Perovskite-inspired photovoltaic materials: Toward best practices in materials characterization and calculations

Robert L. Z. Hoye,*‡◊ Philip Schulz,‡◊ Laura T. Schelhas,**◊ Aaron M. Holder,‡◊◊ Kevin H. Stone,‡◊ John D. Perkins,‡◊ Derek Vigil-Fowler,‡◊ Sebastian Siol,‡◊ David O. Scanlon,§◊ Andriy Zakutayev,‡ Aron Walsh,§ Ian C. Smith,§ Brent C. Melot,§ Rachel C. Kurchin,† Yiping Wang,‡ Jian Shi,‡ Francisco C. Marques,‡ Joseph J. Berry,‡ William Tumas,‡ Stephan Lany,‡◊ Vladan Stevanovic,†◊ Michael F. Toney,**◊ and Tonio Buonassisi*‡◊

†Massachusetts Institute of Technology, Cambridge, MA 02139, USA
‡National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, Colorado 80401, USA
◊SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
‡Department of Chemical and Biological Engineering, University of Colorado, Boulder, Colorado 80309, USA
§University College London, Kathleen Lonsdale Materials Chemistry, Department of Chemistry, 20 Gordon Street, London WC1H 0AJ, UK
^Diamond Light Source, Ltd., Diamond House, Harwell Science and Innovation Campus, Didcot, Oxfordshire OX11 0DE, UK
ßImperial College London, Department of Materials, Exhibition Road, London SW7 2AZ, UK
θDepartment of Chemistry, Stanford University, Stanford, CA 94305, USA
§Department of Chemistry, University of Southern California, Los Angeles, CA 90089, USA
¶Department of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy, NY, 12180, USA
¥University of Campinas, Physics Institute, 13083-859, Campinas, SP, Brazil
ǁColorado School of Mines, 1500 Illinois Street, Golden, Colorado 80401, USA

◊The article primarily reflects the views of these authors, who were primarily responsible for writing the text and compiling the data

ABSTRACT: Recently, there has been an explosive growth in research based on hybrid lead-halide perovskites for photovoltaics owing to rapid improvements in efficiency. The advent of these materials for solar applications has led to widespread interest in understanding the key enabling properties of these materials. This has resulted in renewed interest in related compounds and a search for materials that may replicate the defect-tolerant properties and long lifetimes of the hybrid lead-halide perovskites. Given the rapid pace of development of the field, the rises in efficiencies of these systems have outpaced the more basic understanding of these materials. Measuring or calculating the basic properties, such as crystal/electronic structure and composition, can be challenging because some of these materials have anisotropic structures, and/or are composed of both heavy metal cations and volatile, mobile, light elements. Some consequences are beam damage during characterization, composition change under vacuum, or compound effects, such as the alteration of the electronic structure through the influence of the substrate. These effects make it challenging to understand the basic properties integral to optoelectronic operation. Compounding these difficulties is the rapid pace with which the field progresses. This has created an ongoing need to continually evaluate best practices with respect to characterization and calculations, as well as to identify inconsistencies in reported values to determine if those inconsistencies are rooted in characterization methodology or materials synthesis. This article describes the difficulties in characterizing hybrid lead-halide perovskites and new materials, and how these challenges may be overcome. The topic was discussed at a seminar at the 2015 Materials Research Society Fall Meeting & Exhibit. This article highlights the lessons learned from the seminar and the insights of some of the attendees, with reference to both recent literature and controlled experiments to illustrate the
challenges discussed. The focus in this article is on crystallography, composition measurements, photoemission spectroscopy and calculations on perovskites and new, related absorbers. We suggest how some of the important artifacts could be avoided and how the reporting for each technique could be streamlined between groups to ensure reproducibility as the field progresses.

INTRODUCTION
Solar cells based on hybrid organic-inorganic lead-halide perovskites have rapidly risen in efficiency in recent years, gaining increasing scientific and industrial interest. The primary challenges with this material are concerns over the toxicity of lead, and the limited photo- thermo- and chemical stability in the presence of moist air. As a result, many researchers are interested in finding lead-free alternatives. A wide variety of materials have recently been explored for photovoltaics. These include compounds previously investigated for other applications, as well as completely new compounds. Rigorously establishing the basic physical properties of these materials is essential, but some compositions present challenges with respect to the use of characterization techniques developed for conventional inorganic materials.

