe_2I_3$ [24], and $Sn_2BiS_2I_3$ [25]) compounds. Now we extend our study to the lead-based $Pb_2M(\mathrm{III})Ch_2X_3$ materials space, reporting 25 previously unstudied compounds alongside the two known materials $Pb_2SbS_2I_3$ [26–28] and $Pb_2BiS_2I_3$ [25,28], which had been the only Pb-based representatives reported to date. We compute the equilibrium properties of the 27 lead-based counterparts to the tin compounds of our previous study with DFT. We then apply a random forest (RF) machine-learning model and the Shapley additive explanations (SHAP) analysis to discover chemical trends for lead-free as well as lead-based perovskite-inspired quaternary MMCHs.

Our goal is to go beyond a purely DFT-based analysis by leveraging a combination of RF models and SHAP analysis to identify and understand the root of chemical trends of the four atomic sites and their elements on material properties for a total of 54 lead-free and lead-based MMCH compounds. To cover a broad range of MMCH materials, we considered Sn^{2+} and Pb^{2+} (both with ns^2 lone-pair electrons) for the M(II) site, and Sb^{3+}, Bi^{3+} (both with ns^2 lone-pair electrons), and In^{3+} (with a $5d^{10}6s^0$ valence electron configuration) for the M(III) site. Additionally, we included S^{2-}, Se^{2-}, and Te^{2-} for the *Ch* site, and Cl^-, Br^-, and I^- for the *X* site. We considered three phases (*Cmcm*, $Cmc2_1$, and $P2_1/c$), in which MMCH materials can crystallize to ensure structural diversity. This allows us to assess the dominance of electron acceptor sites [*Ch* and *X*] over electron donor sites [M(II) and M(III)]. An additional objective is to determine whether and how the chemical trends depend on the material parameters, such as the consistent decrease in formation energies (thermodynamic stabilities), band gaps, as well as effective electron and hole masses from Cl via Br to I. Finally, we will investigate the effect of the MMCH structure and phase on the chemical trends.

The structure of this article is as follows: In Sec. II, we outline our computational workflow, starting with data generation using DFT, followed by the creation of appropriate RF models to examine the influence of atomic sites, and concluding with an analysis of the impact of the elements on material properties using SHAP. Section III reports the effects of atomic sites and elements on material properties. Section IV discusses the identified effects and design rules on one hand within the MMCH materials space and on the other hand in terms of chemical origin. Finally, Sec. V provides a summary of our findings.

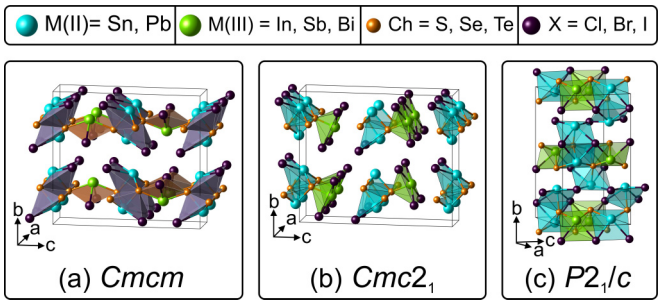

FIG. 1. Illustration of different phases: (a) *Cmcm*, (b) $Cmc2_1$, and (c) $P2_1/c$ in MMCH structures. The respective M(II) coordination polyhedrons are depicted in cyan and the M(III) coordination polyhedrons in green. The coordinates of each phase were taken from the Materials Project [30] and were visualized with VESTA [31].

II. MODELS AND COMPUTATIONAL DETAILS

We used a combination of different methods to study trends in the MMCH materials space. Here, we describe how we calculated the material properties (formation energy, fundamental band gap, optical band gap, and effective electron and hole masses) for different MMCHs crystal structures by means of DFT. Next, we outline how we determined the influence of atomic sites [M(II), M(III), *Ch*, and *X*] on the materials properties with RF regression models. Finally, we present details of the SHAP analysis to assess the impact of different elemental substitutions on the materials properties.

A. $M(\mathrm{II})_2M(\mathrm{III})Ch_2X_3$ structures

The structural framework of $M(\mathrm{II})_2M(\mathrm{III})Ch_2X_3$ materials consists of four distinct atomic sites (see Fig. 1): the two metal sites M(II) and M(III), the chalcogen site *Ch*, and the halogen site *X*. Both metal sites are occupied by ns^2 lone-pair metals, except In for M(III), serving as electron donors, while the *Ch* and *X* sites function as electron acceptors.

MMCH materials can crystallize, as revealed by x-ray diffraction measurements and DFT calculations, in three different phases: *Cmcm* (#63), $Cmc2_1$ (#36), and $P2_1/c$ (#14) [17,20–28], see Fig. 1. Throughout the manuscript, the three MMCH structures are distinguished by their respective space group labels. The orthorhombic *Cmcm* phase is prevalent at room temperature [21,24–28] and transitions to a monoclinic $P2_1/c$ phase <100 K, as demonstrated by Doussier *et al.* [27] for $Pb_2SbS_2I_3$. DFT calculations for $Sn_2SbS_2I_3$ suggest that the *Cmcm* structure is thermodynamically unstable at 0 K due to imaginary phonon modes and can be viewed as an average of the more energetically favorable, lower symmetry $Cmc2_1$ phase [17]. *Ab initio* molecular dynamic simulations further demonstrated a transition from the *Cmcm* to the $Cmc2_1$ phase at 500 K [17], highlighting the entropic stabilization effect of the *Cmcm* phase. Additionally, the $P2_1/c$ phase is energetically favorable over *Cmcm* and $Cmc2_1$ [22]. A more detailed description of the structural characteristics of MMCH materials crystallizing in the three phases (*Cmcm*, $Cmc2_1$, and $P2_1/c$) is given in Sec. SM 1 in the Supplemental Material (SM) [29].

The initial structures for all MMCHs have been generated based on known materials listed in the Materials Project [30].

For the *Cmcm* and $Cmc2_1$ phases, we used the corresponding $Sn_2SbS_2I_3$ structures (Refs. No. mp-561134 and No. mp-1219046) and, for $P2_1/c$ phases, the corresponding $Pb_2SbS_2I_3$ structure (Ref. No. mp-578882). We carried out substitutions on the *M*(II), *M*(III), *Ch*, and *X* sites to generate a total of 54 MMCHs per phase.

B. DFT calculations

We determined the material properties (formation energy, fundamental band gaps, optical gaps, and effective electron and hole masses) of all 54 MMCHs in the three phases for our RF-based analysis with DFT. We performed spin-unpolarized DFT calculations with the all-electron, numeric atom-centered orbital code FHI-AIMS [32–37]. We chose the Perdew-Burke-Ernzerhof exchange-correlation functional for solids (PBEsol) [38,39] due to its computational efficiency and its good agreement with experimental lattice constants for various halide perovskites [40–44] and the five known MMCHs. On average, PBEsol underestimates the lattice constants for the known MMCHs by ∼1.7 % for *Cmcm*, ∼1.6 % for *Cmcm*, and ∼0.4 % for $P2_1/c$. Atomic geometries were relaxed with the Broyden-Fletcher-Goldfarb-Shanno algorithm [45] in two steps. We first preoptimized each structure with light real-space grid settings and a tier-1 basis and then refined the geometry using tight settings and a tier-2 basis. The structures were relaxed until all forces acting on the atoms were $<5 \times 10^{-3}$ eV Å^{-1}. The electronic self-consistency convergence threshold was set to 1×10^{-6} eV. A Gaussian broadening of 0.01 eV was applied to all electronic occupations, and relativistic effects were included based on the zeroth-order regular approximation [32,46]. In addition to PBEsol geometry optimization, we performed single point calculations for the band structures, the absorption spectra, and the density of state (DOS) with the range-separated hybrid Heyd-Scuseria-Ernzerhof (HSE06) functional (with 25% exact exchange) [47–49] including spin-orbit coupling [50] and a tier-2 basis. The effective masses were extracted from the band structure with a quadratic least-squares fit to the respective band edges assuming a parabolic dispersion with the implementation of Whalley in the EFFMASS package [51]. In addition, the absorption spectra were calculated based on the linear macroscopic dielectric tensor [52] within the independent particle approximation.

