




RESEARCH ARTICLE | AUGUST 09 2023

## Pyrovskite: A software package for the high-throughput construction, analysis, and featurization of two- and three-dimensional perovskite systems

Robert Stanton  ; Dhara J. Trivedi  



*J. Chem. Phys.* 159, 064803 (2023)

<https://doi.org/10.1063/5.0159407>



CrossMark

500 kHz or 8.5 GHz?  
And all the ranges in between.

Lock-in Amplifiers for your periodic signal measurements



Find out more



# Pyroovskite: A software package for the high-throughput construction, analysis, and featurization of two- and three-dimensional perovskite systems

Cite as: J. Chem. Phys. 159, 064803 (2023); doi: 10.1063/5.0159407

Submitted: 23 May 2023 • Accepted: 24 July 2023 •

Published Online: 9 August 2023



Robert Stanton  and Dhara J. Trivedi<sup>a)</sup> 

## AFFILIATIONS

Department of Physics, Clarkson University, Potsdam, New York 13699, USA

<sup>a)</sup> Author to whom correspondence should be addressed: [dtrivedi@clarkson.edu](mailto:dtrivedi@clarkson.edu)

## ABSTRACT

The increased computational and experimental interest in perovskite systems comprising novel phases and reduced dimensionality has greatly expanded the search space for this class of materials. In similar fields, unified frameworks exist for the procedural generation and subsequent analysis of these complex condensed matter systems. Given the relatively recent rise in popularity of these novel perovskite phases, such a framework is yet to be created. In this work, we introduce Pyroovskite, an open source software package, to aid in both the high-throughput and fine-grained generation, simulation, and subsequent analysis of this expanded family of perovskite systems. Additionally, we introduce a new descriptor for octahedral distortions in systems, including, but not limited to, perovskites. This descriptor quantifies diagonal displacements of the B-site cation in a  $BX_6$  octahedral coordination environment, which has been shown to contribute to increased Rashba–Dresselhaus splitting in perovskite systems.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0159407>

## I. INTRODUCTION

In the last decade, perovskite systems have become ubiquitous in the search for high-efficiency photovoltaic devices.<sup>1–4</sup> With more recent developments, two-dimensional perovskites (2DPKs) have emerged as candidates, which maintain or improve upon many advantageous properties of their bulk counterparts, such as low excitation binding energies, geometrically tunable electronic properties, and inexpensive building components, while addressing the key concerns of bulk perovskites, which are long-term stability and device longevity.<sup>5–9</sup> Additionally, 2DPKs serve as a platform for the realization of physical properties seldom seen in their bulk counterparts, which make them favorable for usage beyond photovoltaics, such as Rashba–Dresselhaus splitting for spintronic applications, and strong electron–phonon coupling while maintaining narrow linewidths for light emitting applications.<sup>10–13</sup> The increased area of applicability, and, thereby, target properties, together with the widely expanded family of perovskites, which serve as potential candidate structures for material scientists, poses both experimental and computational

challenges. Large-scale, data driven, computational approaches in the realm of hybrid perovskites that contain organic cations remain limited in the literature, particularly in the realm of 2DPKs.<sup>14–18</sup> This is, in part, due to the complexity of the resultant structure, as well as the size of simulation cells, which frequently enter into the hundreds of atoms for even modest layer thicknesses.

Another challenge in dealing with perovskite systems is the structural diversity associated with the family of materials. Bulk  $ABX_3$  perovskites, double  $A_2BB'X_6$  perovskites, 2DPKs, and 2D-double perovskites form an extremely diverse set of materials in terms of phase, chemical composition, and their resulting geometric properties.<sup>19–21</sup> With this being the case, structure generation and characterization becomes a challenge, for which a comprehensive suite of tools, robust to the varied perovskite phases, dimensionalities, and diversity of organic cations is yet to be created to the author's knowledge. Characterizing the geometric properties associated with the relaxed perovskite crystal structures is of vital importance, as numerous structural features have been established as effective means of tuning the electronic properties of the resulting

material. For example, the literature is abundant with investigations characterizing the impact of organic cation orientation, octahedral tilting, organic spacer selection in 2D-perovskites, and octahedral distortions on the resulting electronic structure associated with the perovskite system.<sup>22–26</sup>

In this work, we introduce a software package, Pyrovskite, as a comprehensive suite of utilities for the (1) generation of perovskite crystal structures and electronic structure input files, (2) featurization of crystal structures with octahedral distortion parameters, (3) production of publication quality figures concerning the geometric properties of the system, and (4) interfacing the above-mentioned information with typical data structures used for subsequent data analysis or machine learning. Additionally, motivated by the work of Maurer *et al.* regarding the link between diagonal B-site cation displacements and increased Rashba–Dresselhaus spin-splitting, we introduce a new descriptor quantifying these diagonal displacements and compare it with pre-existing octahedral distortion parameters.<sup>27</sup>

The paper is structured as follows: In Sec. II, we discuss the relevance of distortions to the octahedral coordination environments present in many systems including perovskites, and we review the ubiquitous descriptors in the literature that quantify such distortions. Additionally, we highlight the difficulty, and, thereby, lack of computational solutions, associated with procedural generation of the expanded family of perovskites discussed in the Introduction. In Sec. III, we introduce a new descriptor for octahedral distortions, which can be computed with the Pyrovskite package, to quantify diagonal displacements of the B-site cation in BX<sub>6</sub> octahedral coordination environments, including, but not limited to, perovskites. Additionally, we discuss the implementation of procedural perovskite generation present within Pyrovskite, which is capable of addressing inorganic, hybrid, and double perovskites of varying phase and dimensionality. In Sec. IV, we comprehensively illustrate the capabilities of the package, including structure generation and characterization, the computation of geometric properties such as octahedral distortions, and the available interface with electronic structure and data analysis packages for machine learning. Finally, we present Sec. V containing conclusions and an outlook for future improvements to the package and its area of applicability.