Some of the alternatives to the hybrid lead-halide perovskites that have been investigated substitute the Pb$^{2+}$ cation with other metal cations in a +2 oxidation state, such as Sn$^{2+}$, Ge$^{2+}$, and Cs$^+$. The AMX$_3$ composition ($A$ = monovalent cation, $M$ = metal cation, $X$ = halide) is thus maintained. Examples include formamidinium cesium tin iodide (as well as lead-tin mixtures), methylammonium tin iodide, formamidinium tin iodide, cesium germanium iodide, methylammonium germanium iodide, and formamidinium germanium iodide. Double perovskites, such as Cs$_2$BiAgCl$_6$, Cs$_2$BiAgBr$_6$, Cs$_2$BiAgCl$_6$, and have also been found to have promising optoelectronic properties.

Others in the community have searched for ‘perovskite-inspired materials’ or PIMs. The search for PIMs is driven by using theory to learn from the hybrid lead-halide perovskites to develop new design rules for solar absorbers. One of these design rules relates to the electronic structure. The Pb$^{2+}$ cation in MAPbI$_3$ has a lone pair of 6s electrons which hybridize with the I$^-$ 5p electrons to form an antibonding orbital at the valence band maximum. The Pb$^{2+}$ 6p orbital also hybridizes with the I$^-$ 5p orbital to form a bonding orbital in the valence band and antibonding orbital in the conduction band minimum. This electronic structure may result in tolerance to intrinsic defects, because dangling bonds may be more likely to form states in the bands or close to the band edges. Materials that can replicate this type of electronic structure, but not necessarily the crystallographic structure, include those with In$^+$, Sn$^{2+}$, Sb$^{3+}$, and Bi$^{3+}$ cations, and do not necessarily need to have the AMX$_3$ composition or crystal structure. Many of these PIMs have recently been explored for photovoltaic applications. These include BiI$_3$, (CH$_3$NH$_3$)$_2$BiI$_4$, BiSI, BiSeI, K$_2$BiI$_4$, Cs$_3$Bi$_2$I$_9$, BiOI, Cs$_2$Sb$_2$I$_9$, SbSI, SbSeI, SbSe$_2$, CuSbSe$_2$, (CH$_3$NH$_3$)$_2$KBiCl$_6$, SnS. Many other recent works have been written on understanding the properties of hybrid lead-halide perovskites that make them excellent solar absorbers, e.g., energetic carrier screening via polarons, which may lead to new design rules.

In many cases, these materials are being investigated as solar absorbers for the first time. Among the multitude of physical properties to be assessed, the most critical for rigorously evaluating the promise of each material for photovoltaics at an early stage are:

- Structure & composition: what is the crystallographic structure and ratio of elements in the synthesized thin film? Is this a single phase?
- Electronic structure: what is the electronic structure, particularly the conduction and valence band energies?
- Optoelectronic properties: what is the bandgap and its nature, what is the optical absorption coefficient, and what is the photoluminescence quantum yield?
- Transport properties: what are the minority carrier lifetime and mobility?
- Device-relevant properties: what is the photovoltaic efficiency potential?

However, a common issue in many of these systems is the extent to which the characterization techniques used to determine these physical properties can damage or significantly perturb the materials. These questions are particularly relevant to hybrid or halide-based materials, which can have formation enthalpies that are low in magnitude. The flux or vacuum used during measurement can cause chemical changes in the material, and we discuss this extensively in the text. Traditional characterization techniques used to measure these physical properties include X-ray diffraction for determining structure and X-ray photoemission spectroscopy under ultrahigh vacuum for measuring composition. Photoemission spectroscopy (PES) is used to elucidate the electronic structure. UV-visible
spectrophotometry and ellipsometry can be used to determine the bandgap and absorption coefficient. The minority carrier lifetime and mobility can be measured through time resolved photoluminescence,\cite{44} terahertz spectroscopy,\cite{36,45,46} and photoconductivity decay measurements.\cite{47}

Challenges also exist in performing quantum mechanical calculations on some of these materials. These calculations are used to predict the thermodynamically stable crystal structure for a given composition,\cite{21} and are also often used to predict the electronic structure.\cite{27,28,30,31} As suggested by recent literature, these calculations have become a critical component in the search for novel photovoltaic materials.\cite{17,20,21,25,28,30,31,48,49} However, the layered structure and presence of partially-oxidized heavy metal cations in many of these materials necessitate special care when performing these calculations.