We used 16-atom unit cells for the *Cmcm* (#63) [see Fig. 1(a)] and $Cmc2_1$ (#36) [see Fig. 1(b)] phases and a 32-atom unit cell for the $P2_1/c$ (#14) [see Fig. 1(c)] phase. For PBEsol geometry optimization, we sampled the Brillouin zone with a Γ-centered $11 \times 11 \times 3$ *k*-point mesh for the *Cmcm* and $Cmc2_1$ phases and $6 \times 3 \times 5$ for $P2_1/c$. For the HSE06 band structure and absorption spectra calculations, we increased the grids to Γ-centered $16 \times 16 \times 4$ for *Cmcm* and $Cmc2_1$ and $9 \times 4 \times 8$ for $P2_1/c$.

The complex nature of MMCH compounds makes a rigorous stability analysis highly demanding, as it requires knowledge of all competing phases within the quaternary phase space. As a computationally feasible DFT-based approximation, we estimate the thermodynamic stability (at 0 K) by calculating the formation energy with respect to elemental decomposition of an individual MMCH compound as

$$E_{\text{form}}[M(\text{II})_2M(\text{III})Ch_2X_3] = E_{\text{tot}}[M(\text{II})_2M(\text{III})Ch_2X_3] - \sum_i x_i\mu_i. \quad (1)$$

$E_{\text{tot}}[M(\text{II})_2M(\text{III})Ch_2X_3]$ is the total energy of $M(\text{II})_2M(\text{III})Ch_2X_3$, x_i the number of atoms of the *i*th element, and μ_i the corresponding chemical potential. The upper limit for the chemical potential is given by $\mu_i \leqslant E_{\text{tot}}(i\text{th element})$, i.e., the total energy per atom in the most stable pure phase [53]. We used the following elemental compounds with the same computational parameters but adjusted Γ-centered *k*-point meshes: Sn $I4_1/amd$ (#141), Pb $Fm\bar{3}m$ (#225), In $I4/mmm$ (#139), Sb $R\bar{3}m$ (#166), Bi $R\bar{3}m$ (#166), S $Fddd$ (#70), Se $P2_1/c$ (#14), and Te $P3_121$ (#152), as well as Cl, Br, and I in the gas phase as X_2.

In the interest of open materials science [54], we make all relevant data publicly available; see Ref. [55] for lead-free and Ref. [56] for lead-based $M(\text{II})_2M(\text{III})Ch_2X_3$ materials.

C. RF regression modeling

We used RF regression of the DFT data as a means to determine the impact of each atomic site [*M*(II), *M*(III), *Ch*, and *X*] on different material properties. RF is a nonlinear and nonparametric ensemble learning method composed of different decision trees [57]. For regression models, as used in this study, the predicted target is averaged over the individual decision trees, which increases prediction accuracy and reduces overfitting [58].

For each phase, we first created a dataset that contains the material properties for each MMCH material. As a descriptor, we use the atomic numbers of all elements within each MMCH, e.g., $Sn_2SbS_2I_3$ is represented by the array [50, 51, 16, 53]. Next, we trained an RF regression model for each material property using the atomic numbers as input and the respective material property as the target value. To determine the optimal RF hyperparameters, we used the cross-validated grid-search (GRIDSEARCHCV) method implemented in SCIKIT-LEARN [59]. A 75%/25% split for the training/test data was used to determine the performance of each hyperparameter combination. The mean absolute error (MAE) and root mean square error between predicted and DFT-calculated material properties are calculated for each hyperparameter set to assess the performance. In addition, a 10-fold cross-validation was carried out to ensure the robustness of each hyperparameter set. The best performing hyperparameter set is applied to analyze the entire dataset. The performance of the RF models for the formation energy, (fundamental) band gap, and effective electron and hole masses for different phases is presented in Fig. SM 1 in the SM [29]. In addition, we trained RF models for the lowest direct band gaps and optical band gaps, see Figs. SM 2 and SM 3 in the SM [29] for their individual performance. We used the final trained RF models to determine the importance of the atomic sites for each material property. Each atomic site is one feature of our input space, and we used the built-in feature importance tools of SCIKIT-LEARN.

Additionally, we tested other descriptors such as electronegativity (EN), electron affinity (EA), ionization energy (IE), as well as atomic, van der Waals, and covalent radii. In the past, descriptors based on such tabulated properties have been successfully used to predict material properties for a variety of material systems, such as N-based semiconductors [60], ABX_3 [61], and $A'A''B'B''X_6$ double perovskites [62]. Gladkikh *et al.* [61] and Pilania *et al.* [62] used such element-derived descriptors in combination with the kernel ridge regression model to predict the band gaps of ABX_3 and $A'A''B'B''X_6$ double perovskites. However, we found, for MMCH materials, that the atomic numbers provided the best performance. Furthermore, we tested the sure independence screening and sparsifying operator (SISSO) approach as implemented in SISSO++ [63], but the results were not as good as those of the RF.

D. SHAP analysis

To examine the impact of elements on material properties within the MMCH materials space, we analyzed the resulting RF models using the SHAP method [64]. SHAP, based on game theory and an additive feature attribution approach, explains machine-learning outputs. Its purpose is to identify the contribution (SHAP value) of each feature (in our case, of each element within a MMCH compound) to the predicted target value. Various methods can approximate SHAP values. We here use TREESHAP, specifically designed for RFs [65,66], to assess the influence of the features (atomic numbers of elements within the four atom sites of MMCH compounds) on the material parameters. For each RF model, a SHAP base value is first determined, representing the mean of the predicted values. Then the contribution of each feature to the predicted target value is calculated relative to the SHAP base value. In this study, we use the SHAP implementation by Lundberg [67].

In this study, we exclusively use the mean SHAP values for each element type to enhance interpretability. Since the contribution of each element varies depending on the specific MMCH compound, minor deviations in element type contributions are observed. For instance, the SHAP contribution of Cl to the formation energy is -75 meV in $Sn_2SbS_2Cl_3$ and -77 meV in $Sn_2BiTe_2Cl_3$. Therefore, we report the mean SHAP contribution along with the corresponding variance, which characterizes the distribution around the mean value.

III. RESULTS

We examine the impact of atomic sites and elements on the formation energy, fundamental band gap, and effective electron and hole masses of MMCH materials in their *Cmcm* phase. This analysis allows us to identify material property trends across the MMCH materials space. The results for the $Cmc2_1$ and $P2_1/c$ phases are shown in Figs. SM 4 and SM 5 in the SM [29] for the formation energy, Figs. SM 6 and SM 7 in the SM [29] for the fundamental band gap, and Figs. SM 14 to SM 17 in the SM [29] for the effective masses of electrons and holes. Additionally, Table SM I in the SM [29] lists all material parameters for all 54 MMCH materials across all phases.

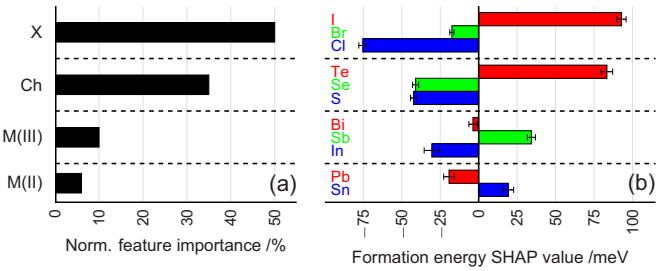

FIG. 2. Summary of the feature importance of the formation energy in the RF model for the *Cmcm* phase: (a) normalized atom site importance for the predicted formation energy in percent; (b) impact of elements on the *M*(II), *M*(III), *Ch*, and *X* sites measured in terms of their mean SHAP values on the predicted base formation energy value in meV. The error bars denote the standard deviation for each mean SHAP value. Since the formation energy is negative, negative mean SHAP values indicate a stabilization of the material and positive SHAP values a destabilization.

