

Machine Learning-Aided Band Gap Engineering of BaZrS₃ Chalcogenide Perovskite

Shyam Sharma, Zachary D. Ward, Kevin Bhimani, Mukul Sharma, Joshua Quinton, Trevor David Rhone, Su-Fei Shi, Humberto Terrones,* and Nikhil Koratkar*



Cite This: <https://doi.org/10.1021/acsami.3c00618>



Read Online

ACCESS |



Metrics & More



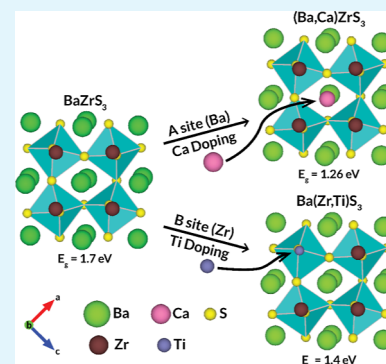
Article Recommendations



Supporting Information

ABSTRACT: The non-toxic and stable chalcogenide perovskite BaZrS₃ fulfills many key optoelectronic properties for a high-efficiency photovoltaic material. It has been shown to possess a direct band gap with a large absorption coefficient and good carrier mobility values. With a reported band gap of 1.7–1.8 eV, BaZrS₃ is a good candidate for tandem solar cell materials; however, its band gap is significantly larger than the optimal value for a high-efficiency single-junction solar cell (~1.3 eV, Shockley–Queisser limit)—thus doping is required to lower the band gap. By combining first-principles calculations and machine learning algorithms, we are able to identify and predict the best dopants for the BaZrS₃ perovskites for potential future photovoltaic devices with a band gap within the Shockley–Queisser limit. It is found that the Ca dopant at the Ba site or Ti dopant at the Zr site is the best candidate dopant. Based on this information, we report for the first time partial doping at the Ba site in BaZrS₃ with Ca (i.e., Ba_{1-x}Ca_xZrS₃) and compare its photoluminescence with Ti-doped perovskites [i.e., Ba(Zr_{1-x}Ti_x)S₃]. Synthesized (Ba,Ca)ZrS₃ perovskites show a reduction in the band gap from ~1.75 to ~1.26 eV with <2 atom % Ca doping. Our results indicate that for the purpose of band gap tuning for photovoltaic applications, Ca-doping at the Ba-site is superior to Ti-doping at the Zr-site reported previously.

KEYWORDS: BaZrS₃ films, band gap engineering, machine learning, random forest regression, density functional theory



INTRODUCTION

Society is increasingly moving toward renewable energy sources to reduce carbon emissions. To achieve this, one of the most attractive approaches is to harvest solar energy. According to the International Energy Agency (IEA), solar energy is on track to set a record for new global deployments after 2022, with an average of ~125 GW of new capacity expected globally between 2021 and 2025.¹ Moreover, with a plethora of policies being proposed to achieve net-zero emissions over the next few decades, solar panel usage is expected to continue to grow worldwide.²

Despite these advances, solar technology has some inherent limitations. Even though solar panel costs are decreasing due to mass production, significant efficiency improvements have yet to take place.³ For decades, silicon (Si) has been the go-to material for solar cells (~90% market share⁴). However, Si has challenges with respect to both cost and efficiency.⁵ For example, the efficiency of Si solar cells has plateaued at ~21–22% (for polycrystalline Si). Also, the manufacturing of Si solar cells is expensive and not environmentally friendly due to the need to remove impurities from Si.⁶ Today, mainstream Si solar technology is reaching its practical efficiency limit when used alone.^{7,8} However, in recent years, perovskite-based materials have emerged as an alternative to Si, which could potentially revolutionize photovoltaic (PV) technology.⁹ Most common perovskite PV cells use a hybrid organic–inorganic

lead or tin halide-based perovskite as the active layer such as methylammonium lead iodide (MAPbI₃) which is cheap and simple to manufacture. Perovskites offer a broad absorption spectrum, fast charge separation, long transport distance of electrons and holes, and long carrier separation lifetimes, which make them highly attractive in PV applications.¹⁰ In fact, halide perovskite-based solar cells have achieved an unprecedented improvement in performance, with a boom in efficiency from about 3% in 2006 to over 29% in 2020.¹¹

A major advantage that perovskite PVs have over conventional solar technology is that they can react to the various wavelengths of light, enabling them to convert more of the available solar energy into electricity. This band gap flexibility opens up other valuable applications for these solar cells in high-performance tandem device configurations that can achieve efficiency above 30%.¹² Moreover, they offer flexibility, semi-transparency, and tailored form factors, and are lightweight.¹³ For these reasons, the perovskite thin-film structure

Received: January 13, 2023

Accepted: March 20, 2023

uses about 20 times less material than Si cells and does not require materials that are limited in supply.¹⁴

In May 2021, Saule launched the world's first industrial production line of halide perovskite solar cells.¹⁵ Jinko Solar, another leader in the PV market, is also working on producing perovskite technology.¹⁶ Even though many companies are investing in this technology, halide perovskites, the current state-of-the-art, have their own set of unique challenges that need to be addressed. The main hurdle for halide perovskite devices is whether they can last as long as a typical Si panel (25 years).¹⁷ Halide perovskites are very sensitive to oxygen, moisture, and heat, which require heavy encapsulation to protect the cells, increasing both the cost and weight.¹⁸ For example, Microquanta perovskite solar cells showed ~80% reduction in efficiency after 1–2 years of operation.⁵ Another concern with lead-based halide perovskites is their toxicity due to the presence of lead.¹⁸

To address the air sensitivity and toxicity issues faced by halide perovskites, interest is growing in chalcogenide perovskites (ABX₃). These perovskites are distinguished by their excellent thermal and chemical stability, non-toxicity, and abundance of elemental components.⁹ With a band gap of ~1.7 eV and being one of the few stable chalcogenide perovskites with a GdFeO₃-type structure, BaZrS₃ is the most popular chalcogenide perovskite synthesized by several independent research teams.^{18–20} This material shows strong absorption in the visible spectrum and is stable against oxidation and moisture. However, the 1.7 eV band gap of BaZrS₃ is larger than the ideal band gap required to achieve the maximum theoretical solar cell efficiency of ~33.7% (Shockley–Queisser limit⁸) for a single-junction PV cell. To achieve optimal efficiency, the band gap of BaZrS₃ must be reduced to ~1.34 eV. Tuning the band gap of BaZrS₃ can be achieved by cation or anion doping.¹² For anion doping, Nishigaki et al. reported the synthesis of BaZr(S,Se)₃.²¹ However, due to the non-perovskite ground state of BaZrSe₃ and the existence of a competing phase, the range of doping was limited. Additionally, partial replacement of the B cation (Zr) with Ti has shown a reduction in the band gap to ~1.4 eV.^{22,23} The range of doping was again limited by the non-perovskite ground state of BaTiS₃. In the following, with the term doping, we will refer to isovalent alloying or atom lattice substitution of the same valency.

It is challenging to predict which chalcogenide perovskite (ABX₃) and dopant combination would lead to a decrease in the band gap, thereby maximizing the PV device efficiency. Considering 100 different perovskite candidate ABX₃ structures with 35 dopants at the A site and another 35 at the B site, 122,500 possible combinations need to be investigated. Since there are many such candidates to investigate, the use of serial experiments or first-principles calculations to search through the space of candidate materials would be both prohibitively time-consuming and expensive. Recently, machine learning (ML) has emerged as an important tool to accelerate the search of ideal candidates for applications in materials discovery.^{24–26} In this study, we combine ML techniques with first-principles calculations [density functional theory (DFT) and Heyd–Scuseria–Ernzerhof (HSE)] to predict the best candidates for the doping of chalcogenide perovskites. ML has been applied to other perovskites such as double perovskites, inorganic perovskites, and hybrid organic–inorganic perovskites^{27–31} to predict useful quantities such as the band gap and formation energies. Most studies using ML

to study perovskites have considered a variety of perovskite structures by using different elements (varying A and B in an ABO₃ perovskite, for example). Just a few studies report using dopants in specific perovskites.³² In this account, ML has been used to consider 35 dopants of the BaZrS₃ perovskite with six different concentrations to predict band gaps and thermodynamic stability (formation energies). Using ML, we were able to narrow down our search to just two dopants: Ti in site B and Ca in site A. To the best of our knowledge, doping of the A-site cation (Ba) in BaZrS₃ has not been experimentally demonstrated to date. Herein, we present a study of partial doping by Ca at A-site cation (Ba) in BaZrS₃ thin films to form (Ba,Ca)ZrS₃. We report the photoluminescence (PL) properties of such A-site cation-doped ((Ba,Ca)ZrS₃) films and compare them to previous reports on B cation-doped (Ba(Zr,Ti)S₃) materials. Our results indicate that Ca-doping at the A-site is a superior strategy to Ti-doping at the B-site for band gap engineering of BaZrS₃ thin films. This study also shows how ML-based approaches can be used to efficiently screen the parameter space in order to guide the optimal doping and band gap tuning of chalcogenide perovskites.