The time required to fully understand and overcome all of these challenges has, in some cases, been incongruous with the rapid pace of the perovskite and perovskite-inspired materials fields.\cite{24} The risks of inaccuracies on the progress and standing of various research fields has been the subject of regular debate.\cite{2} To date, much of the attention for streamlining the research efforts in the field of hybrid perovskite photovoltaics has focused on device efficiencies, leading to a checklist being established on best practices in photovoltaic device characterization (refer to Ref. 50).

A similar reflection and debate among researchers on the pitfalls and best practices in materials characterization and calculations is needed. Such a discussion was held at the 2015 Materials Research Society Fall Meeting & Exhibit by some of the researchers working on hybrid perovskites and PIMs. Here we discuss the important insights gained from debating collective experience in undertaking characterization and computational efforts. Many of the learnings from these discussions are communicated here, both with reference to recent literature addressing specific components and techniques, as well as controlled experiments performed by the authors to specifically highlight the challenges. In this article, we focus on three important areas of materials characterization for new photovoltaic materials: crystallography, composition and photoemission spectroscopy. We also focus on first principles calculations of new materials. This article concludes with our suggestions on protocols for avoiding artifacts and minimum reporting for each technique.

**EXPERIMENTAL SECTION**

**Materials.** The MAPbI$_3$ thin films for X-ray diffraction analysis were prepared from a solution of methylammonium iodide (Dyesol) and PbCl$_2$ (Sigma-Aldrich, 99,999%) in N,N-dimethylformamide (Sigma-Aldrich, 29%) using a previously reported method.\cite{50,51} The solutions were spun on clean soda lime glass substrates at 2000 rpm for 60 s followed by 6000 rpm for 10 s in a N$_2$ glovebox. The spin cast films were annealed under helium for 30 min at 100 °C.

The thin films synthesized for composition analysis were deposited on quartz (Quartz Scientific, Inc.), which was ultrasonically cleaned in acetone and isopropanol for 15 min each sequentially.

Bi$_x$ thin films were grown by physical vapor transport in a two-zone horizontal tube furnace according to a previous report.\cite{20} Prior to deposition, the quartz substrate was UV-ozone cleaned for 15 min. Bi$_x$ powder (Alfa Aesar, 99.998% ultra dry) was loaded into the source zone of the horizontal furnace and the quartz substrate to the substrate zone. The furnace was evacuated to a base pressure of 24 mTorr. N$_2$ gas was flowed through the furnace at 10 mL min$^{-1}$ such that the operating pressure was 5 Torr. The source temperature was 250 °C and substrate zone temperature 150 °C, and the growth time 90 min.

Methylammonium bismuth iodide (MBI) thin films were grown by solution-assisted conversion using a previously reported method.\cite{2} The substrate was O$_2$ plasma cleaned for 10 min. Bi$_x$ was deposited from a solution of 400 mg mL$^{-1}$ Bi$_x$ (Alfa Aesar, 99.999%) in N,N-dimethylformamide (Sigma-Aldrich, 99.8%). The solution was spun at 3000 rpm for 5 s, followed by 6000 rpm for 5 s inside a N$_2$ glovebox. The film was dried for 30 min at room temperature in a purging glovebox, following by annealing at 100 °C for 30 min on a hotplate.

PbI$_2$ was deposited on quartz that was O$_2$ plasma cleaned for 10 min. 450 mg mL$^{-1}$ PbI$_2$ (Sigma-Aldrich, 99%) was dissolved in N,N-dimethylformamide at 80 °C under stirring. The substrate was heated at 80 °C before 25 μL of the solution was dispensed onto the substrate and spun at 6000 rpm for 35 s. The film was annealed at 100 °C for 50 min.