A. Formation energies

We have previously determined that lead-free MMCH materials are stable with respect to elemental decomposition [20]. Here, we show that this stability also extends to the lead-based counterparts as evidenced by their negative formation energies, listed in Table SM I in the SM [29]. Therefore, all 54 MMCH materials are stable-independent of their phase. We observe that the $P2_1/c$ phase is most prevalent (i.e., has the highest absolute formation energies) followed by $Cmc2_1$ then *Cmcm*, see Table SM I in the SM [29].

Figure 2(a) shows the atomic site importance of the formation energy for all 54 MMCH materials in the *Cmcm* phase as determined by the RF model. The *X* site has the largest influence (50%) followed by the *Ch* site (34%). In contrast, both *M*(II) (6%) and *M*(III) (10%) metal sites have only minor effects on the formation energy.

The corresponding SHAP analysis is presented in Fig. 2(b) and offers more detail on the elemental contributions to the formation energy. SHAP reveals that Cl and Br decrease the formation energy (Cl more than Br) and I increases it. On the *Ch* site, S and Se decrease the formation energy (almost by the same amount on average), while Te increases it. For Bi, In, and Pb, we also observe decreases, while Sb and Sn increase the formation energy. In correspondence with the feature importance, the formation energy change caused by the metal atoms on *M*(II) and *M*(III) sites is less than that of the *X* and *Ch* site elements. We observed similar behavior for the $Cmc2_1$ and the $P2_1/c$ phases, as shown correspondingly in Figs. SM 4 ($Cmc2_1$) and SM 5 ($P2_1/c$) in the SM [29].

B. Fundamental band gaps

MMCH materials are mainly semiconductors, with the exception of $Sn_2InTe_2Br_3$ and $Sn_2BiTe_2I_3$ in the *Cmcm* phase (see Table SM I in the SM [29]). The majority of the computed, fundamental band gaps are indirect. Only ∼33% of lead-free and ∼40% of lead-based MMCH compounds feature a direct fundamental band gap in their *Cmcm* and $Cmc2_1$ phases. Also, the vast majority of MMCH compounds in the $P2_1/c$ phase have an indirect fundamental

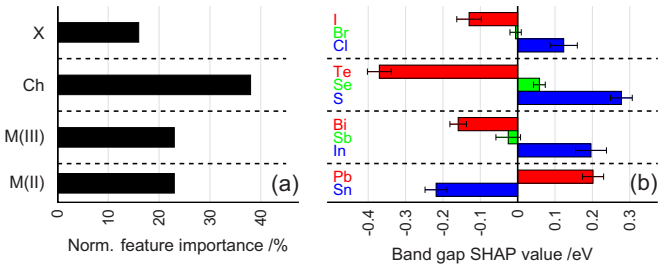

FIG. 3. Summary of the feature importance of the fundamental band gap in the RF model for the *Cmcm* phase: (a) normalized atom site importance for the predicted fundamental band gap in percent; (b) impact of elements on the *M*(II), *M*(III), *Ch*, and *X* sites quantified in terms of their mean SHAP values for the predicted base fundamental band gap in eV. The error bars denote the standard deviation for each mean SHAP value.

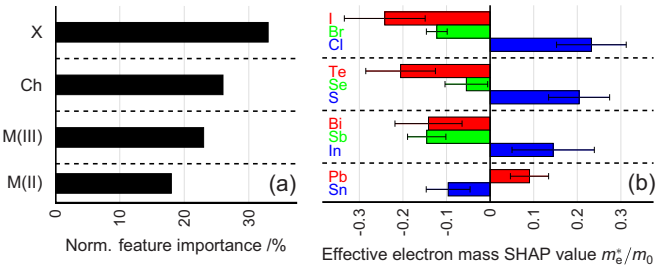

FIG. 4. Summary of the feature importance of the effective electron mass in the RF model for the *Cmcm* phase: (a) normalized atom site importance for the predicted electron mass in percent; (b) impact of elements on the *M*(II), *M*(III), *Ch*, and *X* sites quantified in terms of their mean SHAP values for the predicted base effective hole mass value per m_0 (free electron mass). The error bars denote the standard deviation for each mean SHAP value.

band gap. Only ∼7% of lead-free and ∼11% of lead-based $M(\mathrm{II})_2M(\mathrm{III})Ch_2X_3$ compounds exhibit a direct fundamental band gap. The average values of the fundamental band gap in the *Cmcm*, $Cmc2_1$, and $P2_1/c$ phases are ∼1.2, ∼1.1, and ∼1.6 eV, respectively.

Figure 3(a) presents the same feature importance analysis done for the formation energy in the previous section now for the fundamental band gap in the *Cmcm* phase. The band gap is mainly influenced by the chalcogen site (38%), followed by both metal sites [23% for both *M*(II) and *M*(III); see Fig. 3(a)]. In contrast with the formation energy, halogens have the smallest (16%) influence on the fundamental band gap. The corresponding SHAP analysis in Fig. 3(b) reveals that, on the *X* site, Cl and Br decrease the fundamental band gap, while I increases it. Similarly, on the *Ch* site, S and Se increase the fundamental band gap, and Te decreases it. For both metal sites, In and Pb lead to a band gap increase, while Bi, Sb, and Sn reduce the gap.

Again, we observed similar behavior for the $Cmc2_1$ and $P2_1/c$ phases (see Figs. SM 6 and SM 7 in the SM [29], respectively). Furthermore, we obtained similar results for an RF model trained on the lowest direct band gap for each material instead of the lowest fundamental band gap (which includes a mix of direct and indirect band gaps). The results for the lowest direct band gap are presented in Figs. SM 8 (*Cmcm*), SM 9 ($Cmc2_1$), and SM 10 ($P2_1/c$) in the SM [29].

We also calculated the absorption spectra for all 54 MMCH compounds in the three phases in the independent particle approximation. We then determined the optical band gaps as onset of absorption from the corresponding Tauc plots. In our previous work [20], we had demonstrated that the optical band gap lies on average 0.29 eV above the fundamental band gap for the lead-free MMCH compounds. For the lead-based systems, the difference is 0.22 eV, which drops the average for the whole MMCH materials system to 0.26 eV. Figures SM 11 (*Cmcm*), SM 12 ($Cmc2_1$), and SM 13 ($P2_1/c$) in the SM [29] show the corresponding feature importance analysis for the optical band gaps for all three phases. We observe that also the optical band gap is determined by the *Ch*, *M*(II), and *M*(III) sites.

C. Effective electron and hole masses

Table SM I in the SM [29] also lists the effective electron and hole masses for the investigated MMCH materials. Unlike for the formation energy (Sec. III A) and the fundamental band gap (Sec. III B), the effective masses behave differently across the three different phases. For $P2_1/c$, the average electron (1.42 m_0 with $m_0 = 9.109 \times 10^{-31}$ kg denoting the free electron mass) and hole (1.46 m_0) masses are almost the same, while for *Cmcm* and $Cmc2_1$, electrons are lighter than holes ($m_e^{Cmcm} = 0.86\ m_0$, $m_h^{Cmcm} = 1.37\ m_0$, $m_e^{Cmc2_1} = 0.93\ m_0$, and $m_h^{Cmc2_1} = 1.38\ m_0$).

The corresponding feature importance plots are shown in Figs. 4(a) and 5(a) for *Cmcm*, Figs. SM 14(a) and SM 16(a) in the SM [29] for $Cmc2_1$, and Figs. SM 15(a) and SM 17(a) in the SM [29] for $P2_1/c$. They reveal that the *M*(II) site has the lowest impact on the effective electron and hole masses regardless of the phase. The order of importance of the other three sites varies for different phases. For the *Cmcm* phase, the *X* site has the largest influence, although all three sites are closer together in their effect (between 23% and 33%) than they were for the formation energy and the fundamental band gap.

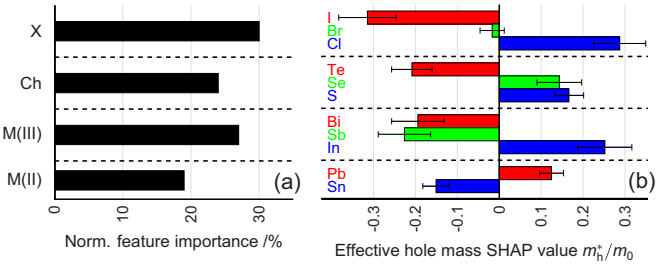

FIG. 5. Summary of the feature importance of the effective hole mass in the RF model for the *Cmcm* phase: (a) normalized atom site importance for the predicted effective hole mass in percent; (b) impact of elements on the *M*(II), *M*(III), *Ch*, and *X* sites quantified in terms of their mean SHAP values for the predicted base effective hole mass value per m_0. The error bars denote the standard deviation for each mean SHAP value.