## II. PROBLEM FORMULATION

### A. Descriptors for distortions in octahedral coordination environments

Octahedral coordination environments are present in a myriad of systems relevant to modern materials science, from spin-crossover and Werner-type complexes, to bulk metal-oxide systems, and, of particular interest to the present work, perovskites.<sup>28–34</sup> Common amongst these materials are distortions of the idealized octahedral coordination environment in the form of deviations in bond lengths and angles away from those present in the perfectly symmetric octahedra. These distortions are typically facilitated by the geometry or electronic structure of the local chemical environment, for example, by organic A-site cations in hybrid perovskites or Jan–Teller distortions in the case of transition-metal complexes.<sup>35–37</sup> The impact of the above-mentioned distortions on the resulting material properties have been the subject of considerable investigation for decades.<sup>38–40</sup> Given that these octahedral distortions typically serve as tuning mechanisms for a target material property,

such as the bandgap in a condensed matter system, a number of descriptors exist, dating back to 1971 by Robinson *et al.*, to distinguish and quantify different types of deviations from the idealized octahedral environment.<sup>41,42</sup> Two such descriptors describing the deviations are  $\Delta$  and  $\Sigma$  representing deviations in bond lengths and angles, given by Eqs. (1) and (2), respectively,

$$\Delta = \frac{1}{6} \sum_{i=1}^6 \left( \frac{d_i - d}{d} \right)^2, \quad (1)$$

$$\Sigma = \frac{1}{6} \sum_{i=1}^{12} |\phi_i - 90^\circ|. \quad (2)$$

In the context of BX<sub>6</sub> octahedra,  $d_i$  is the distance from the B-cation to the  $i$ 'th X-anion,  $d$  is the average B–X distance, and  $\phi_i$  is the  $i$ 'th *cis* X–B–X angle in the octahedra. The notation, as presented, is adopted from Ketkaew *et al.*, and the quantities are, in essence, the same as those described by Robinson *et al.*<sup>42,43</sup> An additional descriptor was formulated by Ketkaew *et al.*, quantifying twisting of the triangular octahedral faces; however, for this feature, we direct the reader to the OctaDist codebase.<sup>43</sup>

A recently highlighted form of octahedral distortion comes with the dislocation of B-site cations found in the BX<sub>6</sub> octahedra present in perovskite systems. The B-site displacement has been demonstrated by Maurer *et al.* to be tied to the strength of the Rashba–Dresselhaus spin-splitting exhibited by the perovskite system, which is of crucial importance to prospective spintronic applications.<sup>27</sup> In particular, the Rashba spin-splitting was demonstrated to correlate with diagonal displacements of the B-site cation relative to the octahedral coordination environment. In Sec. III, we formulate a descriptor to quantify diagonal displacements of the B-site cations and discuss their implementation in the Pyrovskite package.

### B. Procedural generation of hybrid perovskite systems

The increasing relevance of double perovskites, hybrid perovskites, 2DPKs, and their combinations for photovoltaic, thermoelectric, and ultraviolet radiation detection, and light-emitting applications has greatly increased the materials' search space relative to that of the previous consideration of only 3D perovskites.<sup>44–48</sup> The 2DPKs, in particular, contribute greatly to this increased structural diversity, as the phase-dependent chemical formula of A'<sub>2</sub>A<sub>*n*–1</sub>BX<sub>3*n*+1</sub> for the 2D Ruddlesden–Popper (2DRP) and monolayer configurations and A'<sub>2</sub>A<sub>*n*–1</sub>BX<sub>3*n*+1</sub> for the Dion–Jacobson (DJ) phase introduce two degrees of freedom not present in bulk perovskites. The A' spacing cation serves to decouple the adjacent inorganic networks of the BX<sub>6</sub> octahedra, both spatially, and electronically.<sup>49</sup> The choice of spacing the cations has been demonstrated to have non-trivial effects on properties ranging from the structural stability, device longevity, to even electronic structure of the resulting material.<sup>50–52</sup> Additionally, the layer thickness  $n$  determines the number of inorganic octahedra per layer, with electronic bandgaps converging from above to their bulk counterpart in the case of  $n \rightarrow \infty$ . These additional degrees of freedom create an entire family of structures for a given bulk ABX<sub>3</sub> perovskite, or A<sub>2</sub>BB'X<sub>6</sub> double perovskite, which often host very different physical properties in their dimensionally reduced form.<sup>53</sup>

This large search space creates the immediately apparent problem for computational researchers of resource limitations. However, a more conspicuous problem is in the reliable crystal structure generation of perovskites containing organic A- and A'-site cations of varied length, size, and shape in the 2DPKs. Interlayer alignment for phase specificity, interlayer spacing dictated by the interactions between BX<sub>6</sub> octahedra and A' cations, and the placement, as well as orientation of the organic A and A' building components, make the procedural generation of the newly diversified family of perovskite systems a challenge.

### III. IMPLEMENTATION

Herein, we introduce Pyroovskite, a Python package facilitating the high-throughput creation, simulation, and analysis of perovskite systems of varied phase, dimensionality, and chemical composition. Additionally, we note that all functionality associated with computation of octahedral distortions is applicable to any condensed phase or molecular system containing octahedral coordination environments. We first formalize the new  $\Lambda_2$  and  $\Lambda_3$  descriptors for quantifying the diagonal displacement of B-cations in BX<sub>6</sub> coordination environments. This is followed by a comprehensive overview of the user-friendly procedural perovskite generation present in the Pyroovskite package for bulk, double, 2D, and 2D-double perovskite phases.

#### A. Descriptor for the diagonal displacement of B-site cation

In a BX<sub>6</sub> octahedra respecting full octahedral symmetry, the midpoint of the three lines connecting *trans*-pairs of X atoms,  $M_{CD}$ ,  $M_{EF}$ , and  $M_{GH}$ , coincide directly with one another, and with the B-site cation at  $R_B$  in the center of the octahedra, termed  $P$  [as depicted in Fig. 1(a)],

$$M_{CD} = M_{EF} = M_{GH} = P = R_B. \quad (3)$$

Upon octahedral distortion, however, the midpoints of *trans*-X-X connections, separate, and the natural center of the octahedra become poorly defined. To remediate this problem, we define the natural center  $\tilde{P}$  of the BX<sub>6</sub> octahedra by considering the center of the three midpoints of the *trans*-pairs of X atoms, as in Eq. (4),

$$\tilde{P} = \frac{1}{3}(M_{CD} + M_{EF} + M_{GH}). \quad (4)$$

The displacement vector,  $S$ , between the B-cation and the natural center is given by Eq. (5),

$$S = R_B - \tilde{P}. \quad (5)$$

$\tilde{P}$  and  $S$  can be seen in the distorted octahedra in Fig. 1(b). In order to characterize the diagonal nature of the displacement vector  $S$ , we require reference to a natural basis of the given octahedra. To achieve this, we form the non-orthogonal basis [Eq. (6)] as the collection of unit vectors associated with the lines connecting the three *trans*-pairs of X atoms,  $CD$ ,  $EF$ , and  $GH$ ,

$$\{e_1, e_2, e_3\} = \left\{ \frac{CD}{|CD|}, \frac{EF}{|EF|}, \frac{GH}{|GH|} \right\}, \quad (6)$$