RESULTS AND DISCUSSION

In order to understand and predict the impact of introducing different dopants in the BaZrS₃ perovskite, DFT calculations were performed using the generalized gradient approximation Perdew–Burke–Ernzerhof (GGA-PBE)³³ and HSE³⁴ functionals to build a database of 35 different dopants at the A and B sites with three different doping concentrations (8.33, 12.5, and 25%) to deduce the best candidate dopants based on whether the band gap falls into the Shockley–Queisser limit (1–1.5 eV) and if the doped BaZrS₃ perovskite is thermodynamically stable (see Computational Methods). This database of 35 dopants is used to predict the band gaps and structural formation energies using ML models (see ML Methods).

Supporting Information Figure S1 shows the defect formation energy for various concentrations for both B site Ti-doping and A site Ca-doping of BaZrS₃ through GGA-PBE functional calculations. The formation energy for the doped BaZrS₃ perovskites is calculated using the relation

$$E_{\text{form}} = E_{\text{Doped-BaZrS}_3} - E_{\text{BaZrS}_3} + E_{\text{Ba,Zr}} - E_{\text{A,B}} \quad (1)$$

where E_{BaZrS_3} is the pristine BaZrS₃ energy, $E_{\text{Doped-BaZrS}_3}$ is the total energy for the A or B site-doped BaZrS₃, $E_{\text{A,B}}$ is the elemental energy for the A or B site dopant, and $E_{\text{Ba,Zr}}$ is the elemental energy for the Ba or Zr atom. Our calculations indicate that Ca-doped BaZrS₃ is more stable than Ti-doped BaZrS₃ and becomes increasingly so with decreasing doping concentrations (Figure S1). This is because the chemical and electronic properties of a Ca atom are better suited to the distorted perovskite structure than those of a Ti atom. This follows from the fact that BaZrS₃ has an orthorhombic crystal structure (a distorted perovskite structure); similarly, CaZrS₃ also has a distorted perovskite structure. In contrast, BaTiS₃ forms a hexagonal structure (a distorted wurtzite structure). However, as the Ca atom concentration increases, the defect formation energy increases faster as compared to Ti-doped BaZrS₃, likely due to its larger size which induces more stress in the system than Ti doping. For both the A and B site dopants, the formation energy (E_{form}) values are all larger than

the pristine BaZrS₃ formation energy, $E_{\text{form}} = -1.866$ eV/atom and the band gaps range from 0.5 to 2.5 eV.

For the ML models, the weighted average of each of the chemical descriptors (η) across all the elements contained in each of the BaZrS₃ supercells is used, which is referred to as the weighted chemical descriptor (see Materials Descriptors). For the band gap ML model, the chemical descriptors for the individual dopant elements are also included in the training set along with the weighted chemical descriptors with a cube root transformation also being applied to the band gaps in the model during predictions. For formation energies, both the random forest and CGCNN models accurately predict the formation energies (Figures 1 and 2). For band gap

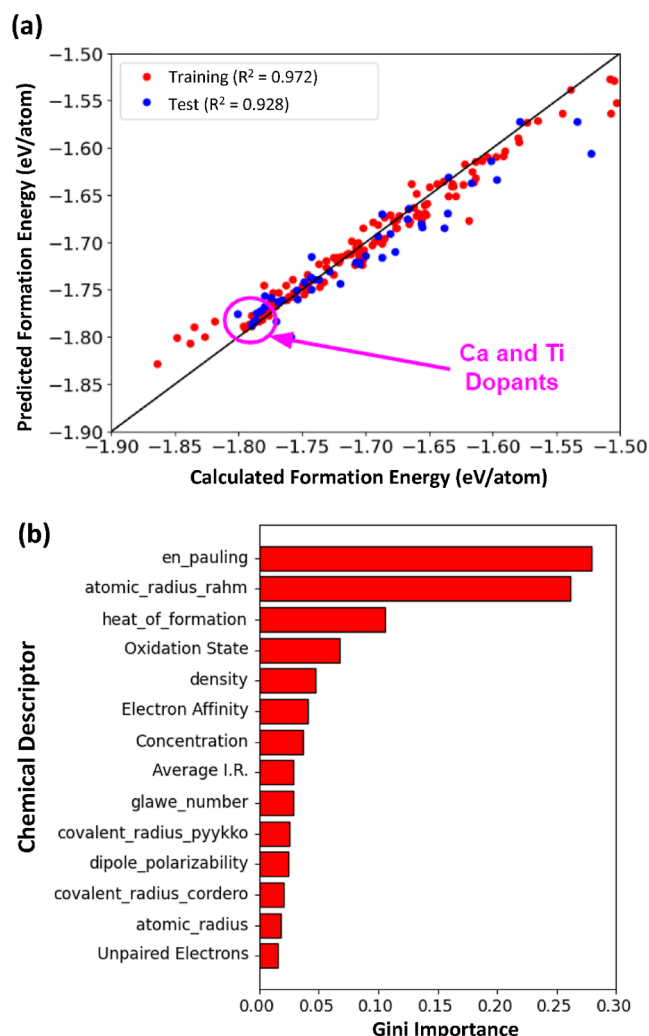


Figure 1. (a) Formation energy predictions for the A- and B-site dopants using the random forest algorithm with the training (red) and test (blue) sets shown with the black line representing the identity line. (b) Gini importance for the corresponding weighted chemical descriptors used for formation energy predictions.

predictions, the random forest algorithm works well to yield accurate predictions (Figure 3). Based on the random forest algorithm, the mean absolute error (MAE) for the band gap predictions is MAE = 0.14 eV, and the MAE for the formation energy predictions is MAE = 0.02 eV/atom. Note that descriptor importance can be extracted using the random forest algorithm. For the formation energy predictions, the

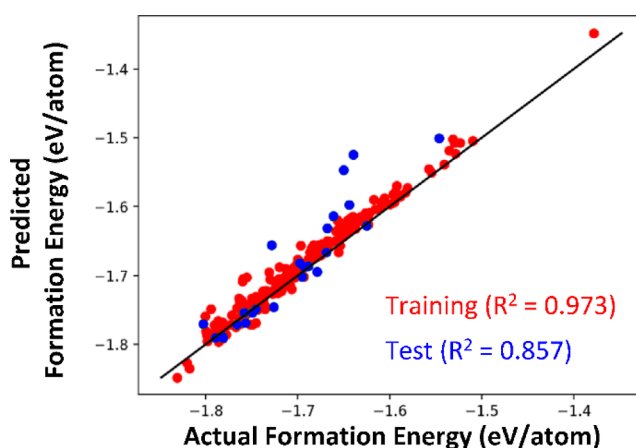


Figure 2. Formation energy predictions for the A- and B-site dopants using the deep learning algorithm with the training (red) and test (blue) sets shown, with the black line representing the identity line.

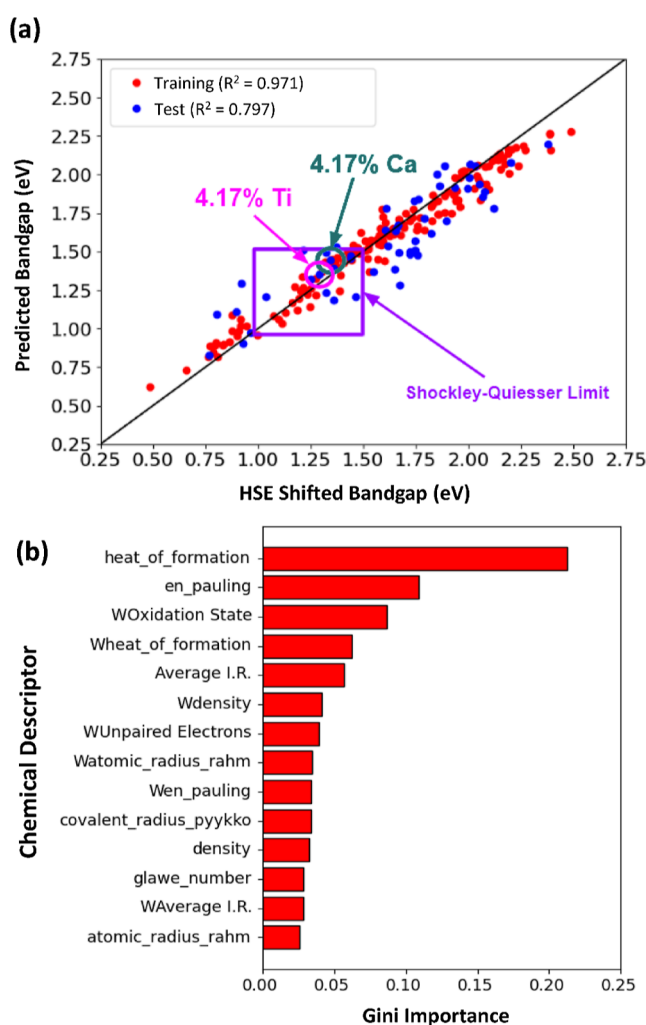


Figure 3. (a) Band gap predictions for the A- and B-site dopants using the random forest algorithm with the training (red) and test (blue) sets shown, with the black line representing the identity line. (b) Gini importance for the corresponding chemical descriptors used for band gap predictions. Note: W in Watomic_radius_rahm represents a weighted chemical descriptor, while the chemical descriptors without the W at the beginning are the chemical descriptors for the individual single dopant.