TABLE I. Summary of the impact of the elements on the formation energy, band gap, as well as effective electron and hole masses within the atom sites. Symbols indicate the relative importance of the atom site. ++: largest impact; +: second-largest impact; −: second-smallest impact; −−: smallest impact. Deviating trends within the space groups (for the effective electron and hole masses) are indicated by (1) *Cmcm*, (2) $Cmc2_1$, and (3) $P2_1/c$; otherwise, the listed element trends are identical in all three space groups.

	M(II) site	*M*(III) site	*Ch* site	*X* site
Formation energy	−−: Pb > Sn	−: In > Bi > Sb	+: S ≃ Se > Te	++: Cl > Br > I
Band gap	−: Pb > Sn	+: In > Sb > Bi	++: S > Se > Te	−−: Cl > Br > I
Eff. electron mass	−−: Pb > Sn	−: (1) In > Bi ≃ Sb, (2) In > Bi > Sb, (3) Bi > Sb > In	+: (1) S > Se > Te, (2) Te > Se > S, (3) S > Se > Te	++: Cl > Br > I
Eff. hole mass	−−: Pb > Sn	+: (1) In > Bi ≃ Sb, (2) Sb > Bi > In, (3) Sb > In ≃ Bi	−: (1) S ≃ Se > Te, (2) S > Se ≃ Te, (3) Se > S > Te	++: Cl > Br > I

Our SHAP analysis [see Figs. 4(b) and 5(b) for *Cmcm*] reveals that, in all three phases, the halogens I and Br decrease the effective electron and hole masses, whereas Cl increases them. Similarly, on the *M*(II) site, Pb increases and Sn decreases the effective masses. For the *Cmcm* phase, Te, Bi, and Sb have a decreasing effect, while S and In always increase both effective masses. Only Se changes its role, providing a small decrease of the electron but a moderate increase of the hole masses. For the other two phases, these elemental effects are slightly different and are summarized in Table I as well as Figs. SM 14(b) and SM 16(b) ($Cmc2_1$) in the SM [29] and Figs. SM 15(b) and SM 17(b) ($P2_1/c$) in the SM [29].

IV. DISCUSSION

First, our analysis of the formation energy in Sec. III A shows that MMCH stability (which can be estimated from the formation energy, as more negative means more stable) is mainly influenced by the two electron acceptor sites [*Ch* and *X*] with the two metal sites [*M*(II) and *M*(III)] playing only a minor role, see Fig. 2(a). From the corresponding SHAP analysis in Fig. 2(b), we can deduce the following trends, which could be used as materials design rules: The formation energy decreases (and thus, the stability increases) for halogens on the *X* site along Cl > Br > I and for chalcogens on the *Ch* site along S ≃ Se > Te. On the *M*(III) site, the formation energy decreases along In > Bi > Sb. Lastly, lead-based MMCHs are more stable than their Sn-based counterparts (Pb > Sn), which is in good agreement with theoretical [68] and experimental [69–71] observations for Sn- and Pb-based perovskites.

These design rules can be rationalized by the strength of the metal-halide and the metal-chalcogen bonds. The EN of the elements, i.e., the strength with which an element can attract valence electrons, can be used as an indicator for the bond strength. Since the *X* and *Ch* sites in MMCH compounds are both electron acceptors, the bond strength increases with EN, as both sites seek to attract electron density in the bond. Therefore, for halogens (Cl > Br > I), the bond-strength progresses from metal-Cl (strongest) to metal-I (weakest) following the EN trend: Cl (3.16) > Br (2.96) > I (2.66) on Pauling's scale [72]. In the chalcogen family, we have S (2.58) ≃ Se (2.55) > Te (2.10), which explains the almost identical effect of S and Se on the formation energy and the weakest stability for the Te-based systems. Conversely, both metal sites in MMCHs are electron donors. With a lower EN, they can transfer their valence electrons more easily to the electron acceptor sites for creating stronger chemical metal-*Ch* and metal-*X* bonds in the process. An inverse relation between the EN of the metal sites and stability ensues, which we see reflected in, e.g., the *M*(II) site elements In (1.78) < Bi (1.90) < Sb (2.05). Also, Pb-based MMCHs are more stable than their Sn-based analogs, in line with the respective EN values (1.80 and 1.96).

Figure 2 shows that the *X* site element has the largest effect on the MMCH formation energy, followed by the *Ch*, *M*(III), then *M*(II) sites. The difference in impact can also be rationalized by the differences in the respective EN values. For the dominant *X* site, they span the largest range (0.50), followed by the *Ch*, *M*(III), then *M*(II) sites (0.48, 0.27, and 0.16, respectively).

Secondly, we found that the fundamental band gap (the same applies to the lowest direct and optical band gaps, see Sec. III B) are mainly influenced by the *Ch* site and both metal sites. Furthermore, we showed that the halogens have the smallest effect on the fundamental band gap [see Fig. 3(a)], a trend opposite to the formation energy. These trends can be traced back to the orbital character of the halogens, chalcogens, and both ns^2 lone-pair metals. Halogen *p* orbitals are generally more localized than those of the chalcogens due to their higher EN and smaller atomic size [72]. This increased localization leads to a weaker orbital overlap and thus hybridization with the ns^2 lone-pair orbitals of the metal sites [73]. In contrast, chalcogen *p* orbitals are more diffused, thus increasing the orbital overlap and the hybridization with the ns^2 lone-pair orbitals of the metal sites and subsequently also the impact on the valence band maximum (VBM). The chalcogens therefore have a larger effect on the fundamental band gap than the halogens. Conversely, both *M*(II) and *M*(III) cations have a similar orbital character [$5s^2$ lone-pair and $5p$ empty orbitals for Sn and Sb, as well as $6s^2$ lone-pair and $6p$ empty orbitals for Pb and Bi) except for In (no lone pair)], and therefore, their contribution to the fundamental band gap is expected to be similar.

The effects of the orbital analysis are reflected in the DOS near the VBM and the conduction band minimum (CBM). This is illustrated in Fig. 6 for lead-based and lead-free M(II)$BiSe_2Br_3$ compounds as examples. Additionally, Fig. SM 18 in the SM [29] shows the corresponding $Cmc2_1$ and $P2_1/c$ phases. The chalcogens (orange) provide the largest contribution to the total DOS (dashed blue) near the VBM, primarily through the closed Se $4p^6$ shell. The contribution of the *M*(II) metals (cyan) is caused by the strong ns^2 antibonding character, whose important effect on

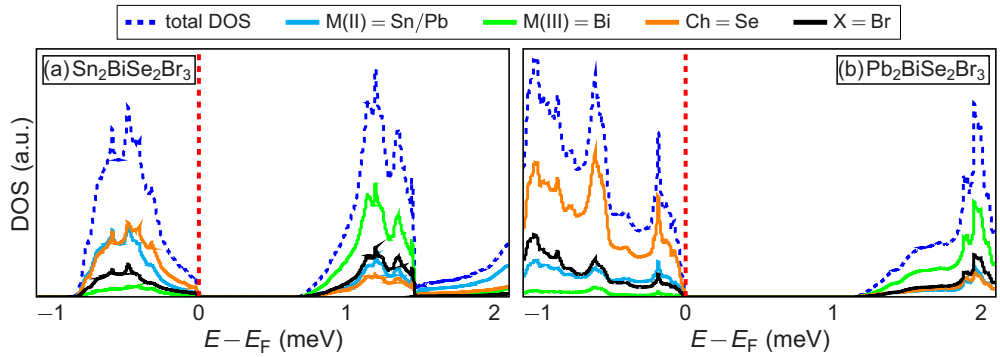

FIG. 6. DOS for (a) $Sn_2BiSe_2Br_3$ and (b) $Pb_2BiSe_2Br_3$ at the *Cmcm* phase.

the electronic structure of halide perovskites has been reported previously [74–76]. In contrast, the *M*(III) metals (light green) provide the largest contribution to the total DOS near the CBM via their empty *p* orbitals. Notably, the halogens (black) contribute significantly to the valence band (although less than the chalcogens, primarily due to the higher EN of halogens) but only minimally to the CBM. In summary, the hole states (near the VBM) are primarily determined by the chalcogens and the *M*(II) metals, while the free electron states (near the CBM) are mainly influenced by the *M*(III) metals. The fundamental band gap in MMCHs is therefore strongly affected by the chalcogen-ns^2 lone-pair bonds. The corresponding band structures are shown in Fig. SM 19 in the SM [29]. This character is different to halide perovskites for which the monovalent metal (or organic) cation provides negligible contribution to both band edges. Their impact on the electronic structure is indirect (via the crystal structure). In this regard, MMCHs offer more options for tailoring the electronic structure than halide perovskites.