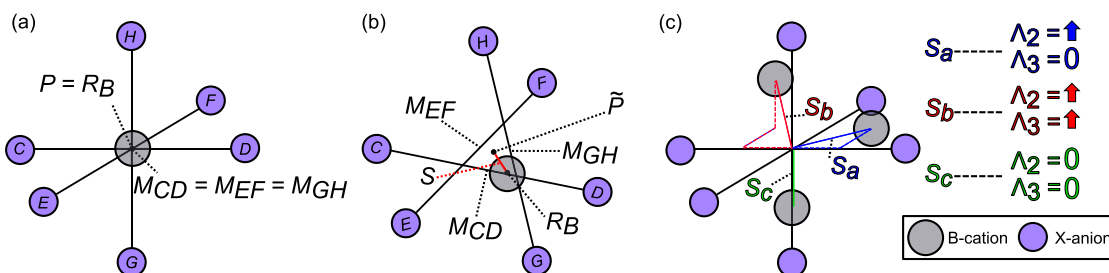
where  $CD = R_C - R_D$  and so on. The characterization of a diagonal displacement of the B-site cation is then informed by the ratio of the components of  $S$  when projected onto the basis of the octahedra. We define  $\lambda'_{ij}$  for the three unique pairs of  $S$  components in Eq. (7) such that  $\lambda'_{ij}$  are bounded by 0 and 1. A  $\lambda'_{ij}$  value approaching 0 then corresponds to a disproportionate displacement, almost entirely along the  $i$  or  $j$  direction, and a  $\lambda'_{ij}$  value approaching 1 corresponds to nearly equal displacement along the  $i$  and  $j$  directions. To differentiate between equally diagonal displacements (that is, collinear displacements relative to the natural center of the octahedra  $\tilde{P}$ ) of different magnitude, the  $\lambda_{ij}$  values are then obtained by scaling  $\lambda'_{ij}$  by a factor of  $(|S_i| + |S_j|)$ ,

$$\lambda'_{ij} = \min \left( \frac{|S_i|}{|S_j|}, \frac{|S_j|}{|S_i|} \right). \quad (7)$$

We then define a new set of descriptors that indicate diagonal displacements within one plane spanned by a pair of basis vectors of the octahedra and along all three planes spanned by the pairs of basis vectors as  $\Lambda_2$  and  $\Lambda_3$ , respectively. Figure 1(c) gives three examples of B-cation displacements and the qualitative corresponding impact on the  $\Lambda_2$  and  $\Lambda_3$  descriptors;

$$\Lambda_2 = \max(\lambda_{12}, \lambda_{13}, \lambda_{23}), \quad (8)$$

$$\Lambda_3 = \lambda_{12}\lambda_{13}\lambda_{23}. \quad (9)$$



**FIG. 1.** Schematic representation of an (a) undistorted, and (b) distorted BX<sub>6</sub> octahedra with the quantities from Sec. III A relevant to the definition of the labeled  $\Lambda_2$  and  $\Lambda_3$  descriptors. Distortions in (b) and (c) are exaggerated relative to typical octahedral distortions for clarity. (c) Three example placements of the B-cation in the BX<sub>6</sub> octahedra and the subsequent qualitative effect on the  $\Lambda_2$  and  $\Lambda_3$  descriptors.

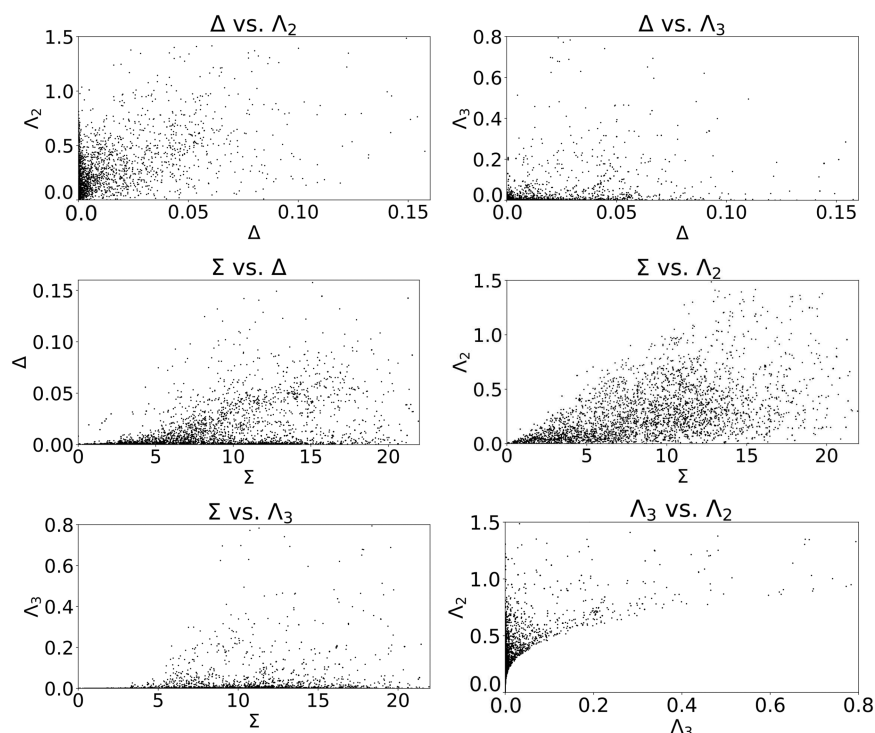


FIG. 2. Pairwise plots of the  $\Delta$ ,  $\Sigma$ ,  $\Lambda_2$ , and  $\Lambda_3$  descriptors, computed using the dataset described in Sec. III, by the Pyrovskite package.