MAPbI$_3$ for composition measurements was deposited on quartz O$_2$ plasma cleaned for 10 min. Lead acetate trihydrate (Sigma-Aldrich, 99.999%) was mixed with methylammonium iodide (Dyesol) in a 13:03 molar ratio inside a N$_2$ glovebox and mixed for 15 min at 50 °C. The solution was filtered with a 0.2 μm PTFE filter. The MAPbI$_3$ films were spin-cast inside a box filled with air, in which the relative humidity was <10%. The perovskite solution was heated at 50 °C and substrate at 85 °C for 15 min. The solution was dynamically spun at 2000 rpm for
Characterization methods. X-ray diffraction (XRD) measurements were made at SSRRL beamline 2-1 with a photon energy of 10.995 keV calibrated using a sintered plate of Al\textsubscript{2}O\textsubscript{3} (NIST SRM 1976b). The samples were mounted inside a chamber filled either with helium or ambient air at 25 °C with the beam passing through X-ray transparent kapton windows. Diffracted X-rays were detected using a Pilatus 100k detector mounted 600 mm from the sample on the 2θ arm of the diffractometer. Two-dimensional scattering was collected at SSRRL at BL1-3 with a MAR345 image plate detector and X-ray energy of 12.7 keV. The grazing incidence angle was 2°. Images were calibrated using a LaB\textsubscript{6} standard and integrated using GSAS-II. Diffraction was collected with a 60 s exposure time in a chamber filled with helium at 25 °C.

In the electronic structure section, hybrid functional calculations were performed using the PAW implementation\textsuperscript{39} of the HSE functional\textsuperscript{39} in VASP. Spin-orbit interactions were included as first-order perturbation (i.e., not self-consistently).\textsuperscript{33} A 6\times6\times6 sampling of the Brillouin zone was used together with the plane wave cutoff of 300 eV. The GW calculations\textsuperscript{30} were performed using a fully relativistic, two-component formalism\textsuperscript{60} with the Quantum ESPRESSO\textsuperscript{62} package for the mean-field PBE calculations and the BerkeleyGW package\textsuperscript{63} for the actual GW calculations. Pseudopotentials were generated using the ONCVPSP package.\textsuperscript{64} Specifically, the iodine pseudopotential was generated by taking the input file from the SG15 pseudopotential library,\textsuperscript{65} but using the fully relativistic version of the ONCVPSP code. This is necessary because the SG15 library contains scalar relativistic, not fully relativistic pseudopotentials. The Bi pseudopotential is an example provided with the ONCVPSP code. The wavefunction cutoff, screened Coulomb cutoff, and number of bands in the perturbation theory sums were 60 rydberg, 15 rydberg, and 3600, respectively. A 4\times4\times4 sampling of the Brillouin zone was used. To obtain the full band structure, the Godby-Needs plasmon pole model\textsuperscript{66,67} was used, but results were checked against full frequency calculations and shown to give a 50 meV difference in the gap. The accuracy of our calculations is 0.1–0.2 eV.
MATERIALS CHARACTERIZATION

Crystallography. This section discusses the best practices in characterizing the atomic structure of lead-halide perovskite and PIMs using crystallographic methods, as well as the potential pitfalls that can be encountered. We primarily focus on X-ray diffraction (XRD) because this is more commonly available, but also make reference to the opportunities available with neutron diffraction. Introductions to these techniques are readily found in a number of texts. While we focus much of the following discussion on MAPbI₃, the proposed best practices to avoid potential artifacts are relevant to PIMs in general.

Beam damage

We begin by discussing the importance of considering the radiolysis of photovoltaic materials, and show the primary factors that can influence the degree to which beam damage affects the quality of the measurements. We then describe best practices in crystal structure determination in single crystals, powders and thin films. It is particularly important that these are performed correctly by the community investigating PIMs, since accurate structural information is essential for accurate density functional theory calculations and ensuring that the material synthesized is of the desired phase. We also describe the opportunities offered by new in-situ and operando XRD measurements of thin films, and the specific best practices that should be adopted.