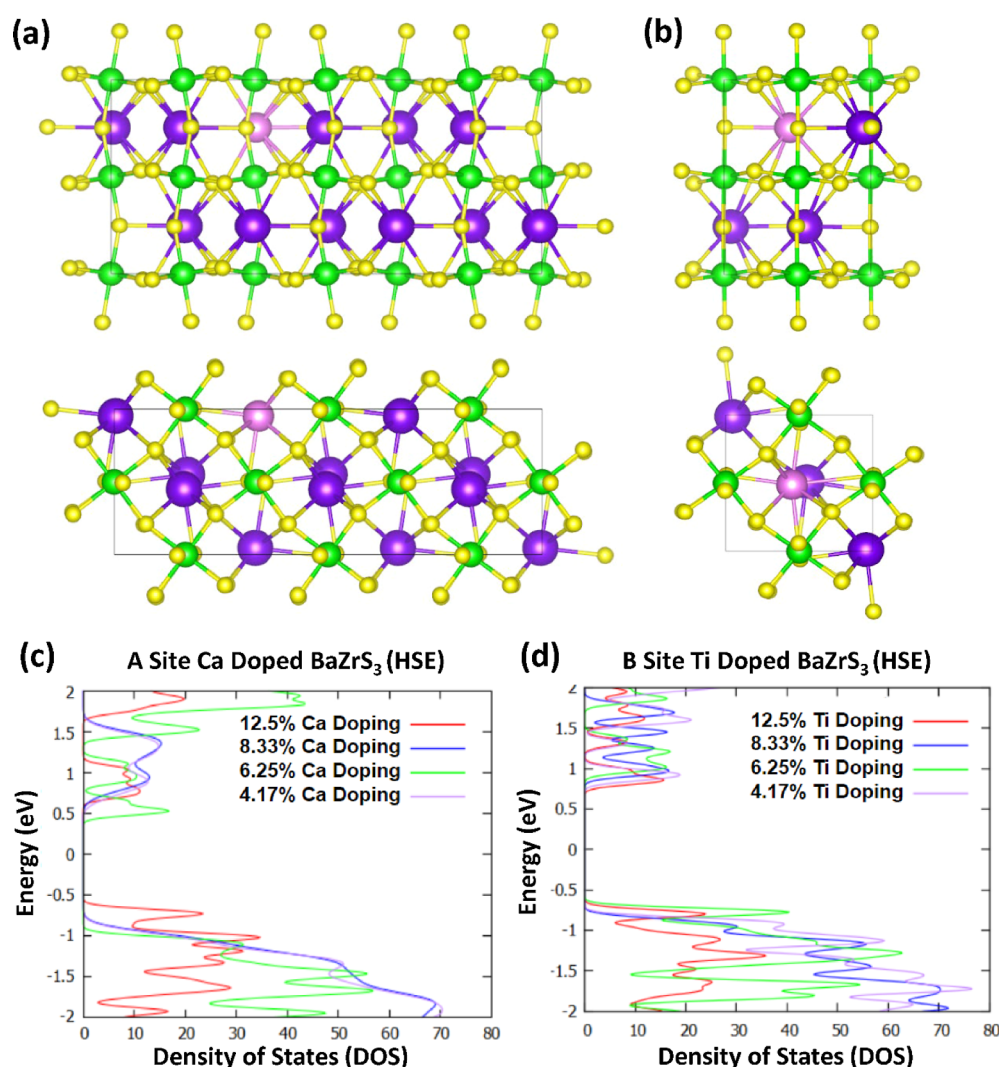


Figure 4. DFT calculations of Ca-doped BaZrS₃. (a,b) Ca-doped BaZrS₃ perovskite supercells with (a) 8.33 and (b) 25 atom % doping, where Ca, Ba, Zr, and S are represented by pink, purple, green, and yellow atoms, respectively. (c,d) DOS using the HSE functional for varying (c) Ca and (d) Ti doping concentrations.

chemical descriptors with the highest Gini importance are the Pauling electro-negativity³⁵ and Rahm atomic radius^{36,37} (Figure 1), while for band gap predictions, the heat of formation³⁵ exhibits the highest Gini importance (Figure 3) (see ML Methods for chemical descriptors used).

Given that ML models can accurately predict the band gaps and formation energies, a filtering process is applied to determine the dopants with the best potential to create stable and efficient PV devices. First, the doped BaZrS₃ perovskite must yield a band gap within the Shockley–Queisser limit (1–1.5 eV),³⁸ and second, the perovskite must be structurally stable. Among all the 35 dopants investigated, only Ti at the B site and Ca at the A site are predicted to have a band gap within the Shockley–Queisser limit, while La and Ca at the A site and Y and Ti at the B site are expected to be energetically stable (Supporting Information Figure S2). The only elements that meet the conditions for the stability and band gap criteria are Ca at the A site and Ti at the B site, each confirmed to be stable and exhibiting a band gap in the Shockley–Queisser limit at five distinct concentrations (8.3, 12.5, and 25% with 4.17 and 6.25% doping included specifically for the A site Ca and B site Ti) with the formation energies being accurately

predicted and the band gap residing in the Shockley–Queisser limit range (Figure 1). Figure 4a,b shows the Ca-doped BaZrS₃ perovskite supercells with 8.33 and 25 Ca atom %, respectively. The La and Y dopants with 25% doping are the most energetically stable dopants in this investigation, which agrees with the finding that La and Y elements are found to yield stable LaYS₃ perovskite structures.³⁹ Provided stable LaYS₃ perovskites can be synthesized; they can potentially operate as a platform for other dopants to tune the band gap into the Shockley–Queisser limit. Double dopants were also investigated, but there were no double dopant combinations found that could potentially create a stable PV device while also yielding a band gap in the Shockley–Queisser limit.

B site Ti- and A site Ca-doped BaZrS₃ are found to be the most promising from the DFT/ML calculations. It should be noted that the DFT/ML-predicted band gaps are verified using experiments, the details of which are presented in the later section. The density of states (DOS) for both Ca and Ti dopants was calculated (Figure 4c,d) using HSE which corrects the band gaps. To best represent the band gaps obtained from the experimental results, the screening parameter (μ) is tuned within the HSE correction.^{34,40–42} The values obtained for μ

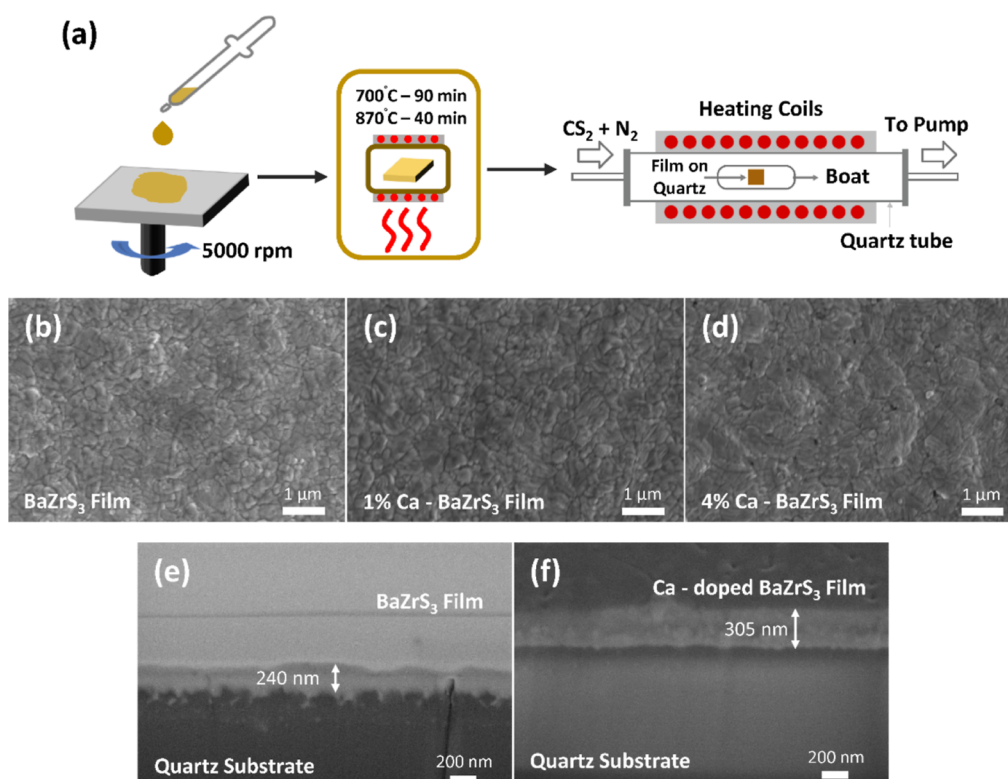


Figure 5. Synthesis and characterization of undoped and Ca-doped BaZrS₃ thin films. (a) Schematic illustrating the synthesis of BaZrS₃-based perovskite films. (b–d) SEM images of the (b) undoped BaZrS₃ film and (c,d) Ca-doped BaZrS₃ film. (e,f) FIB cross-sectional SEM images.