The Pyrovskite package can be used to automatically detect the B-site and X-site ions, identify the corresponding octahedra, and compute the  $\Delta$ ,  $\Sigma$ ,  $\Lambda_2$ , and  $\Lambda_3$  descriptors. Figure 2 and Table I show the correlation of the above four parameters computed for a diverse dataset comprising over 1000 perovskites and, thereby, over 5000 perovskite octahedra. The full details of this dataset, as well as example code for the reproduction of such analysis, can be found in supplementary material Secs. I–III. The correlation coefficients between the  $\Lambda_2$  descriptor and the  $\Delta$  and  $\Sigma$  descriptors are the highest at 0.57, as is to be expected, as the in-plane diagonal displacement necessitates an increased deviation from average bond lengths and angles. Nonetheless, the descriptors still maintain a substantial degree of noncollinearity. Additionally, the correlation between  $\Lambda_3$ , and the  $\Delta$  and  $\Sigma$  descriptors is quite low at 0.36, and 0.21, respectively. The large difference in correlation of the new  $\Lambda_2$  and  $\Lambda_3$  parameters with that of the previously established

$\Delta$  and  $\Sigma$  descriptors offers a convenient way to select a feature that adequately quantifies the diagonal B-cation displacement, while avoiding collinearity with another feature describing the octahedral distortions present in the system.

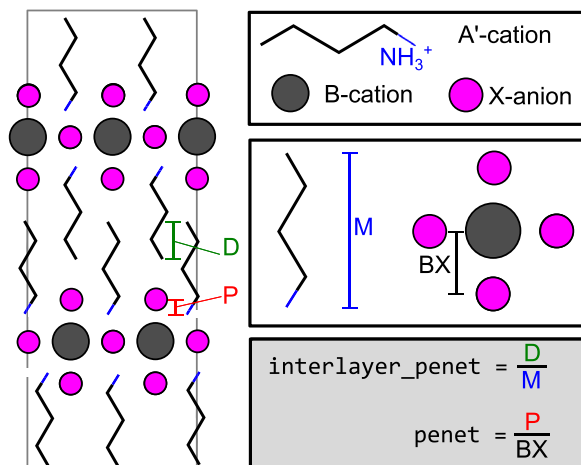
## B. Procedural generation of perovskite crystal structures

The Pyrovskite package contains the builder module, which allows for the rapid generation of both 2D and 3D perovskites in multiple phases, while including organic A-site cations and A' site cations where applicable. The builder module leverages the atomic simulation environment (ASE) Atoms object in order to facilitate construction of the perovskite systems.<sup>54</sup> For purely inorganic and bulk 3D perovskites, multiple tools exist for the generation of crystal structures of varying components, such as pymatgen.<sup>55</sup> However, for the incorporation of 2DPKs, and organic A-site cations, the authors are not aware of any software packages that address this need. The Pyrovskite builder module utilizes the functionality of the ASE Atoms object for the generation of all perovskite structures. A-site cations can be passed as either an Atoms object, or as a path to one of various input structure files such as .xyz, .sdf, and .pdb. A-site cation orientation can be adjusted, and the center of mass is situated at the corresponding crystal site of the given perovskite phase. In the case of 2DPKs, Fig. 3 demonstrates the key factors that determine the crystal structure. For example, a 2DPK in the  $n = 2$ , 2DRP phase can be created by simply specifying the path to any organic cations, the elemental symbol of any inorganic ions, the layer thickness  $n$ , and the desired B–X bond length:

TABLE I. Pearson correlation coefficients of the  $\Sigma$ ,  $\Delta$ ,  $\Lambda_2$ , and  $\Lambda_3$  descriptors across the dataset used in the present study.

Correlation coefficients	$\Sigma$ (°)	$\Delta$ (Å)	$\Lambda_2$ (Å)	$\Lambda_3$ (Å <sup>3</sup> )
$\Sigma$	1.00	0.41	0.57	0.21
$\Delta$	0.41	1.00	0.57	0.36
$\Lambda_2$	0.57	0.57	1.00	0.48
$\Lambda_3$	0.21	0.36	0.48	1.00





**FIG. 3.** Schematic representation detailing the procedural generation of 2DRPs by the Pyroovskite builder module. The `interlayer_penet` parameter describes the degree of interlocking present between the  $A'$  cations in the 2DRP phase, and the `penet` parameter describes the penetration depth of the  $A'$  cation into the adjacent inorganic octahedra. Hydrogens of the  $A'$ -cation are omitted for clarity, but we note that they are included for the computation of the molecule length  $M$ .

```

from pyroovskite.builder import make_2drp
from ase.io import read

A_file = "./methyllumonium.xyz"
Ap_file = "./butylammonium.xyz"

Ap = read(Ap_file)
A = read(A_file)

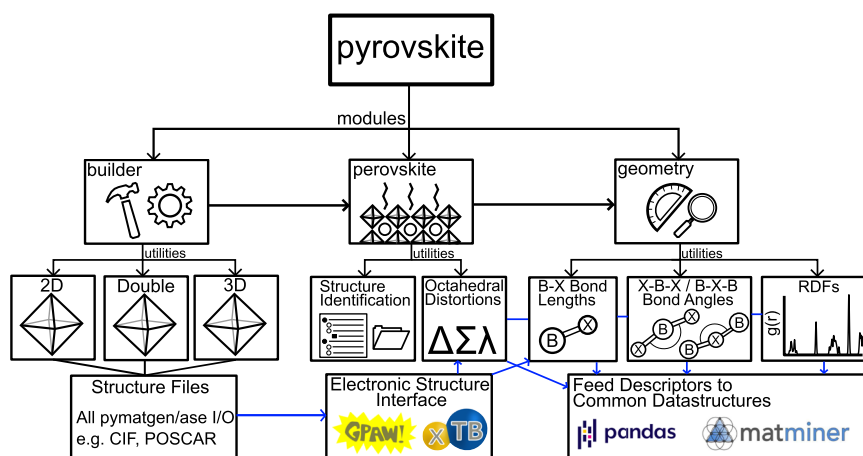
org_A_2drp = make_2drp(Ap, A, "Pb", "I", 2, 3)
in_A_2drp = make_2drp(Ap, "Cs", "Pb", "I", 2, 3)