Before discussing “beam damage”, it is useful to recall the definition of flux – the number of photons per second per unit area. The flux of X-ray sources can vary by many orders of magnitude and is highly dependent on the specific source, as well as the instrument/beamline optics. However, in general, a synchrotron source will provide between one to more than three orders of magnitude greater flux than conventional laboratory X-ray sources. Beam damage, or, more precisely, radiation-induced structural damage, is one negative effect associated with the large flux generated at synchrotron sources. For this reason it is important to consider what effects the radiation could be having on the system being used. Beam damage is often manifested by a decrease in the scattering intensity of Bragg peaks as was shown by Polvino et al. in a study of thin film capping layers on silicon-on-insulator (SOI) substrates. One must carefully consider what the radiation effects are on their system so one can properly interpret the data. In contrast, for neutron diffraction, the damage caused by neutron beams is negligible, because of weak neutron interactions, and can be safely ignored.
Figure 1. Beam damage study of MAPbI$_3$ films at room temperature taken by the authors. (a) Image of MAPbI$_3$ films after approximately 24 h of X-ray exposure in air (top) and in helium (bottom). (b) Full powder XRD patterns both before and after 1 h of X-ray exposure in air. The difference between the two patterns is plotted below, showing significant differences due to beam damage. The fitted intensities and positions of the (0, 0, 2) and (1, 1, 0) peaks are shown as a function of X-ray exposure time in panels (c-f), with error bars plotted at representative points. The error bars for the (1, 1, 0) peak intensity are multiplied by a factor of 10 for clarity. We note that the (0, 0, 2) and (1, 1, 0) peak positions are at different 2θ values to lab-standard diffraction patterns because these measurements were taken using a 10.995 keV X-ray beam, rather than a 8.049 keV Cu K$_\alpha$ X-ray beam. The raw data is available as Supporting Information.

Similar to synchrotron measurements on other halide materials, and biological materials, significant beam damage can occur when working with MAPbI$_3$ and other halide-based PIMs. It is therefore important to consider optimal conditions for measuring these new absorbers. Here, we study MAPbI$_3$ thin films to understand the factors that may influence beam damage. In Figure 1, we show a summary of a beam damage study performed on thin films of MAPbI$_3$. One film was cut in half and dosed with X-rays in both air (32% relative humidity) and helium environments at approximately 23°C. Details of the measurement can be found in the experimental section. In summary, the incident X-ray energy from a synchrotron source was 10.995 keV, and the flux was 4.9×10$^{10}$ photon s$^{-1}$mm$^{-2}$. After approximately 24 h of exposure, the originally dark films visually showed damage. The film exposed to synchrotron radiation in air was nearly completely ablated (Figure 1a). This beam damage exceeds that reported for methylammonium lead iodide measured in air with lower-flux Cu K$_\alpha$ X-rays in <1 h, which again emphasizes that the flux of the X-ray source is a critical factor for beam damage. Closer inspection of our synchrotron-based XRD data shows that the atmosphere also greatly influences the magnitude of
the beam damage seen in these samples. After approximately 1 h of exposure, both samples show a decrease in the scattering intensity, as well as a shift in the peak position, although PbI$_2$ peaks do not appear (Figure 1). However, it is clear the sample measured in air is much more prone to beam damage, which is consistent with the susceptibility of MAPbI$_3$ to chemical degradation in the presence of humid air.\textsuperscript{50} Another factor to consider is the energy of the incident X-rays. X-ray absorption is element and energy specific. For example Pb has an absorption edge near 13.03 keV, meaning that in the case of MAPbI$_3$, the Pb in the sample will be absorbing a much larger amount of X-rays above 13 keV compared to 11 keV. This stronger absorption can result in a higher dose, and hence a higher rate of beam damage.

It is important to understand what measures can be taken to minimize the effects of X-ray beam damage, as well as how to properly report these effects. To minimize damage it is best to consider the flux and atmosphere of the measurement (e.g., He vs. air) as well as the X-ray energy. If possible, avoiding an absorption edge can help reduce the rate of damage. It is important to clearly report all of the conditions during the measurement. This includes: X-ray energy, flux, atmosphere, temperature, and relative humidity if run in air.