which best match the experimental band gaps are $\mu = 0.20$ and $\mu = 0.08$ for Ca and Ti, respectively. The HSE-calculated band gaps agree well with the experimental band gaps for low (<8.33 atom %) doping concentrations using this method. The B-site Ti-doped BaZrS₃ system yields a band gap at the upper end of the Shockley–Queisser limit (1.4–1.5 eV), while the A-site Ca-doped BaZrS₃ system yields a band gap in the middle of the Shockley–Queisser limit (1.23–1.28 eV). The average Ba–S and Zr–S bond lengths formed in the $3 \times 2 \times 1$ supercell are 3.29 and 2.57 Å, respectively. For the 4.17% Ca-doped perovskite, the average Ca–S bond length is 3.14 Å, while for the 4.17% Ti doped perovskite, the average Ti–S bond length is 2.49 Å. The acute S–Ca–S bond angles in the 4.17% Ca-doped perovskite tend to deviate by about 2–3° from the corresponding S–Ba–S bond angles in pristine BaZrS₃, while in the 4.17% Ti-doped perovskite, the S–Ti–S bond angles remain mostly unchanged, each measuring about 90°. The larger reduction in the band gap in (Ba,Ca)ZrS₃ when compared to Ba(Zr,Ti)S₃ may be due to more significant changes in the bonds and greater induced stress on the BaZrS₃ perovskite. The lowest concentration studied by first-principles calculations is 4.17% doping; lower concentrations would be too computationally expensive for HSE calculations. The band structure is also calculated on the X- Γ -Z-Y- Γ -U-S- Γ -R k -point path for the 4.17% Ca-doped BaZrS₃ with the PBE functional and reveals that the band gap is direct for 4.17% Ca doped BaZrS₃ at the Γ point (Supporting Information Figure S1a). The lower-concentration (4.17 and 6.25%) A site Ca and B site Ti-doped BaZrS₃ perovskites are also included in the ML dataset.

DFT/ML suggests that the two best candidate dopants for PV devices in BaZrS₃ perovskites are Ti at the B sites or Ca at the A sites. In a recent work, we successfully synthesized B site

Ti-doped BaZrS₃ thin films.²² In the present study, we have synthesized A site Ca-doped BaZrS₃ thin films using chemical vapor deposition (CVD). The methodology used for CVD synthesis of BaZrS₃ and Ca-doped BaZrS₃ thin films by the sulfurization of oxide precursor films is provided in the Methods section and has been illustrated in Figure 5a. The morphology of the as-synthesized films was studied using scanning electron microscopy (SEM), the images of which are shown in Figure 5b–d. These results reveal the films to be polycrystalline and continuous (with some cracks which can be attributed to the high-temperature synthesis). The cross-sectional SEM image of Ca-doped BaZrS₃ shown in Figure 5e,f confirms that the film's thickness lies in the range of 250–300 nm, which is in the ideal range for a single-junction PV device.

The X-ray diffraction (XRD) pattern of Ca-doped BaZrS₃ thin films with different Ca doping (0 to 4%) is shown in Figure 6a. For 0% doping, the peaks match with the standard reference pattern of BaZrS₃ (ICDD 15-0327), confirming that the sample is polycrystalline and possesses an orthorhombic distorted perovskite structure with a $Pnma$ space group. No secondary phase is present in all the samples. The background signal pattern emerges from the quartz substrate, as shown in Supporting Information Figure S3. The XRD pattern for the substrate and precursor Ca-doped BaZrO₃ films is provided in the Supporting Information Figure S4. An enlarged view of the four highest intensity peaks of BaZrS₃ is shown in Figure 6b–e. A peak shift is observed at lower Ca-doping concentration. Besides the peak shift, an increase in the full width at half maximum (fwhm) of the BaZrS₃ peaks with an increase in Ca concentration is also evident (calculations are shown in the Supporting Information Figure S5). This indicates reduced stability of BaZrS₃ at higher doping concentrations of Ca,

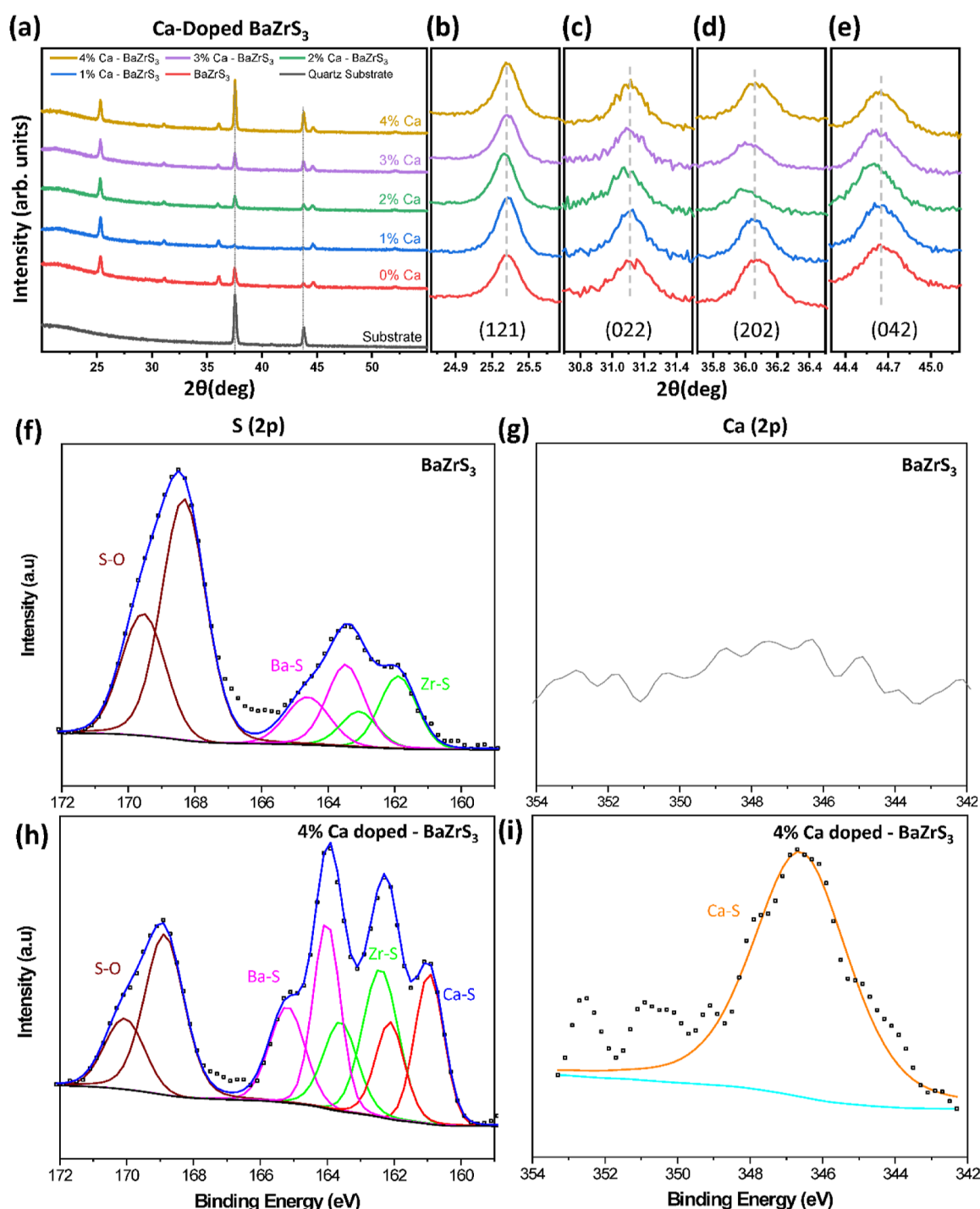


Figure 6. (a–e) XRD patterns of undoped and doped BaZrS₃ films on a quartz substrate at different Ca concentrations in (a) 20–55, (b) 25.0–25.8, (c) 30.6–31.3, (d) 35.6–36.6, and (e) 44.3–45.2 (2θ degrees). (f–i) High-resolution XPS spectra of S 2p and Ca 2p.

which is also reflected in the defect formation energy calculations shown in Supporting Information Figure S1b.

X-ray photoelectron spectroscopy (XPS) was employed to investigate chemical bonds in Ca-doped BaZrS₃ and BaZrS₃ films (Figure 6f–i). A C–C bond, observed due to adsorbed carbon species on films at 284.6 eV binding energy, was used as a reference. High-resolution S 2p XPS spectra of BaZrS₃ show the appearance of three doublet peaks, which can be attributed to O–S, Ba–S, and Zr–S bonds (Figure 6f). In the case of Ca-doped BaZrS₃, four doublet peaks were observed, which can be attributed to O–S, Ba–S, Zr–S, and Ca–S

bonds (Figure 6h). High-resolution Ca 2p XPS spectra of Ca-doped BaZrS₃ show the appearance of a Ca–S peak (Figure 6i) and only noise in the case of BaZrS₃ (Figure 6g). An XPS survey and high-resolution XPS of Ba and Zr are provided in the Supporting Information Figures S6–S8, respectively. The XRD and XPS results in Figure 6 confirm the successful doping of Ca in the BaZrS₃ structure.