```

$A'$  spacing cation locations are determined by the perovskite phase, as well as by the degree to which they penetrate into the inorganic  $BX_6$  octahedra. Additionally, in the case of 2DRP phases, the  $A'$  cation placement depends on the degree to which the adjacent organic  $A'$  cations interlock with one another. Sensible defaults are chosen for the interlayer penetration, `interlayer_penet`, and the octahedral penetration, `penet`, of 0.4 and 0.3, respectively. As shown in Fig. 3, `penet` is given in terms of a fraction of the  $B-X$  bond length for the system, and `interlayer_penet` is given in terms of a fraction of the  $A'$  cation length. The builder module allows for structure generation requiring only specification of the elemental and molecular building components, the perovskite phase, and the layer thickness in the case of 2DPKs. Additionally, more fine-grained control is available to the user, such as the ability to (1) modify the orientation of all organic  $A$  and  $A'$  cations, (2) adjust the layer penetration, and (3) introduce disorder to the created structure. The builder module is capable of creating (1) 3D  $ABX_3$  perovskites, (2) 3D  $A_2BB'X_6$  double perovskites, (3) 2DRP perovskites, (4) DJ perovskites, (5) 2D monolayer perovskites, (6) 2DRP double perovskites, (7) DJ double perovskites, and (8) 2D monolayer double perovskites. The package comes with numerous organic  $A$ -site and  $A'$  spacing cations, and including additional organic cations is straightforward and described in the documentation. Structures created using the builder module can be saved as any typical crystal structure file compatible with the pymatgen or ASE packages. Additionally, the builder module can supply the user with input files for subsequent electronic structure calculations using either GPAW or xTB via CP2K.<sup>56–59</sup>

#### IV. CAPABILITIES

The Pyroovskite codebase contains numerous utilities to aid in both high-throughput and fine-grained computational workflows involving perovskite systems, including crystal structure creation, electronic structure input file generation, structural analysis, and interfacing with commonly used tools, for subsequent data analysis or machine learning. The central datatype of the package is



**FIG. 4.** Schematic representation for the structure of the Pyroovskite codebase. A typical workflow beginning from structure generation, subsequent electronic structure calculations, computation of geometric properties, and exporting of data for further analysis is highlighted in blue.

the Perovskite class from which most functionality of the geometry and builder modules is exposed directly [as schematically shown in Fig. 4]. The Perovskite class can be instantiated either through the creation of a structure in the builder module or by passing any perovskite crystal structure by file name or as an ASE Atoms object. If instantiated using a crystal structure, additional information can be passed, including the identity of the A, A', B, and X building components, but is not necessary. The class is robust to the automatic detection of B-site cations and X-site anions, even in the presence of organic molecules that contain common X-site anions, such as fluorinated spacer molecules. Furthermore, the class detects B' cations in the case of double perovskites of either three or two dimensions. This greatly aids in applications that parse large datasets using tools such as the Materials Project API, where such information may not be directly available.<sup>60</sup> If the user has additional information regarding the oxidation state of the elements constituting the system, it is also straightforward to compute metrics such as the Goldschmidt tolerance factor and octahedral factor for the perovskite system of interest.<sup>61,62</sup> The automatic identification of all unique octahedra for computation of the distortion parameters discussed in Secs. II and III is also robust to a broad range of perovskite phases and chemical compositions, making the codebase suitable for most researchers in the field. Much of the functionality of the Pyrovskite package is also applicable to other systems that host octahedral coordination environments, including molecular systems containing transition-metal complexes, as well as condensed matter systems such as metal-organic frameworks.

A common use case for the package in structure generation has already been detailed in Sec. III, but, here, we highlight a more complete use case beginning with structure preparation and writing an xTB input file for usage with CP2K:

```
from pyrovskite.builder import make_2drp
from pyrovskite.perovskite import Perovskite
from ase.io import read

A_file = "./methyllummonium.xyz"
Ap_file = "./butylammonium.xyz"
Ap = read(Ap_file)
A = read(A_file)

new_perov = make_2drp(Ap, A, "Pb", "I", 2, 3.1)
perovskite_2drp = Perovskite(new_perov)
perovskite_2drp.write_xTB("perovskite_2drp.inp")
```

At this point, the user can run the electronic structure calculation, which, in the above scenario, defaults to a geometry optimization. The subsequent optimized structure can then be read by the Pyrovskite package to compute octahedral distortions, visualize the isolated octahedra, and create publication quality figures of the B–X bond length, X–B–X bond angle, and B–X–B bond angle distributions, as well as partial radial distribution functions of the resulting system. The B–X bond lengths and B–X–B bond angles, in particular, are known in the literature to have significant impact on the resultant electronic properties of the perovskite

system, largely owing to the fact that they manipulate the degree of overlap between frontier orbitals (often consisting of halogen p-bands and B-site cation p-bands for the valence and conduction bands, respectively):<sup>63</sup>

```
# Import and load in the optimized perovskite
from pyrovskite.perovskite import Perovskite
opt = Perovskite("./xTB_opt_perov_2drp.cif")

# Compute octahedral distortion parameters.
delta = opt_perovskite.compute_delta()
sigma = opt_perovskite.compute_sigma()
lambda3, lambda2 = \
    opt_perovskite.compute_lambda()

# Visualize the identified octahedra.
opt_perovskite.plot_vertices()

# Plot bond, angle distributions, and pRDFs
opt_perovskite.plot_distances()
opt_perovskite.plot_angles()
opt_perovskite.plot_rdf()
```

These features also work seamlessly with double perovskite systems, where B'–X bonds, as well as X–B'–X, B'–X–B, and B'–X–B' angles, will be included by default where applicable if a double perovskite structure is provided. All computed properties are stored in the Perovskite object, and in-built functions for exporting structures to pandas DataFrames exist for easy integration into, for example, MatMiner workflows.

## V. CONCLUSION AND OUTLOOK

Our open-source package, Pyrovskite, provides an all-inclusive platform for the generation and analysis of perovskite systems, as well as a seamless interface with multiple electronic structure codes and external resources, to aid in large-scale data analysis. Additionally, the octahedral distortion descriptors introduced here generalize to any octahedral coordination environment, greatly expanding the utility of the package. The octahedra of both molecular systems, as well as those of non-perovskite condensed matter systems, can be automatically detected, allowing for the facile computation of octahedral distortion parameters beyond the scope of perovskites. The codebase is under active development and will, in the future, contain automatic phase detection, characterization of octahedral tilt patterns, and pre-trained models for the prediction of material properties from both chemical formulas and crystal structures.