Crystal structure determination

It is imperative that the crystal structure of a new photovoltaic material is correctly solved. This is because an understanding of the electronic structure, transport properties, mechanical properties and defect chemistry relies on knowing the arrangement of atoms in the material.\textsuperscript{1,7,13,24,30,79,81,82} While crystal structure determination is standard in crystallography, it is important to highlight the critical considerations for evaluating the quality and accuracy of the solved structure.

Analysis of single crystals

Single crystal X-ray diffraction has become a powerful and fairly routine method for determining the structures of new materials. Advances in crystal growth coupled with modern synchrotron sources and analysis tools have made it possible to determine the crystal structure on smaller crystals and on increasingly complex materials. This is a good method to determine the crystal structure of new materials and has been adopted in many recent works on PIMs.\textsuperscript{13,24,83} There are, however, some pitfalls and challenges in determining crystallographic structure from single-crystal diffraction data. Among many others, the most common difficulties are: 1) crystals are often twinned, leading to the projection of two or more diffraction domains onto one data set. This is especially difficult for MAPbI$_3$ and similar perovskite materials due to their relatively fast solution-phase crystallization as well as their nearly-cubic phase, which obfuscates the three crystallographic directions to facilitate twin boundary formation in "single" crystals; 2) the combination of both very heavy and very light elements in a single structure. X-ray diffraction intensity increases with the atomic number, which makes determining the positions of heavier elements (e.g., I and Pb) significantly easier than light elements (e.g., C and N). Additionally, this can obfuscate the C and N atoms of the methylammonium cation, which have similar atomic numbers and have traditionally been difficult to resolve.

The elements may be disordered over multiple positions in the lattice. This is affected by the temperature and phase, and is difficult to model, especially for the lightly-diffracting methylammonium cations. In addition to observing general single-crystal structure determination guidelines (such as those reported in Ref. 84 and 85), it is imperative that researchers are aware of the above potential difficulties when performing crystallographic studies of hybrid organic-inorganic single crystals. Relevant parameters for data collection and structure analysis should be reported along with the article. For example, authors should provide in the experimental details or in the supporting information: 1) details about the crystal, such as its growth conditions, dimensions, and appearance; 2) data collection parameters including sample, X-ray energy, and instrument or beamline equipment; 3) details about structure determination, including the software package, solution and refinement methods, and how disorder and crystal twinning were modelled. A crystallographic information file (*.cif) should be generated and checked for quality using checkCIF.\textsuperscript{86} In general, high-level alerts (i.e., all A- and B-level alerts) should be avoided entirely and lower-level alerts should be addressed if possible. Finally, the *.cif file should be distributed with the report as a supplementary file. It is important that studies that utilize newly solved crystal structures should report the crystallographic data as described above. Waiting for a future publication to reveal those details (or not reporting them at all) should be avoided. Some recent examples of good reporting of single crystal X-ray diffraction are given in Ref. 11, 87 and 88.

Analysis of powders

Powder methods are another option for elucidating the crystal structure of new materials if it is difficult to grow a single crystal. For photovoltaic materials, the sample form factor is generally a polycrystalline thin film, which can be readily scraped off as a powder sample. Powder diffraction is therefore the preferred method to study device-relevant form factors. Some groups also prefer to crush single crystal samples to powder for simpler analysis.\textsuperscript{89,90} While structure determination from powder diffraction is increasingly common, and has been demonstrated to give comparable accuracy to single
crystal methods, its application is significantly complicated by the collapse of three dimensions of reciprocal space into a single dimension and subsequent loss of information due to peak overlap.\textsuperscript{6} Since powder X-ray diffraction data is commonly used to report new structures, and is not always performed or reported in a rigorous way, it is important to consider best practices in this technique.

Prior to any analysis, it is critical to ensure that sufficiently high quality data is collected. No amount of careful analysis will make up for poor quality data, and may often lead to erroneous results from over interpretation. What constitutes sufficient data quality will always be determined by the question being asked of the data, however a few general guidelines are useful. The point density used for a measurement is important for proper sampling of a diffraction feature. This will depend on the sharpness of the diffraction peaks, but should be such that there are approximately 6 points over the sharpest feature in the data; this would translate to having 8–12 data points over the full width at half maximum for a diffraction peak. Sparser point densities will yield inaccurate peak positions and intensities while higher density will simply result in increased data collection times. For good structural refinement, diffraction data should be collected to high resolution, typically meaning a d-spacing close to one Angstrom or $Q \approx 6.3 \, \text{Å}^{-1}$, where $Q = \frac{4\pi \sin(\theta)}{\lambda}$ or $Q = \frac{2\pi}{d}$, where $\lambda$ is X-ray wavelength and $d$ is the d-spacing. Not all samples will give measureable diffraction at these high $Q$-values, and for simple determinations such as phase confirmation, much more limited data may be sufficient.