The design rules we derive from the SHAP analysis [see Fig. 3(b)] for the band gap are generally similar to the ones for the formation energy, i.e., the larger EN for the nonmetals and the smaller EN for the metals, the higher the stability and the larger the band gap. For *M*(III), we observe an exception. The reverse EN trend In > Bi > Sb for the formation energy does not hold for the band gap (In > Sb > Bi). The absence of the ns^2 lone pair in In^{3+} gives rise to a larger band gap in In-based MMCHs, breaking the trend. Conversely, the strong spin-orbital coupling induced by Bi significantly reduces the band gaps of Bi-based compounds to lower than those of the isostructural Sb-based analogs, a trend opposite to that observed for Pb vs Sn.

The observation that the fundamental band gap in MMCH compounds decreases from Cl via Br to I and from Pb to Sn is also found in ABX_3 perovskites [77–81]. Furthermore, Nie *et al.* [16] and Islam *et al.* [25] reported a decrease of the optical band gap from $Sn_2SbS_2I_3$ (1.41 eV) to $Sn_2BiS_2I_3$ (1.22 eV), in agreement with our observed design rule for the *M*(III) site (Sb > Bi). This agrees well with our calculated values 1.547 and 1.390 eV, see Table SM I in the SM [29]. Moreover, we observe similar band gap trends in chalcogenides and sesquioxides for the *M*(III) site. Huerta-Flores *et al.* [82] measured a band gap decrease from In_2S_3 to Sb_2S_3 to Bi_2S_3 (2.10 > 1.80 > 1.70 eV) that was found also in DFT calculations by Matsumoto *et al.* [83] for oxides (As_2O_3 > Sb_2O_3 > Bi_2O_3). These comparisons illustrate the consistency of our band gap design rules with the available literature and across different materials spaces.

Thirdly, our results show that the conduction bands in MMCH materials are more curved than the valence bands, meaning smaller effective electron masses than hole masses, especially in the *Cmcm* and $Cmc2_1$ phases. The curvature at the VBM is almost the same regardless of the phase. However, small changes in the chemical nature of MMCH compounds can shift their band structure and curvature, leading to inconsistent trends in effective masses across all three phases. Despite this, we found consistent patterns for the impact of halogens and metals at the *M*(II) site on electron and hole effective masses. Specifically, the effective masses decrease in the order of Cl > Br > I for halogens and Pb > Sn for metals. These trends match well with previous studies for ABX_3 [84–86] and two-dimensional halide perovskites [87]. Ashari-Astani *et al.* [86] used tight binding and $\mathbf{k}\cdot\mathbf{p}$ theory to establish a strong linear relationship between the effective electron and hole masses and the fundamental band gap in lead-free and lead-based ABX_3 perovskites. They attributed this correlation, especially for holes, to the orbital overlap between halogens on the *X* site and metal atoms on the *B* site. We observe a similar correlation in MMCHs. Regardless of the phase, effective masses decrease in the order of Cl > Br > I for the *X* site, and Pb > Sn for the *M*(II) site, matching the fundamental band gap trends. The correlation between effective masses and band gap suggests a significant orbital contribution from the *X*-ns^2 metal bond, consistent with the explanation of Ashari-Astani *et al.* [86] for ABX_3 perovskites. However, we find no consistent correlations for the *Ch* and *M*(III) sites, which we attribute to the different chemical environments in the different phases. The increasing number of *M*(III)-*Ch* bonds per stoichiometric unit (one in $Cmc2_1$, two in *Cmcm*, and three in $P2_1/c$) gives rise to different coordination polyhedra and therefore a different environment around the *M*(III)-*Ch* bonds. Due to this difference in local environments, the effective mass trends differ across the three phases.

The fact that our purely data-driven model qualitatively agrees with the chemical intuition of a human scientist (e.g., based on the well-established EN or orbital character) is reassuring and illustrates that machine learning can find chemical trends without needing the auxiliary information of element specific attributes. While site-specific dependencies (e.g., Cl > Br > I for stability and band gap) and their rationalization

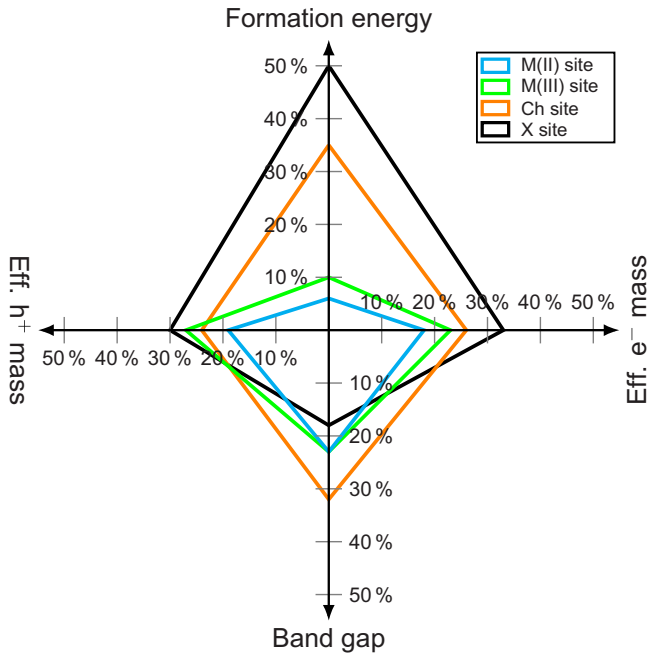

FIG. 7. Summary of atom site importance on the formation energy, band gap, as well as the effective masses of electrons and holes in *Cmcm*.

in terms of chemical descriptors like EN could have easily been deduced by human scientists, the quaternary nature of the MMCH materials space makes the deduction of chemical trends more challenging. Our data-driven models revealed the chemical trends, established their relative importance (e.g., *X* and *Ch* sites dominate the metal sites for the formation energy), and enabled the formulation of a correlation between different material parameters. This allows us to deepen our MMCH material spaces knowledge.

Finally, we can now use the atom site importance and design rules we have derived for materials design recommendations. Figure 7 summaries the atom site importance of the formation energies, fundamental band gaps, as well as of effective electron and hole masses for *Cmcm*. Also, see Figs. SM 20 ($Cmc2_1$) and SM 21 ($P2_1/c$) in the SM [29] for the other two phases.

We have demonstrated that the two electron acceptor sites [*Ch* and *X*] affect the materials properties more than the two electron donor sites [*M*(II) and *M*(III)]. Authors of future MMCH design studies can set the basic material properties with the two acceptor sites and fine-tune them for desired applications with appropriate choices for the two donor sites.

Furthermore, it is remarkable that the element impact across the atomic sites is fairly consistent across all material parameters. While a correlation between band gaps and the effective masses of electrons and holes might be expected, it is rather surprising that this relationship also extends to structural stability. This trend is particularly evident for choosing elements within the *M*(II) and *X* sites, see Table I. In contrast, the design rules vary to some extent at the *M*(III) and *Ch* sites (see Table I), leading to a more pronounced trade-off between stability and electronic properties. For example, Sb and Se reduce the stability, but $M(\mathrm{II})_2\mathrm{SbSe}_2X_3$-based materials exhibit preferable band gaps for PV applications with suitable electron effective masses, yet high hole masses. In contrast, swapping Sb for Bi increases the stability, or Se for Te decreases it, but both substitutions lower the electron and hole masses and reduce the band gaps.