## SUPPLEMENTARY MATERIAL

The supplementary material contains additional information about the perovskite dataset used to evaluate the octahedral distortion parameters, information regarding the usability and extensibility of the code base, and additional code examples demonstrating the functionality of the package.

## ACKNOWLEDGMENTS

This work was supported by the U. S. National Science Foundation Engineering Research Initiation Award No. ECCS-2138728.

## AUTHOR DECLARATIONS

## Conflict of Interest

The authors have no conflicts to disclose.

## Author Contributions

**Robert Stanton:** Formal analysis (equal); Methodology (equal); Validation (equal); Writing – original draft (equal). **Dhara J. Trivedi:** Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Supervision (equal); Writing – review & editing (equal).

## DATA AVAILABILITY

This software package is openly available in GitHub at <https://github.com/r2stanton/pyrovskite>.

## REFERENCES

- 1 R. Wang, M. Mujahid, Y. Duan, Z.-K. Wang, J. Xue, and Y. Yang, “A review of perovskites solar cell stability,” *Adv. Funct. Mater.* **29**, 1808843 (2019).
- 2 X. Zhang and S.-H. Wei, “Origin of efficiency enhancement by lattice expansion in hybrid-perovskite solar cells,” *Phys. Rev. Lett.* **128**, 136401 (2022).
- 3 J.-P. Correa-Baena, A. Abate, M. Saliba, W. Tress, T. Jesper Jacobsson, M. Grätzel, and A. Hagfeldt, “The rapid evolution of highly efficient perovskite solar cells,” *Energy Environ. Sci.* **10**, 710–727 (2017).
- 4 Y.-C. Hsiao, T. Wu, M. Li, Q. Liu, W. Qin, and B. Hu, “Fundamental physics behind high-efficiency organo-metal halide perovskite solar cells,” *J. Mater. Chem. A* **3**, 15372–15385 (2015).
- 5 A. Krishna, S. Gottis, M. K. Nazeeruddin, and F. Sauvage, “Mixed dimensional 2D/3D hybrid perovskite absorbers: The future of perovskite solar cells?,” *Adv. Funct. Mater.* **29**, 1806482 (2019).
- 6 G. Wu, R. Liang, Z. Zhang, M. Ge, G. Xing, and G. Sun, “2D hybrid halide perovskites: Structure, properties, and applications in solar cells,” *Small* **17**, 2103514 (2021).
- 7 J.-C. Blancon, J. Even, C. C. Stoumpos, M. G. Kanatzidis, and A. D. Mohite, “Semiconductor physics of organic-inorganic 2D halide perovskites,” *Nat. Nanotechnol.* **15**, 969–985 (2020).
- 8 R. Stanton, S. K. Gupta, and D. J. Trivedi, “Probing strain-induced effects on performance of low-dimensional hybrid perovskites for solar energy harvesting,” *ACS Appl. Mater. Interfaces* **14**, 34603–34611 (2022).
- 9 E. S. Vasileiadou, I. Hadar, M. Kepenekian, J. Even, Q. Tu, C. D. Malliakas, D. Friedrich, I. Spanopoulos, J. M. Hoffman, V. P. Dravid, and M. G. Kanatzidis, “Shedding light on the stability and structure-property relationships of two-dimensional hybrid lead bromide perovskites,” *Chem. Mater.* **33**, 5085–5107 (2021).
- 10 J. A. Steele, P. Puech, B. Monserrat, B. Wu, R. X. Yang, T. Kirchartz, H. Yuan, G. Fleury, D. Giovanni, E. Fron, M. Keshavarz, E. Debroye, G. Zhou, T. C. Sum, A. Walsh, J. Hofkens, and M. B. J. Roeffaers, “Role of electron-phonon coupling in the thermal evolution of bulk Rashba-like spin-split lead halide perovskites exhibiting dual-band photoluminescence,” *ACS Energy Lett.* **4**, 2205–2212 (2019).
- 11 J. Li, G. Lu, H. Xie, X. Jiang, Z. Fang, Y.-J. Zeng, H. He, and Z. Ye, “Effect of electron-phonon coupling on the color purity of two-dimensional Ruddlesden-Popper hybrid lead iodide perovskites,” *J. Phys. Chem. C* **127**, 6380–6388 (2023).
- 12 K. Leng, R. Li, S. P. Lau, and K. P. Loh, “Ferroelectricity and Rashba effect in 2D organic-inorganic hybrid perovskites,” *Trends Chem.* **3**, 716–732 (2021).
- 13 S. Bhattacharya and Y. Kanai, “Spin-orbit-coupling-induced band splitting in two-dimensional hybrid organic-inorganic perovskites: Importance of organic cations,” *Phys. Rev. Mater.* **7**, 055001 (2023).
- 14 J. Ovčar, L. Grisanti, B. Mladineo, A. B. Djurišić, J. Popović, and I. Lončarić, “Crystal structure prediction of quasi-two-dimensional lead halide perovskites,” *Phys. Rev. B* **107**, 174109 (2023).
- 15 G. Zhao, J. Xie, K. Zhou, B. Xing, X. Wang, F. Tian, X. He, and L. Zhang, “High-throughput computational material screening of the cycloalkane-based two-dimensional Dion-Jacobson halide perovskites for optoelectronics,” *Chin. Phys. B* **31**, 037104 (2022).
- 16 E. I. Marchenko, S. A. Fateev, A. A. Petrov, V. V. Korolev, A. Mitrofanov, A. V. Petrov, E. A. Goodilin, and A. B. Tarasov, “Database of two-dimensional hybrid perovskite materials: Open-access collection of crystal structures, band gaps, and atomic partial charges predicted by machine learning,” *Chem. Mater.* **32**, 7383–7388 (2020).
- 17 C. Kim, T. D. Huan, S. Krishnan, and R. Ramprasad, “A hybrid organic-inorganic perovskite dataset,” *Sci. Data* **4**, 170057 (2017).
- 18 W. Hu and L. Zhang, “High-throughput calculation and machine learning of two-dimensional halide perovskite materials: Formation energy and band gap,” *Mater. Today Commun.* **35**, 105841 (2023).
- 19 N. R. Wolf, B. A. Connor, A. H. Slavney, and H. I. Karunadasa, “Doubling the stakes: The promise of halide double perovskites,” *Angew. Chem., Int. Ed.* **60**, 16264–16278 (2021).
- 20 W. Guo, Z. Yang, J. Dang, and M. Wang, “Progress and perspective in Dion-Jacobson phase 2D layered perovskite optoelectronic applications,” *Nano Energy* **86**, 106129 (2021).
- 21 J. Qian, S. Yu, B. Xing, X. He, and L. Zhang, “Two-dimensional Ruddlesden-Popper halide perovskite solar absorbers with short-chain interlayer spacers,” *Phys. Rev. Mater.* **6**, 065405 (2022).
- 22 J. Yin, Z. Xu, Q. Hu, G. Teobaldi, L.-M. Liu, and O. V. Prezhdo, “Tuning octahedral tilting by doping to prevent detrimental phase transition and extend carrier lifetime in organometallic perovskites,” *J. Am. Chem. Soc.* **145**, 5393–5399 (2023).
- 23 S. Krach, N. Forero-Correa, R.-I. Biega, S. E. Reyes-Lillo, and L. Lepert, “Emergence of Rashba-/Dresselhaus effects in Ruddlesden-Popper halide perovskites with octahedral rotations,” *J. Phys.: Condens. Matter* **35**, 174001 (2023).
- 24 K. Ji, Z. Wu, X. Shen, J. Wang, and J. Zhang, “Jahn-Teller distortion driven ferromagnetism in a perovskite fluoride monolayer,” *Phys. Rev. B* **107**, 134431 (2023).
- 25 R. Stanton and D. J. Trivedi, “Atomistic description of the impact of spacer selection on two-dimensional (2D) perovskites: A case study of 2D Ruddlesden-Popper CsPbI<sub>3</sub> analogues,” *J. Phys. Chem. Lett.* **13**, 12090–12098 (2022).
- 26 J. Li and P. Rinke, “Atomic structure of metal-halide perovskites from first principles: The chicken-and-egg paradox of the organic-inorganic interaction,” *Phys. Rev. B* **94**, 045201 (2016).
- 27 B. Maurer, C. Vorwerk, and C. Draxl, “Rashba and Dresselhaus effects in two-dimensional Pb-I-based perovskites,” *Phys. Rev. B* **105**, 155149 (2022).
- 28 S. Jang and H. Kim, “Understanding the origin of the chiral recognition of esters with octahedral chiral cobalt complexes,” *Asian J. Org. Chem.* **10**, 886–890 (2021).
- 29 B. Kim, J. Yu, H. Koh, I. Nagai, S. Ikeda, S.-J. Oh, and C. Kim, “Missing *xy*-band fermi surface in 4d transition-metal oxide Sr<sub>2</sub>RhO<sub>4</sub>: Effect of the octahedra rotation on the electronic structure,” *Phys. Rev. Lett.* **97**, 106401 (2006).
- 30 L. Su, H. Huan, A. Sarkar, W. Gao, X. Yan, C. Addiego, R. Kruk, H. Hahn, and X. Pan, “Direct observation of elemental fluctuation and oxygen octahedral