In many powder samples, preferred orientation can be quite strong, giving rise to large changes in the relative peak intensities in a diffraction pattern. While there are certainly models that can account for this in Rietveld refinement, it is always preferable to avoid the problem at the sample preparation and data collection stage. Forming loose powders from scraped films and loading into a capillary or side-loading into a flat plate holder will typically reduce or eliminate preferred orientation. Mixing loose powders with another component, either an internal crystalline standard or ground glass, can further reduce preferred orientation effects. Should preferred orientation be impossible to eliminate, either the March-Dollase model or spherical harmonics can be incorporated into most Rietveld refinement software to correct for this effect.\textsuperscript{92,93} However, it is imperative to ensure that the texture model used is physically meaningful in addition to giving a better fit to the data. This is especially true in the case of spherical harmonics where non-positive definite results are possible if the parameters used are not properly constrained.

Once a material has been structurally characterized and reported in the literature, it becomes common practice to use powder XRD as a means of confirming the phase purity of a sample. This typically follows a fingerprinting approach, where characteristic peaks are identified with a crystalline phase and identified in the diffraction pattern. This approach ignores a significant fraction of the diffraction pattern and runs the risk of missing phases whose peaks may closely overlap those of other phases. Assuming proper sample preparation and data collection, analysis should be performed by fitting the entire diffraction pattern. This could be a LeBail or Pawley fit if lattice parameters are of interest, or Rietveld refinement if structural information is also desired.\textsuperscript{94,95} Rietveld refinement is the preferred approach to confirm phase purity, as it considers peak intensities as well as locations, which will prove impossible to fit well should there exist multiple phases with overlapped peaks. The Rietveld approach will also naturally provide a quantitative measure of the various phases present in a sample in the case of secondary phases. Care must be taken not to overfit the data, and to gradually add parameters to the refinement to avoid settling into local minima. While refinement routines are designed to minimize a figure of merit, generally either $\chi^2$ or a weighted $R$ factor, these values are not very reliable for determining a good fit. These values are highly impacted by disagreements in areas of the XRD pattern that may not affect the structural results, such as a poorly matched peak shape or background, or even simply poorly determined uncertainties in the measured data. Rather, the fit should be judged by the difference curve between the measured and calculated patterns, the so-called ‘chi-by-eye’ approach. Any significant features in the difference curve should be cause for concern and must be accounted for to ensure a proper fit. This illustrates the importance of reporting not only the structural results of a refinement, but the data and calculated diffraction patterns as well. It should be strongly encouraged to include all measured data as well as all details of a refinement with any published result so that others may judge for themselves the quality of a fit.\textsuperscript{96}

Analysis of thin films

When collecting XRD data on thin films, an important consideration is the orientation of crystallites in the sample. Thin film samples often show preferential orientation of domains in the film, which is often referred to as texture. When a sample is textured the diffracted signal is no longer isotropic. As a result, one-dimensional or line scan data can be misleading (Figure 2). Two-dimensional XRD of a MAPbI$_3$ thin film is shown in Figure 2a. The red line represents the data that would be collected from a standard out-of-plane line scan and the yellow outlined “cake-slice” shows the full pattern integration that can be obtained by two-dimensional area diffraction. When these integrated patterns are
compared it is clear that the line scan does not accurately capture the relative peak intensities of the sample (Figure 2b, green box). In extreme cases peaks may be completely missed in the line scan (Figure 2b, purple box). If a two-dimensional detector is not readily available one can rotate the sample with respect to the X-ray beam and take a series of line scans. If the peak intensity fluctuates between scans at different angles the sample is most likely textured and care should be taken in interpreting the data.