Overall, this allows adjustments toward either more stable materials with wider band gaps and heavier electrons and holes or less stable materials with narrower band gaps and lighter charge electrons and holes. Specifically, $\mathrm{Pb}_2M(\mathrm{III})Ch_2\mathrm{Cl}_3$-based materials tend to be more stable with wider band gaps, whereas $\mathrm{Sn}_2M(\mathrm{III})Ch_2\mathrm{I}_3$-based compounds shift to the opposite. In this context, $\mathrm{Pb}_2\mathrm{SbSe}_2\mathrm{Cl}_3$ and $\mathrm{Sn}_2\mathrm{BiS}_2\mathrm{I}_3$ stand out, combining relatively low charge carrier masses with band gaps of ∼1.4 and 1.0 eV, respectively, making them promising candidates for further experimental and theoretical investigation.

V. CONCLUSIONS

In conclusion, in this study, we aimed at investigating the chemical trends governing material properties in perovskite-inspired quaternary MMCHs. By analyzing thermodynamic stability, band gaps, and effective electron and hole masses through a combination of DFT, RF, and SHAP, we found that the electron acceptor sites [*Ch* and *X*] play a primary role in determining MMCH characteristics, while the electron donor sites [*M*(II) and *M*(III)] provide opportunities for fine-tuning material properties for specific applications. The choice of *M*(III) and *Ch* site occupants often introduces a trade-off when optimizing stability and electronic properties, whereas *M*(II) and halogen sites tend to affect both properties in the same way. Our findings across three crystal structures (*Cmcm*, $Cmc2_1$, and $P2_1/c$) further highlight the importance of structural diversity, as $P2_1/c$ is the equilibrium phase at low temperatures with larger and mostly indirect fundamental band gaps, whereas *Cmcm* and $Cmc2_1$ exhibit smaller and more direct band gaps. While effective hole masses are similar in all three phases, electron effective masses are smaller in *Cmcm* and $Cmc2_1$ than in $P2_1/c$. These results deepen our understanding of the relationship between composition, structure, and material properties in MMCHs, opening avenues for optimizing optoelectronic applications, such as indoor and outdoor PVs. Future researchers could build on these identified chemical rules to explore design strategies for enhancing MMCH performance.

ACKNOWLEDGMENTS

The authors thank Milica Todorović, Armi Tiihonen, and Jarno A. Laakso for fruitful discussions. This study was supported by the Academy of Finland through Project No. 334532, the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 101152684, the National Natural Science Foundation of China (Grant No. 22473088), and the Natural Science Foundation of Shaanxi Province of China (Grant No. 2023-YBGY-447). P.H., J.L., and P.R. further acknowledge CSC-IT Center for Science, Finland, the Aalto Science-IT project, Xi'an Jiaotong University's HPC platform, and the

Computing Center in Xi'an of China for generous computational resources.

DATA AVAILABILITY

The data that support the findings of this article are openly available [55,56].

[1] S.-P. Feng *et al.*, Roadmap on commercialization of metal halide perovskite photovoltaics, J. Phys.: Mater. **6**, 032501 (2023).

[2] W. Zhang, G. E. Eperon, and H. J. Snaith, Metal halide perovskites for energy applications, Nat. Energy **1**, 16048 (2016).

[3] National Renewable Energy Laboratory, Best research-cell efficiencies chart, https://www.nrel.gov/pv/cell-efficiency.html.

[4] J. Seo, J. H. Noh, and S. I. Seok, Rational strategies for efficient perovskite solar cells, Acc. Chem. Res. **49**, 562 (2016).

[5] R. Nie, K. S. Lee, M. Hu, M. J. Paik, and S. I. Seok, Heteroleptic tin-antimony sulfoiodide for stable and lead-free solar cells, Matter **3**, 1701 (2020).

[6] K. J. Savill, A. M. Ulatowski, and L. M. Herz, Optoelectronic properties of tin-lead halide perovskites, ACS Energy Lett. **6**, 2413 (2021).

[7] Y.-T. Huang, S. R. Kavanagh, D. O. Scanlon, A. Walsh, and R. L. Z. Hoye, Perovskite-inspired materials for photovoltaics and beyond—From design to devices, Nanotechnology **32**, 132004 (2021).

[8] Y.-T. Huang, S. R. Kavanagh, D. O. Scanlon, A. Walsh, and R. L. Z. Hoye, Corrigendum: Perovskite-inspired materials for photovoltaics and beyond—From design to devices [*Nanotechnology* **32**, 132004 (2021)], Nanotechnology **32**, 379501 (2021).

[9] A. M. Ganose, C. N. Savory, and D. O. Scanlon, Beyond methylammonium lead iodide: Prospects for the emergent field of ns^2 containing solar absorbers, Chem. Commun. **53**, 20 (2017).

[10] J. Cao and F. Yan, Recent progress in tin-based perovskite solar cells, Energy Environ. Sci. **14**, 1286 (2021).

[11] S. H. Im, H.-j. Kim, S. W. Kim, S.-W. Kim, and S. I. Seok, All solid state multiply layered PbS colloidal quantum-dot-sensitized photovoltaic cells, Energy Environ. Sci. **4**, 4181 (2011).

[12] Y.-L. Lee and Y.-S. Lo, Highly efficient quantum-dot-sensitized solar cell based on co-sensitization of CdS/CdSe, Adv. Funct. Mater. **19**, 604 (2009).

[13] Y. C. Choi, D. U. Lee, J. H. Noh, E. K. Kim, and S. I. Seok, Highly improved Sb_2S_3 sensitized-inorganic-organic heterojunction solar cells and quantification of traps by deep-level transient spectroscopy, Adv. Funct. Mater. **24**, 3587 (2014).

[14] S. R. Kavanagh, A. Walsh, and D. O. Scanlon, Rapid recombination by cadmium vacancies in CdTe, ACS Energy Lett. **6**, 1392 (2021).

[15] V. Yadav, S. Kashyap, R. Pandey, and J. Madan, Impact of doping variation on the performance of Sb_2S_3 based solar cell using device simulations, in *2023 IEEE Devices for Integrated Circuit (DevIC)* (IEEE, Kalyani, India, 2023), pp. 52–55.

[16] R. Nie, R. R. Sumukam, S. H. Reddy, M. Banavoth, and S. I. Seok, Lead-free perovskite solar cells enabled by hetero-valent substitutes, Energy Environ. Sci. **13**, 2363 (2020).

[17] S. R. Kavanagh, C. N. Savory, D. O. Scanlon, and A. Walsh, Hidden spontaneous polarisation in the chalcohalide photovoltaic absorber $Sn_2SbS_2I_3$, Mater. Horizons **8**, 2709 (2021).

[18] K. V. Sopiha, C. Comparotto, J. A. Márquez, and J. J. S. Scragg, Chalcogenide perovskites: Tantalizing prospects, challenging materials, Adv. Opt. Mater. **10**, 2101704 (2022).

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

[20] P. Henkel, J. Li, G. K. Grandhi, P. Vivo, and P. Rinke, Screening mixed-metal $Sn_2M(\mathrm{III})Ch_2X_3$ chalcohalides for photovoltaic applications, Chem. Mater. **35**, 7761 (2023).

[21] J. Olivier-Fourcade, J. C. Jumas, M. Maurin, and E. Philippot, Mise en évidence d'un nouveau sulfoiodure d'étain et d'antimoine $Sn_2SbS_2I_3$: Étude structurale, Z. anorg. allg. Chem. **468**, 91 (1980).

[22] A. Nicolson, J. Breternitz, S. R. Kavanagh, Y. Tomm, K. Morita, A. G. Squires, M. Tovar, A. Walsh, S. Schorr, and D. O. Scanlon, Interplay of static and dynamic disorder in the mixed-metal chalcohalide $Sn_2SbS_2I_3$, J. Am. Chem. Soc. **145**, 12509 (2023).