PL spectroscopy results of Ca-doped BaZrS₃ thin films on the quartz substrate in Figure 7a show a reduction in band gap due to Ca-doping. This is reflected in the PL peak shift from ~725 nm for BaZrS₃ to around ~990 nm for Ca-doped

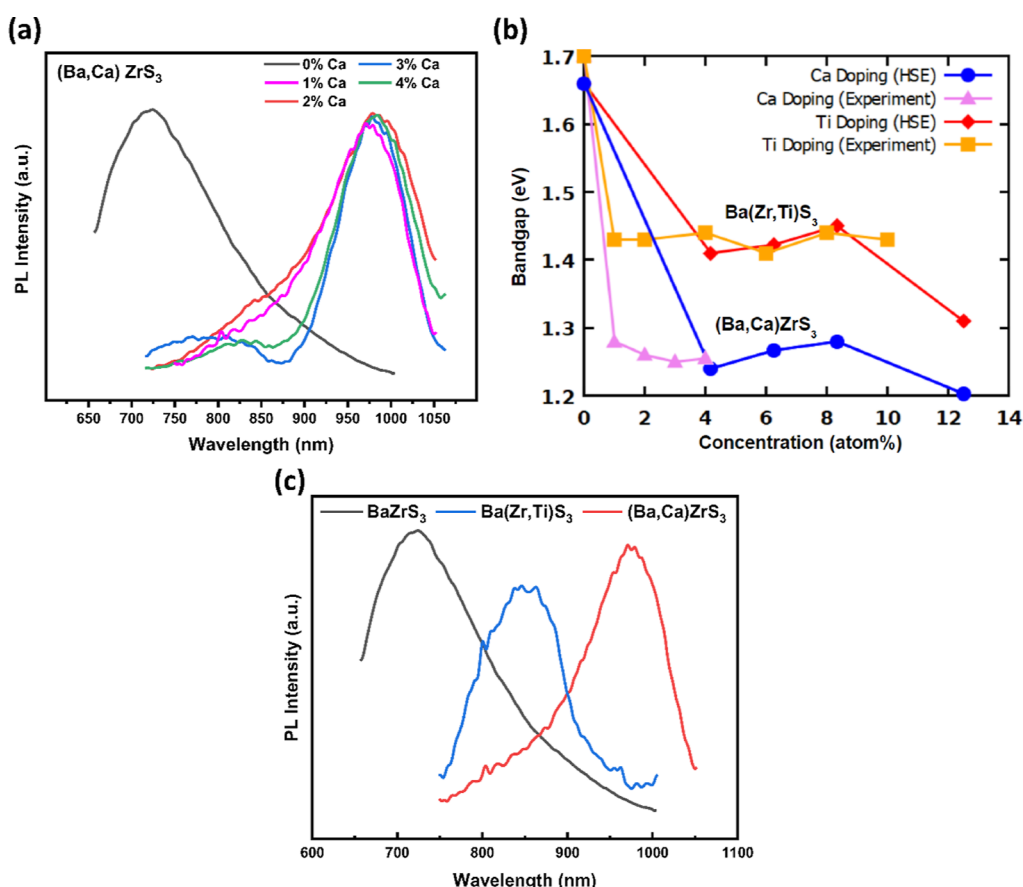


Figure 7. (a) PL of (Ba,Ca)ZrS₃ films on the quartz substrate at different Ca doping concentrations. (b) Experimental and calculated (using the HSE functional) band gap variation with Ca and Ti doping concentrations. (c) PL comparison for pristine BaZrS₃, (Ba,Ca)ZrS₃ with 1% Ca, and Ba(Zr,Ti)S₃ with 4% Ti doping concentration.

BaZrS₃, corresponding to a significant band gap reduction of ~490 meV. Apart from the redshift, some PL intensity drop is observed, which indicates the formation of recombination centers due to an increase in defect concentration resulting from the strain generated by Ca-doping. Finally, looking at the fwhm for all the PL peaks, they become narrower at higher Ca-doping concentrations, corresponding to a better open-circuit voltage potential.

Band gap variation with different doping concentrations and PL peaks of Ca- and Ti-doped and non-doped BaZrS₃ is shown in Figure 7b,c. Figure 7b also shows the HSE-calculated band gaps as a function of the doping concentration for both the Ca and Ti dopants along with the experimental band gaps. Three major observations can be made from these figures: (1) it is seen that Ca-doping results in a significantly higher band gap reduction than Ti-doping. We attribute this to the larger difference in the relative sizes of Ba/Ca as compared to Zr/Ti, which induces more stress on the perovskite structure; to release this induced stress, the ZrS₆ octahedra rotate, resulting in the band gap change. The greater stress induced by Ba/Ca doping compared to Zr/Ti will therefore induce a larger band gap change. (2) PL intensity drop is much higher in the case of Ti doping than in the case of Ca doping, which can be attributed to the structures formed by the doping elements. BaTiS₃ is not a GdFeO₃-type perovskite, but both BaZrS₃ and CaZrS₃ form GdFeO₃-type perovskite structures. Consequently, in the case of Ti-doping, impure secondary phases with different structures and narrower band gaps are formed,

which act as recombination centers. By contrast for Ca, due to the same GdFeO₃-type structure of the secondary phase and a high band gap of CaZrS₃, we see relatively less quenching of the PL intensity. (3) The fwhm for Ti-doped BaZrS₃ (124.8 nm) is relatively higher than that for the Ca-doped BaZrS₃ (112.5 nm) (calculations shown in Supporting Information Figure S9), resulting in a larger penalty in the open-circuit voltage potential.

SUMMARY

In conclusion, this study finds unique A site and B site substitutions for BaZrS₃ that tune the band gap into the Shockley–Queisser limit while retaining structural stability. The BaZrS₃ dopants are Ca at the A site and Ti at the B site. The ML algorithms presented in this study utilize chemical descriptors, comprising the properties of elements on the periodic table, to generate accurate predictions and discern promising candidate dopants from undesired dopants. The most promising candidate dopants are deduced to be Ti at the B site and Ca at the A site given their stability and band gaps residing in the Shockley–Queisser limit. ML/DFT predictions were verified using experiments.

The methodology presented in this study has the potential to work in other perovskites and the usage of weighted chemical descriptors has the potential to provide better predictions and more distinguishability when studying systems with varying stoichiometries. It may be the case that the dopants identified in this study are the only dopants which can

yield both stable doped perovskite structures in BaZrS₃ and a band gap in the Shockley–Queisser limit; however, there are a myriad of combinations of dopants and perovskite systems different from BaZrS₃ left to explore that may have potential for PV devices. ML could significantly accelerate the search for novel-doped perovskites for PV applications.

To demonstrate the technical feasibility of Ca doping at the A-site in BaZrS₃, we have successfully synthesized 250–300 nm thick Ca-doped BaZrS₃ thin films through a facile CVD technique based on the sulfurization of oxide films. We find that just 1 atom % Ca concentration is sufficient to red-shift the PL peak to a 980–990 nm wavelength, which corresponds to a considerable band gap reduction of ~490 meV to ~1.26 eV from ~1.75 eV for the non-doped BaZrS₃. The band gap of the Ca-doped BaZrS₃ material lies close to the optimal band gap (~1.34 eV) required to achieve the maximum theoretical efficiency (Shockley–Queisser limit) for a single-junction PV cell.

In contrast, for Ti doping at the B (Zr) site, we find that despite a relatively higher Ti doping concentration of 4 atom %, a band gap reduction of only ~350 meV to ~1.4 eV was possible. This shows larger band gap reduction at lower doping concentrations for Ca relative to Ti, which indicates greater potency for Ca doping. Furthermore, the comparative PL peak intensity drop was less in the case of Ca doping, suggesting less defect formation for Ca doping relative to Ti. The fwhm is also relatively lower for Ca compared to Ti doping, which indicates a reduced penalty in the open circuit potential. For these reasons, we conclude that Ca doping at the Ba site in BaZrS₃ is a superior doping strategy when compared to Ti doping at the Zr site for solar cell applications.

METHODS

Computational Methods. All first-principles calculations are performed using Quantum Espresso (QE)⁶⁵ to describe the electronic structure of doped BaZrS₃ perovskites. The GGA by the PBE functional⁶⁶ is used with the HSE correction⁶⁷ and applied to obtain the band gaps from the DOS. The relaxation calculations are performed on a 3 × 3 × 1 *k*-point mesh using the PBE functional with a kinetic energy cutoff of 1088 eV, and HSE calculations are performed at the Γ point. The atomic coordinates are relaxed until the total force and scf energy convergence are within the thresholds of 0.01 eV/Å and 10^{−6} eV, respectively. The unit cell for the BaZrS₃ perovskite structure contains 20 atoms with dimensions of *a* = 7.029 Å, *b* = 7.153 Å, and *c* = 10.077 Å, which is then expanded to generate supercells of size *m* × *n* × 1 with dimensions of *ma* Å × *nb* Å × *c* Å. Single Ca and Ti dopants are substituted at the Ba and Zr sites, respectively, to describe doping concentrations of 25, 12.5, 8.33, 6.25, and 4.17% with supercells of sizes 1 × 1 × 1, 2 × 1 × 1, 3 × 1 × 1, 2 × 2 × 1, and 3 × 2 × 1, respectively. The Ca site (A site) doping forms a Ba_{1−*x*}Y_{*x*}ZrS₃ perovskite for a dopant Y, while Ti site (B site) doping forms a BaZr_{1−*x*}Zr_{*x*}S₃ perovskite for a dopant Z with a doping concentration of (*x*). A similar procedure is performed with all dopants investigated in the ML algorithms, but with the default screening parameter of $\mu = 0.106$.