distortion-dependent charge distribution in high entropy oxides," *Nat. Commun.* **13**, 2358 (2022).

- <sup>31</sup>Y. Zhou, S. Dong, C. Shan, K. Ji, and J. Zhang, "Two-dimensional ferroelectricity induced by octahedral rotation distortion in perovskite oxides," *Phys. Rev. B* **105**, 075408 (2022).
- <sup>32</sup>B. Zhou, J. Zhang, X. Ye, G. Liu, X. Xu, J. Wang, Z. Liu, L. Zhou, Z. Liao, H. Yao, S. Xu, J. Shi, X. Shen, X. Yu, Z. Hu, H. Lin, C. Chen, X. Qiu, C. Dong, J. Zhang, R. Yu, P. Yu, K. Jin, Q. Meng, and Y. Long, "Octahedral distortion and displacement-type ferroelectricity with switchable photovoltaic effect in a  $3d^3$ -electron perovskite system," *Phys. Rev. Lett.* **130**, 146101 (2023).
- <sup>33</sup>R. W. Hogue, S. Singh, and S. Brooker, "Spin crossover in discrete polynuclear iron(II) complexes," *Chem. Soc. Rev.* **47**, 7303–7338 (2018).
- <sup>34</sup>T. Wei, L. Wang, K. Mao, J. Chen, J. Dai, Z. Zhang, L. Liu, and X. Wu, "Polarization-induced efficient charge separation in an electromagnetic coupling MOF for enhancing CO<sub>2</sub> photocatalytic reduction," *J. Colloid Interface Sci.* **622**, 402–409 (2022).
- <sup>35</sup>X. Wang, W. Meng, W. Liao, J. Wang, R.-G. Xiong, and Y. Yan, "Atomistic mechanism of broadband emission in metal halide perovskites," *J. Phys. Chem. Lett.* **10**, 501–506 (2019).
- <sup>36</sup>X. Xiao, M. Wu, Z. Ni, S. Xu, S. Chen, J. Hu, P. N. Rudd, W. You, and J. Huang, "Ultrafast exciton transport with a long diffusion length in layered perovskites with organic cation functionalization," *Adv. Mater.* **32**, 2004080 (2020).
- <sup>37</sup>S. Liu, R. Guo, and F. Xie, "The effects of organic cation rotation in hybrid organic-inorganic perovskites: A critical review," *Mater. Des.* **221**, 110951 (2022).
- <sup>38</sup>P. Radaelli, M. Marezio, H. Hwang, S. Cheong, and B. Batlogg, "Charge localization by static and dynamic distortions of the MnO<sub>6</sub> octahedra in perovskite manganites," *Phys. Rev. B* **54**, 8992 (1996).
- <sup>39</sup>J. He, A. Borisevich, S. V. Kalinin, S. J. Pennycook, and S. T. Pantelides, "Control of octahedral tilts and magnetic properties of perovskite oxide heterostructures by substrate symmetry," *Phys. Rev. Lett.* **105**, 227203 (2010).
- <sup>40</sup>J.-S. Zhou and J. Goodenough, "Universal octahedral-site distortion in orthorhombic perovskite oxides," *Phys. Rev. Lett.* **94**, 065501 (2005).
- <sup>41</sup>L.-y. Huang and W. R. Lambrecht, "Electronic band structure trends of perovskite halides: Beyond Pb and Sn to Ge and Si," *Phys. Rev. B* **93**, 195211 (2016).
- <sup>42</sup>K. Robinson, G. Gibbs, and P. Ribbe, "Quadratic elongation: A quantitative measure of distortion in coordination polyhedra," *Science* **172**, 567–570 (1971).
- <sup>43</sup>R. Ketkaew, Y. Tantirungrotechai, P. Harding, G. Chastanet, P. Guionneau, M. Marchivie, and D. J. Harding, "Octadist: A tool for calculating distortion parameters in spin crossover and coordination complexes," *Dalton Trans.* **50**, 1086–1096 (2021).
- <sup>44</sup>M. Sajjad, Q. Mahmood, N. Singh, and J. A. Larsson, "Ultralow lattice thermal conductivity in double perovskite Cs<sub>2</sub>PtI<sub>6</sub>: A promising thermoelectric material," *ACS Appl. Energy Mater.* **3**, 11293–11299 (2020).
- <sup>45</sup>F. Aslam, B. Sabir, and M. Hassan, "Structural, electronic, optical, thermoelectric, and transport properties of indium-based double perovskite halides Cs<sub>2</sub>InAgX<sub>6</sub> (X = Cl, Br, I), for energy applications," *Appl. Phys. A* **127**, 112 (2021).
- <sup>46</sup>Z. Xu, W. Weng, Y. Li, X. Liu, T. Yang, M. Li, X. Huang, J. Luo, and Z. Sun, "3D-to-2D dimensional reduction for exploiting a multilayered perovskite ferroelectric toward polarized-light detection in the solar-blind ultraviolet region," *Angew. Chem., Int. Ed.* **59**, 21693–21697 (2020).