[23] A. N. Roth, A. P. Porter, S. Horger, K. Ochoa-Romero, G. Guirado, A. J. Rossini, and J. Vela, Lead-free semiconductors: Phase-evolution and superior stability of multinary tin chalcohalides, Chem. Mater. **36**, 4542 (2024).

[24] A. Ibanez, J.-C. Jumas, J. Olivier-Fourcade, and E. Philippot, Mise en évidence d'un désordre statistique dans les structures chalcogénoiodures d'étain et d'antimoine, J. Solid State Chem. **55**, 83 (1984).

[25] S. M. Islam, C. D. Malliakas, D. Sarma, D. C. Maloney, C. C. Stoumpos, O. Y. Kontsevoi, A. J. Freeman, and M. G. Kanatzidis, Direct gap semiconductors $Pb_2BiS_2I_3$, $Sn_2BiS_2I_3$, and Sn_2BiSI_5, Chem. Mater. **28**, 7332 (2016).

[26] V. Dolgikh, New chalkogen-halogenides of the type $M_2SbS_2I_3$, Izv. Akad. Nauk SSSR, Neorg. Mater. **21**, 1211 (1985).

[27] C. Doussier, Y. Moëlo, P. Léone, A. Meerschaut, and M. Evain, Crystal structure of $Pb_2Sb_22I_3$, and re-examination of the crystal chemistry within the group of (Pb/Sn/Sb) chalcogenoiodides, Solid State Sci. **9**, 792 (2007).

[28] A. N. Roth, J. Opare-Addo, E. Gi, S. Mena, G. Guirado, R. D. Schaller, E. A. Smith, and J. Vela, Solution-phase synthesis and photoluminescence of quaternary chalcohalide semiconductors, Chem. Mater. **35**, 2165 (2023).

[29] See Supplemental Material at http://link.aps.org/supplemental/10.1103/qjwt-29w9 for (I) RF regression model performance per phase for the studied material properties (formation energies, fundamental band gaps, and effective electron and hole masses); (II) tabulated computed values for formation energies, fundamental band gaps, minimal direct band gaps, optical band gaps, as well as effective electron and hole masses for all MMCH materials in all three phases; (III) RF regression model performance and results for minimum direct band gaps and optical band gaps in all three phases; (IV) DOS of lead-free and lead-based $M(\mathrm{II})_2BiSe_2Br_3$ compounds in $Cmc2_1$ and $P2_1/c$ phases; (V) band structures of lead-free and lead-based $M(\mathrm{II})_2BiSe_2Br_3$ compounds in all three phases; and

(VI) comparison of MMCHs material properties in $Cmc2_1$ and $P2_1/c$ phases, which includes Ref. [88].

[30] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder *et al.*, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, APL Mater. **1**, 011002 (2013).

[31] K. Momma and F. Izumi, VESTA: A three-dimensional visualization system for electronic and structural analysis, J. Appl. Crystallogr. **41**, 653 (2008).

[32] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, *Ab initio* molecular simulations with numeric atom-centered orbitals, Comput. Phys. Commun. **180**, 2175 (2009).

[33] V. Havu, V. Blum, P. Havu, and M. Scheffler, Efficient integration for all-electron electronic structure calculation using numeric basis functions, J. Comput. Phys. **228**, 8367 (2009).

[34] X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, and M. Scheffler, Resolution-of-identity approach to Hartree-Fock, hybrid density functionals, RPA, MP2 and *GW* with numeric atom-centered orbital basis functions, New J. Phys. **14**, 053020 (2012).

[35] S. V. Levchenko, X. Ren, J. Wieferink, R. Johanni, P. Rinke, V. Blum, and M. Scheffler, Hybrid functionals for large periodic systems in an all-electron, numeric atom-centered basis framework, Comput. Phys. Commun. **192**, 60 (2015).

[36] V. W.-z. Yu, F. Corsetti, A. García, W. P. Huhn, M. Jacquelin, W. Jia, B. Lange, L. Lin, J. Lu, W. Mi *et al.*, ELSI: A unified software interface for Kohn-Sham electronic structure solvers, Comput. Phys. Commun. **222**, 267 (2018).

[37] A. C. Ihrig, J. Wieferink, I. Y. Zhang, M. Ropo, X. Ren, P. Rinke, M. Scheffler, and V. Blum, Accurate localized resolution of identity approach for linear-scaling hybrid density functionals and for many-body perturbation theory, New J. Phys. **17**, 093020 (2015).

[38] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Restoring the density-gradient expansion for exchange in solids and surfaces, Phys. Rev. Lett. **100**, 136406 (2008).

[39] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Erratum: Restoring the density-gradient expansion for exchange in solids and surfaces [Phys. Rev. Lett. **100**, 136406 (2008)], Phys. Rev. Lett. **102**, 039902(E) (2009).

[40] R. X. Yang, J. M. Skelton, E. L. da Silva, J. M. Frost, and A. Walsh, Spontaneous octahedral tilting in the cubic inorganic cesium halide perovskites $CsSnX_3$ and $CsPbX_3$ (X = F, Cl, Br, I), J. Phys. Chem. Lett. **8**, 4720 (2017).

[41] M. Bokdam, J. Lahnsteiner, B. Ramberger, T. Schäfer, and G. Kresse, Assessing density functionals using many body theory for hybrid perovskites, Phys. Rev. Lett. **119**, 145501 (2017).

[42] J. Laakso, M. Todorović, J. Li, G.-X. Zhang, and P. Rinke, Compositional engineering of perovskites with machine learning, Phys. Rev. Mater. **6**, 113801 (2022).

[43] F. Pan, J. Zhai, J. Chen, L. Yang, H. Dong, F. Yuan, Z. Jiang, W. Ren, Z.-G. Ye, G.-X. Zhang *et al.*, Mixed-halide perovskite alloys $CsPb(I_{1-x}Br_x)_3$ and $CsPb(Br_{1-x}Cl_x)_3$: New insight of configurational entropy effect from first-principles and phase diagrams, Chem. Mater. **36**, 3957 (2024).

[44] J. Li, F. Pan, G.-X. Zhang, Z. Liu, H. Dong, D. Wang, Z. Jiang, W. Ren, Z.-G. Ye, M. Todorović *et al.*, Structural disorder by octahedral tilting in inorganic halide perovskites: New insight with Bayesian optimization, Small Struct. **5**, 2400268 (2024).

[45] F. Knuth, C. Carbogno, V. Atalla, V. Blum, and M. Scheffler, All-electron formalism for total energy strain derivatives and stress tensor components for numeric atom-centered orbitals, Comput. Phys. Commun. **190**, 33 (2015).

[46] E. van Lenthe, E. J. Baerends, and J. G. Snijders, Relativistic regular two-component Hamiltonians, J. Chem. Phys. **99**, 4597 (1993).

[47] J. Heyd, G. E. Scuseria, and M. Ernzerhof, Hybrid functionals based on a screened Coulomb potential, J. Chem. Phys. **118**, 8207 (2003).

[48] J. Heyd, G. E. Scuseria, and M. Ernzerhof, Erratum: "Hybrid functionals based on a screened Coulomb potential" [J. Chem. Phys. **118**, 8207 (2003)], J. Chem. Phys. **124**, 219906 (2006).

[49] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, Influence of the exchange screening parameter on the performance of screened hybrid functionals, J. Chem. Phys. **125**, 224106 (2006).

[50] W. P. Huhn and V. Blum, One-hundred-three compound band-structure benchmark of post-self-consistent spin-orbit coupling treatments in density functional theory, Phys. Rev. Mater. **1**, 033803 (2017).

[51] L. D. Whalley, effmass: An effective mass package, J. Open Source Software **3**, 797 (2018).

[52] C. Ambrosch-Draxl and J. O. Sofo, Linear optical properties of solids within the full-potential linearized augmented planewave method, Comput. Phys. Commun. **175**, 1 (2006).

[53] C. G. Van de Walle and J. Neugebauer, Structure and energetics of nitride surfaces under MOCVD growth conditions, J. Cryst. Growth **248**, 8 (2003).

[54] L. Himanen, A. Geurts, A. S. Foster, and P. Rinke, Data-driven materials science: Status, challenges, and perspectives, Adv. Sci. **6**, 1900808 (2019).