Database and Data Cleaning. The ML workflow and filtering process to identify stable doped BaZrS₃ perovskites with a band gap in the Shockley–Queisser limit is shown in Supporting Information Figure S2. The structural stability for the doped perovskites is measured by using the structural formation energy, which is represented by

$$E_{\text{form,A}} = -[(4mn - 1)E_{\text{Ba}} + mn(4E_{\text{Zr}} + 12E_{\text{S}}) + E_{\text{D}} - E_{\text{BaZrS}_3\text{-doped}}]/(20mn) \quad (2)$$

$$E_{\text{form,B}} = -[(4mn - 1)E_{\text{Zr}} + mn(4E_{\text{Ba}} + 12E_{\text{S}}) + E_{\text{D}} - E_{\text{BaZrS}_3\text{-doped}}]/(20mn) \quad (3)$$

where *m* and *n* are derived from the *x*- and *y*-directional expansion in a supercell with dimensions *m* × *n* × 1, *E_D* is the energy per atom of the most stable form of the dopant element, *E_{Ba}* and *E_{Zr}* are the energy per atom of the Ba and Zr atoms, respectively, and *E_{BaZrS₃}* is the total energy of the doped BaZrS₃ perovskite. The band gaps are all determined from the DOS (see Computational Methods).

The dopant concentration (“Concentration”) is also included in the dataset as well. In this study, 34 distinct dopants substituted at either the A or B site at different concentrations (8.33, 12.5, and 25%), with 4.17 and 6.25% A site Ca doping and all five concentrations (4.17, 6.25, 8.33, 12.5, and 25%) for B site Ti dopants (35 distinct dopants total, 211 dopants total) are investigated to build the ML models presented in this study. A correction is applied to the PBE band gaps by using multivariable linear regression to shift the band gap to the HSE band gap based on a relationship between the HSE and PBE band gaps based on a sample (56 dopants) of the dopants investigated. The correction applied to the PBE band gaps is $E_{\text{HSE}} = 6.815 - 9.699\eta_{\text{en}} + 0.155\eta_{\text{glawe}} + 1.293\eta_{\text{ue}} + 0.873E_{\text{PBE}}$ where η_{en} , η_{glawe} , and η_{ue} are the weighted chemical descriptors for the Pauling electronegativity (“en_pauling”), glawe number (“glawe_number”), and unpaired electrons (“Unpaired Electrons”), and *E_{PBE}* is the doped perovskite PBE band gap. The correlation strength of the predictions based on the PBE band gaps (*R*² = 0.762 for the test set and *R*² = 0.964 for the training set) is similar to the correlation strength obtained from the HSE shift (*R*² = 0.797 for the test set and *R*² = 0.971 for the training set). This relationship transforms the range of band gaps from 0.2 to 2.0 eV with PBE to 0.5–2.5 eV with the corrected HSE band gaps, which is similar to the effect seen when transforming from PBE to the corrected band gap when studying a variety of materials.⁶³ The MAE for the HSE vs PBE band gaps using this method has an MAE of 0.157 eV in comparison to simple linear regression with an MAE of 0.200 eV, which is comparable to the improvement seen with neural networks.⁶⁴

Materials Descriptors. An analogous relation to the structural formation energy (based on taking the average of the sum of the product of the number of each element and that element’s corresponding chemical descriptor for each doped perovskite) proved to be the best method of combining the chemical descriptors to predict the band gaps and formation energies. This analogous relationship forms weighted chemical descriptors, which are represented by

$$\eta_{\text{A}} = [(4mn - 1)\eta_{\text{Ba}} + mn(4\eta_{\text{Zr}} + 12\eta_{\text{S}}) + \eta_{\text{D}}]/(20mn) \quad (4)$$

$$\eta_{\text{B}} = [(4mn - 1)\eta_{\text{Zr}} + mn(4\eta_{\text{Ba}} + 12\eta_{\text{S}}) + \eta_{\text{D}}]/(20mn) \quad (5)$$

where η_{X} is the corresponding chemical descriptor value for an element X = Ba, Zr, and S, and η_{D} is the corresponding chemical descriptor value for the dopant. For example, for 6.25% A site Ca-doped BaZrS₃, with the chemical formula, Ba₁₅Ca₁Zr₁₆S₄₈, in a 2 × 2 × 1 supercell, the weighted chemical descriptor for the Pauling electronegativity (“en_pauling”) is

$$\eta_{\text{A}} = [x_{\text{Ba}}\eta_{\text{Ba}} + x_{\text{Ca}}\eta_{\text{Ca}} + x_{\text{Zr}}\eta_{\text{Zr}} + x_{\text{S}}\eta_{\text{S}}]/(20mn) \quad (6)$$

$$= [15 \cdot 0.89 + 1 \cdot 1.00 + 16 \cdot 1.33 + 48 \cdot 2.58]/(20 \cdot 2 \cdot 2) = 1.993375 \quad (7)$$

where *x* and η are the number and Pauling electronegativity for each element in this case. This procedure is performed for all the doped BaZrS₃ perovskites for each chemical descriptor investigated in this study. The chemical descriptors obtained from the Mendeleev database used in this study are the “atomic_radius_rahm”,^{36,37} “heat_of_formation”, “density”, “Electron Affinity”, “en_pauling”,⁵⁵ “glawe_number”,⁵⁶ “covalent_radius_pyykko”,⁵⁷ “dipole_polarizability”,⁵⁸ “covalent_radius_cordero”,⁵⁹ and “atomic_radius”.⁶⁰ Since the ionic radii (“Average I.R.”),^{61,62} “Oxidation States”, and “Unpaired

Electrons" can each yield multiple values, the average of each of those chemical descriptors for each element is used as an equivalent chemical descriptor for those elements in the dataset.

CGCNN is a materials representation that is created using a neural network. The materials descriptors and property prediction are combined into one algorithm. The materials descriptors are based on a graph theory representation of the crystal structures of the materials. Additional details on CGCNN are given below.

ML Methods. The complete workflow for the ML algorithms is shown in Supporting Information Figure S2. The ML algorithms including multivariable linear regression,^{43,44} random forest,^{45,46} and Crystal Graph Convolutional Neural Networks (CGCNN)^{47–50} are implemented using the scikit-learn Python module⁵¹ and PyTorch to predict both the structural formation energies and band gaps for all doped BaZrS₃ systems. All ML runs are performed by dividing the data into a test and training set with a 25%/75% split, with the correlation-squared coefficient (R^2) used to measure the strength of the correlation between the predicted and calculated values with the identity line shown as well. The descriptors used are constructed from the chemical element properties obtained from the Mendeleev database.

The band gaps and structural formation energies are studied using different ML algorithms to find the best way to predict these energies. The following models were examined: (i) linear regression forms a multidimensional linear combination from the inputted chemical descriptors by minimizing the residuals between the predicted and computed data. (ii) The random forest algorithm generates and trains an ensemble of decision trees where the importance of each descriptor in the tree is determined by the Gini index.^{52,53} (iii) CGCNN is a deep learning approach based on graph convolutional neural networks. Using a specialized graph convolution function, a CGCNN model encodes both properties of atomic sites and relationships between atoms in order to make predictions. The data from an input structure passes through a number of convolutional layers, followed by a number of fully connected layers, where the exact number of each has been optimized for best performance. Backpropagation is then used over a number of epochs to gradually optimize the parameter weights for both the fully connected and convolutional layers with respect to the loss function.⁵⁴ For the purposes of this research, a mean squared error loss function is used with respect to the target prediction.

Synthesis of the Ca-Doped BaZrO₃ Thin Film. 1.92 g of barium acetate (99%, AlfaAesar), 3.51 g of zirconium(IV) acetylacetonate (97%, Sigma-Aldrich), 0.07 g of calcium acetylacetonate (99.95%, Sigma-Aldrich), and 0.90 g of polyvinyl butyral (Sigma-Aldrich) were stirred and dissolved in 25 mL of propionic acid (99.5%, Sigma-Aldrich) at 60 °C to obtain the precursor solution for 4% Ca-doped films. The mass of zirconium(IV), acetylacetonate, and calcium acetylacetonate was varied accordingly to get 1, 2, and 3 Ca-atom % solutions. The resulting yellowish solution was spin-coated on a clean quartz substrate (1 cm × 1 cm × 2 mm) at 2000 rpm for 1 min, followed by 5000 rpm for 5 min. The spin-coated film was annealed in the air in a Thermolyne FB1315M muffle furnace at 700 °C for 90 min followed by 40 min at 870 °C.

Synthesis of the Ti-Doped BaZrO₃ Thin Film. The synthesis was done as described in our earlier work,²² details of which are described in the Supporting Information.

Sulfurization Procedure. The procedure was performed in an MTI OTF-1200X three-zone tube furnace (a quartz tube with 3" diameter) using CS₂ as the sulfur source at a temperature of 1050 °C for 4 h. A mass flow controller was used to keep the flow rate of the CS₂–N₂ mixture at ~5 sccm, while a pressure of ~2 Torr was maintained inside the tube. The whole process was carried out under a N₂ atmosphere. CS₂ was introduced at 600 °C and turned off when the furnace was cooled down to 600 °C. The tube was purged with ultrahigh purity N₂ until the furnace was cooled down completely. The tube was then brought to atmospheric pressure to take out the sulfurized film.