- <sup>47</sup>L. Zhang, C. Sun, T. He, Y. Jiang, J. Wei, Y. Huang, and M. Yuan, "High-performance quasi-2D perovskite light-emitting diodes: From materials to devices," *Light Sci. Appl.* **10**, 61 (2021).
- <sup>48</sup>R. Chakraborty, G. Paul, and A. J. Pal, "Quantum confinement and dielectric deconfinement in quasi-two-dimensional perovskites: Their roles in light-emitting diodes," *Phys. Rev. Appl.* **17**, 054045 (2022).
- <sup>49</sup>M. R. Filip, D. Y. Qiu, M. Del Ben, and J. B. Neaton, "Screening of excitons by organic cations in quasi-two-dimensional organic-inorganic lead-halide perovskites," *Nano Lett.* **22**, 4870–4878 (2022).
- <sup>50</sup>Z. Xu, L. Li, X. Dong, D. Lu, R. Wang, W.-J. Yin, and Y. Liu, "CsPbI<sub>3</sub>-based phase-stable 2D Ruddlesden-Popper perovskites for efficient solar cells," *Nano Lett.* **22**, 2874–2880 (2022).
- <sup>51</sup>K. Zheng and T. Pullerits, "Two dimensions are better for perovskites," *J. Phys. Chem. Lett.* **10**, 5881 (2019).
- <sup>52</sup>Y. Yan, S. Yu, A. Honarfar, T. Pullerits, K. Zheng, and Z. Liang, "Benefiting from spontaneously generated 2D/3D bulk-heterojunctions in Ruddlesden-Popper perovskite by incorporation of S-bearing spacer cation," *Adv. Sci.* **6**, 1900548 (2019).
- <sup>53</sup>Y. Lei, G. Peng, H. Wang, G. Wang, S. Yang, Q. Wang, Z. Li, and Z. Jin, "2D Dion-Jacobson CsPbI<sub>3</sub> with enhanced interlayer coupling for stable and efficient photovoltaics," *Adv. Mater. Interfaces* **9**, 2201501 (2022).
- <sup>54</sup>A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dulak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, "The atomic simulation environment—A Python library for working with atoms," *J. Phys.: Condens. Matter* **29**, 273002 (2017).
- <sup>55</sup>S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, "Python materials genomics (pymatgen): A robust, open-source Python library for materials analysis," *Comput. Mater. Sci.* **68**, 314–319 (2013).
- <sup>56</sup>J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Dulak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. Hansen, H. H. Kristoffersen, M. Kuisma, A. H. Larsen, L. Lehtovaara, M. Ljungberg, O. Lopez-Acevedo, P. G. Moses, J. Ojanen, T. Olsen, V. Petzold, N. A. Romero, J. Stausholm-Møller, M. Strange, G. A. Tritsarlis, M. Vanin, M. Walter, B. Hammer, H. Häkkinen, G. K. H. Madsen, R. M. Nieminen, J. K. Nørskov, M. Puskas, T. T. Rantala, J. Schiøtz, K. S. Thygesen, and K. W. Jacobsen, "Electronic structure calculations with GPAW: A real-space implementation of the projector augmented-wave method," *J. Phys.: Condens. Matter* **22**, 253202 (2010).
- <sup>57</sup>C. Bannwarth, S. Ehlert, and S. Grimme, "GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions," *J. Chem. Theory Comput.* **15**, 1652–1671 (2019).
- <sup>58</sup>J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, "CP2K: atomistic simulations of condensed matter systems," *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **4**, 15–25 (2014).
- <sup>59</sup>T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann *et al.*, "CP2K: An electronic structure and molecular dynamics software package-quickstep: Efficient and accurate electronic structure calculations," *J. Chem. Phys.* **152**, 194103 (2020).
- <sup>60</sup>A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, "Commentary: The materials project: A materials genome approach to accelerating materials innovation," *APL Mater.* **1**, 011002 (2013).
- <sup>61</sup>V. M. Goldschmidt, "Die gesetze der krystallochemie," *Naturwissenschaften* **14**, 477–485 (1926).
- <sup>62</sup>C. Li, K. C. K. Soh, and P. Wu, "Formability of ABO<sub>3</sub> perovskites," *J. Alloys Compd.* **372**, 40–48 (2004).
- <sup>63</sup>G. Liu, J. Gong, L. Kong, R. D. Schaller, Q. Hu, Z. Liu, S. Yan, W. Yang, C. C. Stoumpos, M. G. Kanatzidis, H. K. Mao, and T. Xu, "Isothermal pressure-derived metastable states in 2D hybrid perovskites showing enduring bandgap narrowing," *Proc. Natl. Acad. Sci. U. S. A.* **115**, 8076–8081 (2018).