[55] P. Henkel, J. Li, and P. Rinke, NOMAD dataset: Lead-free $Sn_2M(II)Ch_2X_3$ materials (2023), 10.17172/NOMAD/2023.08.24-1.

[56] P. Henkel, J. Li, and P. Rinke, NOMAD dataset: Lead-based $Pb_2M(II)Ch_2X_3$ materials (2025), 10.17172/NOMAD/2025.05.09-1.

[57] L. Breiman, Random forests, Mach. Learn. **45**, 5 (2001).

[58] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, Springer Texts in Statistics Vol. 103 (Springer, New York, 2013).

[59] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, SCIKIT-LEARN: Machine learning in Python, J. Machine Learning Res. **12**, 2825 (2011).

[60] Y. Huang, C. Yu, W. Chen, Y. Liu, C. Li, C. Niu, F. Wang, and Y. Jia, Band gap and band alignment prediction of nitride-based semiconductors using machine learning, J. Mater. Chem. C **7**, 3238 (2019).

[61] V. Gladkikh, D. Y. Kim, A. Hajibabaei, A. Jana, C. W. Myung, and K. S. Kim, Machine learning for predicting the band gaps of ABX_3 perovskites from elemental properties, J. Phys. Chem. C **124**, 8905 (2020).

[62] G. Pilania, A. Mannodi-Kanakkithodi, B. P. Uberuaga, R. Ramprasad, J. E. Gubernatis, and T. Lookman, Machine

learning bandgaps of double perovskites, Sci. Rep. **6**, 19375 (2016).

[63] T. A. R. Purcell, M. Scheffler, C. Carbogno, and L. M. Ghiringhelli, SISSO++: A C++ implementation of the sure-independence screening and sparsifying operator approach, J. Open Source Software **7**, 3960 (2022).

[64] S. M. Lundberg and S.-I. Lee, A unified approach to interpreting model predictions, in *Advances in Neural Information Processing Systems*, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc., Red Hook, 2017), Vol. 30.

[65] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee, From local explanations to global understanding with explainable AI for trees, Nat. Mach. Intell. **2**, 56 (2020).

[66] S. M. Lundberg, G. G. Erion, and S.-I. Lee, Consistent individualized feature attribution for tree ensembles, arXiv:1802.03888.

[67] S. M. Lundberg, SHAP, https://shap.readthedocs.io/en/latest/index.html.

[68] W. Kaiser, D. Ricciarelli, E. Mosconi, A. A. Alothman, F. Ambrosio, and F. De Angelis, Stability of tin- versus lead-halide perovskites: *Ab initio* molecular dynamics simulations of perovskite/water interfaces, J. Phys. Chem. Lett. **13**, 2321 (2022).

[69] J. Pascual, G. Nasti, M. H. Aldamasy, J. A. Smith, M. Flatken, N. Phung, D. Di Girolamo, S.-H. Turren-Cruz, M. Li, A. Dallmann *et al.*, Origin of Sn(II) oxidation in tin halide perovskites, Mater. Adv. **1**, 1066 (2020).

[70] A. F. Akbulatov, S. A. Tsarev, M. Elshobaki, S. Y. Luchkin, I. S. Zhidkov, E. Z. Kurmaev, S. M. Aldoshin, K. J. Stevenson, and P. A. Troshin, Comparative intrinsic thermal and photochemical stability of Sn(II) complex halides as next-generation materials for lead-free perovskite solar cells, J. Phys. Chem. C **123**, 26862 (2019).

[71] T. Leijtens, R. Prasanna, A. Gold-Parker, M. F. Toney, and M. D. McGehee, Mechanism of tin oxidation and stabilization by lead substitution in tin halide perovskites, ACS Energy Lett. **2**, 2159 (2017).

[72] Edited by J. R. Rumble, T. J. Brunno, and M. J. Doa, *CRC Handbook of Chemistry and Physics: A Ready-Reference Book of Chemical and Physical Data*, 104th ed. (CRC Press, Boca Raton, 2023).

[73] H. Shi, D. Han, S. Chen, and M.-H. Du, Impact of metal ns^2 lone pair on luminescence quantum efficiency in low-dimensional halide perovskites, Phys. Rev. Mater. **3**, 034604 (2019).

[74] J. M. Frost and A. Walsh, What is moving in hybrid halide perovskite solar cells? Acc. Chem. Res. **49**, 528 (2016).

[75] F. Brivio, A. B. Walker, and A. Walsh, Structural and electronic properties of hybrid perovskites for high-efficiency thin-film photovoltaics from first-principles, APL Mater. **1**, 042111 (2013).

[76] A. Walsh, D. J. Payne, R. G. Egdell, and G. W. Watson, Stereochemistry of post-transition metal oxides: Revision of the classical lone pair model, Chem. Soc. Rev. **40**, 4455 (2011).

[77] F. Hao, C. C. Stoumpos, R. P. H. Chang, and M. G. Kanatzidis, Anomalous band gap behavior in mixed Sn and Pb perovskites enables broadening of absorption spectrum in solar cells, J. Am. Chem. Soc. **136**, 8094 (2014).

[78] Z. Yang, A. Rajagopal, C.-C. Chueh, S. B. Jo, B. Liu, T. Zhao, and A. K.-Y. Jen, Stable low-bandgap Pb-Sn binary perovskites for tandem solar cells, Adv. Mater. **28**, 8990 (2016).

[79] J. Li, Q. Wang, and A. Abate, Perovskite solar cells, *Nanomaterials for Solar Cell Applications* (Elsevier, Amsterdam, 2019), pp. 417–446.

[80] R. Kour, S. Arya, S. Verma, J. Gupta, P. Bandhoria, V. Bharti, R. Datt, and V. Gupta, Potential substitutes for replacement of lead in perovskite solar cells: A review, Global Challenges **3**, 1900050 (2019).

[81] Z. Yang, A. Rajagopal, S. B. Jo, C.-C. Chueh, S. Williams, C.-C. Huang, J. K. Katahara, H. W. Hillhouse, and A. K.-Y. Jen, Stabilized wide bandgap perovskite solar cells by tin substitution, Nano Lett. **16**, 7739 (2016).

[82] A. M. Huerta-Flores, N. A. García-Gómez, S. M. De La Parra, and E. M. Sánchez, Comparative study of Sb_2S_3, Bi_2S_3 and In_2S_3 thin film deposition on TiO_2 by successive ionic layer adsorption and reaction (SILAR) method, Mater. Sci. Semicond. Process. **37**, 235 (2015).

[83] A. Matsumoto, Y. Koyama, A. Togo, M. Choi, and I. Tanaka, Electronic structures of dynamically stable As_2O_3, Sb_2O_3, and Bi_2O_3 crystal polymorphs, Phys. Rev. B **83**, 214110 (2011).

[84] J. Feng and B. Xiao, Effective masses and electronic and optical properties of nontoxic $MASnX_3$ (X = Cl, Br, and I) perovskite structures as solar cell absorber: A theoretical study using HSE06, J. Phys. Chem. C **118**, 19655 (2014).

[85] Y. Chen, L. Xiao, L. Shi, and P. Qian, High-throughput screening of the transport behavior of tetragonal perovskites, Phys. Chem. Chem. Phys. **26**, 9378 (2024).

[86] N. Ashari-Astani, S. Meloni, A. H. Salavati, G. Palermo, M. Grätzel, and U. Rothlisberger, Computational characterization of the dependence of halide perovskite effective masses on chemical composition and structure, J. Phys. Chem. C **121**, 23886 (2017).

[87] M. Dyksik, H. Duim, X. Zhu, Z. Yang, M. Gen, Y. Kohama, S. Adjokatse, D. K. Maude, M. A. Loi, D. A. Egger *et al.*, Broad tunability of carrier effective masses in two-dimensional halide perovskites, ACS Energy Lett. **5**, 3609 (2020).

[88] V. I. Starosta, J. Kroutil, and L. Beneš, Preparation and fundamental physical properties of $Sn_2SbS_2I_3$ and $Pb_2SbS_2I_3$ compounds, Cryst. Res. Technol. **25**, 1439 (1990).