Materials Characterization. XRD results were obtained through a PANalytical X'Pert Pro diffractometer using Cu K α (λ = 1.5405 Å)

radiation with an X-ray generator at 45 kV and 40 mA. SEM was done on a Supra 55 FESEM at 2.5 kV. PL spectra were acquired using 532 nm laser excitation and an Andor spectrograph with a Peltier-cooled CCD detector. The laser power was measured by using a power meter. XPS was conducted using Al K α radiation (1486.7 eV) in a PHI 5000 Versa probe system. The spectra were collected at a pass energy of 23.5 eV and corrected for charging using the adventitious C 1s band.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsami.3c00618>.

Experimental details on synthesis of a Ti-doped BaZrO₃ thin film, DFT-computed band structures and defect formation energy, ML workflow, process flowchart to identify potential dopants, XRD of quartz substrates, XRD of doped and undoped BaZrO₃ films, variation of XRD peak fwhm with Ca-doping concentration, XPS survey spectra of doped and undoped BaZrS₃, high-resolution XPS spectra of Ba (5d) and Zr (3d), and PL patterns of BaZrS₃, Ti-doped BaZrS₃ and Ca-doped BaZrS₃ (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Humberto Terrones – Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, New York 12180, United States; orcid.org/0000-0003-0515-0721; Email: terroh@rpi.edu

Nikhil Koratkar – Department of Mechanical, Aerospace and Nuclear Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180, United States; orcid.org/0000-0002-4080-3786; Email: koratn@rpi.edu

Authors

Shyam Sharma – Department of Mechanical, Aerospace and Nuclear Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180, United States; orcid.org/0000-0001-9168-1857

Zachary D. Ward – Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, New York 12180, United States

Kevin Bhimani – Department of Mechanical, Aerospace and Nuclear Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180, United States

Mukul Sharma – Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180, United States

Joshua Quinton – Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, New York 12180, United States

Trevor David Rhone – Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, New York 12180, United States; orcid.org/0000-0002-0198-9952

Su-Fei Shi – Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180, United States; orcid.org/0000-0001-5158-805X

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acsami.3c00618>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

N.K. and H.T. acknowledge funding support from the USA National Science Foundation (Award 2013640). S.-F.S. acknowledges the support from NSF through the Career Award DMR-1945420 and grant DMR-2104902. TR acknowledges support from the National Science Foundation CAREER award under grant no 21 2044842. H.T. and Z.D.W. are grateful to the “Extreme Science and Engineering Discovery Environment” (XSEDE Now ACCESS), which is supported by National Science Foundation grant number ACI-1548562 through proposal TG-DMR170008.

REFERENCES

- (1) IEA (2021). *Renewables 2021—Analysis and Forecast to 2026*; IEA: Paris, 2021. <https://www.iea.org/reports/renewables-2021> (accessed Aug 12, 2022).
- (2) IEA (2021). *Net Zero by 2050—A Roadmap for the Global Energy Sector*; IEA: Paris, 2021. <https://www.iea.org/reports/net-zero-by-2050> (accessed Aug 12, 2022).
- (3) Kabir, E.; Kumar, P.; Kumar, S.; Adelodun, A. A.; Kim, K. H. Solar Energy: Potential and Future Prospects. *Renewable Sustainable Energy Rev.* **2018**, *82*, 894–900.
- (4) Wilson, G. M.; Al-Jassim, M.; Metzger, W. K.; Glunz, S. W.; Verlinden, P.; Xiong, G.; Mansfield, L. M.; Stanbery, B. J.; Zhu, K.; Yan, Y.; Berry, J. J.; Ptak, A. J.; Dimroth, F.; Kayes, B. M.; Tamboli, A. C.; Peibst, R.; Catchpole, K.; Reese, M. O.; Klinga, C. S.; Denholm, P.; Morjaria, M.; Deceglie, M. G.; Freeman, J. M.; Mikofski, M. A.; Jordan, D. C.; Tamizhmani, G.; Sulas-Kern, D. B. The 2020 Photovoltaic Technologies Roadmap. *J. Phys. D: Appl. Phys.* **2020**, *53*, 493001.
- (5) Extnance, A. The Reality behind Solar Power's next Star Material. *Nature* **2019**, *570*, 429–432.
- (6) Dubey, S.; Jadhav, N. Y.; Zakirova, B. Socio-Economic and Environmental Impacts of Silicon Based Photovoltaic (PV) Technologies. *Energy Procedia* **2013**, *33*, 322–334.
- (7) Andreani, L. C.; Bozzola, A.; Kowalczewski, P.; Liscidini, M.; Redorici, L. Silicon Solar Cells: Toward the Efficiency Limits. *Adv. Phys. X* **2018**, *4*, 1548305.
- (8) Shockley, W.; Queisser, H. J. Detailed Balance Limit of Efficiency of P-n Junction Solar Cells. *J. Appl. Phys.* **1961**, *32*, 510–519.
- (9) Sopiha, K. V.; Comparotto, C.; Márquez, J. A.; Scragg, J. J. S. Chalcogenide Perovskites: Tantalizing Prospects, Challenging Materials. *Adv. Opt. Mater.* **2022**, *10*, 2101704.
- (10) Sahare, S.; Pham, H. D.; Angmo, D.; Ghoderao, P.; MacLeod, J.; Khan, S. B.; Lee, S. L.; Singh, S. P.; Sonar, P. Emerging Perovskite Solar Cell Technology: Remedial Actions for the Foremost Challenges. *Adv. Energy Mater.* **2021**, *11*, 2101085.
- (11) Best Research-Cell Efficiency Chart Photovoltaic Research; NREL. <https://www.nrel.gov/pv/cell-efficiency.html> (accessed Dec 21, 2022).
- (12) Niu, S.; Huan, H.; Liu, Y.; Yeung, M.; Ye, K.; Blankemeier, L.; Orvis, T.; Sarkar, D.; Singh, D. J.; Kapadia, R.; Ravichandran, J. Band gap Control via Structural and Chemical Tuning of Transition Metal Perovskite Chalcogenides. *Adv. Mater.* **2017**, *29*, 1604733.
- (13) Jung, H. S.; Han, G. S.; Park, N. G.; Ko, M. J. Flexible Perovskite Solar Cells. *Joule* **2019**, *3*, 1850–1880.
- (14) Li, G.; Shittu, S.; Diallo, T. M. O.; Yu, M.; Zhao, X.; Ji, J. A Review of Solar Photovoltaic-Thermoelectric Hybrid System for Electricity Generation. *Energy* **2018**, *158*, 41–58.
- (15) Saule Technologies Launches The First Production Line. <https://sauletech.com/launching-production-line/> (accessed Aug 12, 2022).
- (16) Bellini, E. JinkoSolar announces perovskite cell technology platform promising efficiencies close to 30%—pv magazine International. <https://www.pv-magazine.com/2021/06/25/jinkosolar-announces-perovskite-cell-technology-platform-promising-efficiencies-close-to-30/> (accessed Aug 12, 2022).
- (17) Chandler, D. L. Explained: Why perovskites could take solar cells to new heights; MIT News Massachusetts Institute of Technology. MIT News Office. <https://news.mit.edu/2022/perovskites-solar-cells-explained-0715> (accessed Aug 12, 2022).
- (18) Gupta, T.; Ghoshal, D.; Yoshimura, A.; Basu, S.; Chow, P. K.; Lakhnot, A. S.; Pandey, J.; Warrender, J. M.; Efsthadias, H.; Soni, A.; Osei-Agyemang, E.; Balasubramanian, G.; Zhang, S.; Shi, S. F.; Lu, T. M.; Meunier, V.; Koratkar, N. An Environmentally Stable and Lead-Free Chalcogenide Perovskite. *Adv. Funct. Mater.* **2020**, *30*, 2001387–2001389.
- (19) Wei, X.; Hui, H.; Zhao, C.; Deng, C.; Han, M.; Yu, Z.; Sheng, A.; Roy, P.; Chen, A.; Lin, J.; Watson, D. F.; Sun, Y. Y.; Thomay, T.; Yang, S.; Jia, Q.; Zhang, S.; Zeng, H. Realization of BaZrS₃ Chalcogenide Perovskite Thin Films for Optoelectronics. *Nano Energy* **2020**, *68*, 104317.
- (20) Yu, Z.; Wei, X.; Zheng, Y.; Hui, H.; Bian, M.; Dhole, S.; Seo, J. H.; Sun, Y. Y.; Jia, Q.; Zhang, S.; Yang, S.; Zeng, H. Chalcogenide Perovskite BaZrS₃ Thin-Film Electronic and Optoelectronic Devices by Low Temperature Processing. *Nano Energy* **2021**, *85*, 105959.
- (21) Nishigaki, Y.; Nagai, T.; Nishiwaki, M.; Aizawa, T.; Kozawa, M.; Hanzawa, K.; Kato, Y.; Sai, H.; Hiramatsu, H.; Hosono, H.; Fujiwara, H. Extraordinary Strong Band-Edge Absorption in Distorted Chalcogenide Perovskites. *Sol. RRL* **2020**, *4*, 1900555.
- (22) Sharma, S.; Ward, Z.; Bhimani, K.; Li, K.; Lakhnot, A.; Jain, R.; Shi, S. F.; Terrones, H.; Koratkar, N. Band gap Tuning in BaZrS₃ Perovskite Thin Films. *ACS Appl. Electron. Mater.* **2021**, *3*, 3306–3312.
- (23) Wei, X.; Hui, H.; Perera, S.; Sheng, A.; Watson, D. F.; Sun, Y. Y.; Jia, Q.; Zhang, S.; Zeng, H. Ti-Alloying of BaZrS₃ Chalcogenide Perovskite for Photovoltaics. *ACS Omega* **2020**, *5*, 18579–18583.
- (24) Ueno, T.; Rhone, T. D.; Hou, Z.; Mizoguchi, T.; Tsuda, K. COMBO: An Efficient Bayesian Optimization Library for Materials Science. *Mater. Discovery* **2016**, *4*, 18–21.
- (25) Rhone, T. D.; Chen, W.; Desai, S.; Torrisi, S. B.; Larson, D. T.; Yacoby, A.; Kaxiras, E. Data-Driven Studies of Magnetic Two-Dimensional Materials. *Sci. Rep.* **2020**, *10*, 15795–15811.
- (26) Xie, Y.; Tritsaris, G. A.; Grånäs, O.; Rhone, T. D. Data-Driven Studies of the Magnetic Anisotropy of Two-Dimensional Magnetic Materials. *J. Phys. Chem. Lett.* **2021**, *12*, 12048–12054.
- (27) Agiorgousis, M. L.; Sun, Y.; Choe, D.; West, D.; Zhang, S. Machine Learning Augmented Discovery of Chalcogenide Double Perovskites for Photovoltaics. *Adv. Theory Simul.* **2020**, *3*, 1900200.
- (28) Ahmad, M. U.; Akib, A. R.; Raihan, M. M. S.; Shams, A. B. ABO₃ Perovskites' Formability Prediction and Crystal Structure Classification Using Machine Learning. *2022 International Conference on Innovations in Science, Engineering and Technology, ICISSET 2022*, 2022; pp 480–485.
- (29) Pilania, G.; Balachandran, P. V.; Gubernatis, J. E.; Lookman, T. Classification of ABO₃ Perovskite Solids: A Machine Learning Study. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2015**, *71*, 507–513.
- (30) Tao, Q.; Xu, P.; Li, M.; Lu, W. Machine Learning for Perovskite Materials Design and Discovery. *npj Comput. Mater.* **2021**, *7*, 23.
- (31) Zhao, J.; Wang, X. Screening Perovskites from ABO₃ Combinations Generated by Constraint Satisfaction Techniques Using Machine Learning. *ACS Omega* **2022**, *7*, 10483–10491.
- (32) Sharma, V.; Kumar, P.; Dev, P.; Pilania, G. Machine Learning Substitutional Defect Formation Energies in ABO₃ Perovskites. *J. Appl. Phys.* **2020**, *128*, 034902.
- (33) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (34) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid Functionals Based on a Screened Coulomb Potential. *J. Chem. Phys.* **2003**, *118*, 8207–8215.

- (35) Haynes, W. M.; Lide, D. R.; Bruno, T. J. *CRC Handbook of Chemistry and Physics*, 97th ed.; CRC Press, 2016; pp 10–147.
- (36) Rahm, M.; Hoffmann, R.; Ashcroft, N. W. Atomic and Ionic Radii of Elements 1–96. *Chem.—Eur. J.* **2016**, *22*, 14625–14632.
- (37) Rahm, M.; Hoffmann, R.; Ashcroft, N. W. Corrigendum: Atomic and Ionic Radii of Elements 1–96. *Chem.—Eur. J.* **2017**, *23*, 4017.
- (38) Zheng, L.; Fu, L.; He, Y.; Li, K.; Lu, Y.; Xue, J.; Liu, Y.; Dong, C.; Chen, C.; Tang, J. Fabrication and characterization of ZnO/Se1-xTex solar cells. *Front Optoelectron* **2022**, *15*, 36.
- (39) Kuhar, K.; Crovetto, A.; Pandey, M.; Thygesen, K. S.; Seger, B.; Vesborg, P. C. K.; Hansen, O.; Chorkendorff, I.; Jacobsen, K. W. Sulfide Perovskites for Solar Energy Conversion Applications: Computational Screening and Synthesis of the Selected Compound LaYs₃. *Energy Environ. Sci.* **2017**, *10*, 2579–2593.
- (40) Moussa, J. E.; Schultz, P. A.; Chelikowsky, J. R. Analysis of the Heyd-Scuseria-Ernzerhof Density Functional Parameter Space. *J. Chem. Phys.* **2012**, *136*, 204117.
- (41) Song, J. W.; Yamashita, K.; Hirao, K. Communication: A New Hybrid Exchange Correlation Functional for Band-Gap Calculations Using a Short-Range Gaussian Attenuation (Gaussian-Perdew–Burke–Ernzerhof). *J. Chem. Phys.* **2011**, *135*, 071103.
- (42) Viñes, F.; Lamiel-García, O.; Chul Ko, K.; Yong Lee, J.; Illas, F. Systematic Study of the Effect of HSE Functional Internal Parameters on the Electronic Structure and Band Gap of a Representative Set of Metal Oxides. *J. Comput. Chem.* **2017**, *38*, 781–789.
- (43) Maulud, D.; Abdulazeez, A. M. A Review on Linear Regression Comprehensive in Machine Learning. *J. Appl. Sci. Technol. Trends* **2020**, *1*, 140–147.
- (44) Kavitha, S.; Varuna, S.; Ramya, R. A Comparative Analysis on Linear Regression and Support Vector Regression. *Proceedings of 2016 Online International Conference on Green Engineering and Technologies, IC-GET 2016*, 2017.
- (45) Biau, G.; Scornet, E. A Random Forest Guided Tour. *TEST* **2016**, *25*, 197–227.
- (46) Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32.
- (47) Williams, L.; Mukherjee, A.; Dasgupta, A.; Rajan, K. Monitoring the Role of Site Chemistry on the Formation Energy of Perovskites via Deep Learning Analysis of Hirshfeld Surfaces. *J. Mater. Chem. C* **2021**, *9*, 11153–11162.
- (48) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **2018**, *120*, 145301.
- (49) Janiesch, C.; Zschech, P.; Heinrich, K. Machine Learning and Deep Learning. *Electron. Mark.* **2021**, *31*, 685–695.
- (50) Raschka, S.; Mirjalili, V. *Python Machine Learning : Machine Learning and Deep Learning with Python, Scikit-Learn, and Tensorflow 2*, 3rd ed.; Packt Publishing: UK, 2017.
- (51) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
- (52) Menze, B. H.; Kelm, B. M.; Masuch, R.; Himmelreich, U.; Bachert, P.; Petrich, W.; Hamprecht, F. A. A Comparison of Random Forest and Its Gini Importance with Standard Chemometric Methods for the Feature Selection and Classification of Spectral Data. *BMC Bioinf.* **2009**, *10*, 213.
- (53) Strobl, C.; Boulesteix, A.-L.; Zeileis, A.; Hothorn, T. Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. *BMC Bioinf.* **2007**, *8*, 25.
- (54) Lecun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444.
- (55) Haynes, W. M. *CRC Handbook of Chemistry and Physics*. 100 Key Points, 95th ed.; CRC Press: London, 2014; pp 9–97.
- (56) Glawe, H.; Sanna, A.; Gross, E. K. U.; Marques, M. A. L. The Optimal One Dimensional Periodic Table: A Modified Pettifor Chemical Scale from Data Mining. *New J. Phys.* **2016**, *18*, 093011.
- (57) Pyykkö, P.; Atsumi, M. Molecular Single-Bond Covalent Radii for Elements 1–118. *Chem.—Eur. J.* **2009**, *15*, 186–197.
- (58) Schwerdtfeger, P.; Nagle, J. K. 2018 Table of static dipole polarizabilities of the neutral elements in the periodic table. *Mol. Phys.* **2019**, *117*, 1200–1225.
- (59) Cordero, B.; Gómez, V.; Platero-Prats, A. E.; Revés, M.; Echeverría, J.; Cremades, E.; Barragán, F.; Alvarez, S. Covalent Radii Revisited. *Dalton Trans.* **2008**, 2832–2838.
- (60) Slater, J. C. Atomic Radii in Crystals. *J. Chem. Phys.* **1964**, *41*, 3199–3204.
- (61) Lundberg, D.; Persson, I. The Size of Actinoid(III) Ions – Structural Analysis vs. Common Misinterpretations. *Coord. Chem. Rev.* **2016**, *318*, 131–134.
- (62) Shannon, R. D. Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Crystallogr., Sect. A* **1976**, *32*, 751–767.
- (63) Wan, Z.; Wang, Q.-D.; Liu, D.; Liang, J. Effectively Improving the Accuracy of PBE Functional in Calculating the Solid Band Gap via Machine Learning. *Comput. Mater. Sci.* **2021**, *198*, 110699.
- (64) Lentz, L. C.; Kolpak, A. M. Predicting HSE Band Gaps from PBE Charge Densities via Neural Network Functionals. *J. Phys.: Condens. Matter* **2020**, *32*, 155901.
- (65) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.
- (66) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (67) Heyd, J.; Scuseria, G. E. Efficient Hybrid Density Functional Calculations in Solids: Assessment of the Heyd-Scuseria-Ernzerhof Screened Coulomb Hybrid Functional. *J. Chem. Phys.* **2004**, *121*, 1187–1192.