XRD measurements are also often used to determine the crystallite size of thin films. The expression for determining crystallite size from XRD peak broadening was derived by Scherrer almost a century ago,97 and many excellent papers have been published describing its successes and limitations since that time.98-100 Nevertheless, the crystallite size analysis using Scherrer’s approach remains one of the most misused techniques in materials science. The most common problem is that many intrinsic and extrinsic factors, other than crystallite size, can contribute to XRD peak broadening. The extrinsic instrument artifacts are often related to the X-ray source (e.g., convolution of Cu Kα and Cu Kβ lines) and X-ray detector (e.g., resolution of 2D detectors). These artifacts can be accounted for by the proper use of instrument calibration standards (e.g., LaB6 powder from NIST) as a minimum best measurement practice. The range of possible intrinsic materials phenomena that can contribute to peak broadening is even larger, including structural defects (e.g., stacking faults, dislocations, twins), microscopic stress, and chemical heterogeneities of solid solutions. Solid solution broadening is particularly important to the Scherrer analysis of hybrid organic-inorganic perovskites, where alloys are often used to tune the optoelectronic properties.100 Most publications conclude that the crystallite size of MAPbI3 is on the order of 25-75 nm.101-104 But these estimates may be affected by broadening factors (e.g., microstrain). Other studies arrived at a crystallite size on the order of 1.5 μm, which is questionable, given that the resolution of common lab-based XRD instruments is not capable of measuring diffraction peaks this sharp.105 It may be tempting to use electron microscopy to validate the crystal sizes obtained from XRD analysis, and this is indeed a common method for measuring the size of grains in the literature.106-108 However, the microstructures observed may be composed of much smaller crystallites, which would make such a comparison invalid. For example, atomic force microscopy and scanning electron microscopy measurements of methylammonium bismuth iodide have shown features that appear to be grains with sizes ranging from 0.5 to 1 μm. Transmission electron microscopy measurements, however, revealed the crystallite size to range from 20 to 30 nm.21 Some works have presented a deeper analysis of the crystallite size of MAPbI3 using Scherrer’s method, combining XRD with pair distribution function measurements.106

**Further considerations in diffraction measurements**

Neutron diffraction can be used as an important complement to XRD and has the distinct advantage that neutrons are more sensitive to light elements than are X-rays. Hence, neutron diffraction can resolve the position of the molecular ions in perovskite crystals, e.g., methylammonium cations in MAPbI3. Although neutron diffraction cannot be used on thin films (because of insufficient sample volume), neutron diffraction of MAPbI3 single crystals and powders provide important insight into the material’s structure and dynamics.107,108 Text and reviews109,110 can be referred to for best practices in neutron diffraction.

Another important consideration when performing crystallography on lead-halide perovskites is the dynamic motion of the organic cation, which has been observed in the lead-halide perovskites. These may also be present in other hybrid organic-inorganic PIMs.21 Quarti et al. have identified fluctuations of atoms (particularly those in the methylammonium cation and iodine in the case of MAPbI3) on the picosecond time scale.112 However, XRD is not sensitive to these rapid structural fluctuations. Due to the nature of XRD, the data only describes spatially and/or temporally averaged atomic positions and it is not possible to decouple these two effects. Therefore, it is important to understand that dynamic disorder and coupled dynamical disorder is not characterized by XRD.
With the realization of brighter X-ray sources and faster detectors, in-situ and operando XRD measurements are becoming more common. In the context of best practices for these techniques, all of the previous discussion applies. The importance of understanding the influence of beam damage and taking care during structure determination is amplified by the longer times and larger data sets. Whenever possible, running a standard sample to determine beam damage effects is recommended, as this will enable a greater understanding of the sample being measured. For example if the goal is to determine the influence of light soaking a sample on the structure over many hours, a sample should be run in the dark with the same X-ray dose and the light-soaked sample. When care is taken, it is possible to learn a great deal of information from in-situ and operando measurements.

**Composition.** In the hybrid perovskite literature, it has become common to report the ratio of the precursors in solution rather than characterizing the composition of the synthesized material. This protocol is insufficient, particularly as new perovskites and PIMs with more complex compositions are explored. Composition measurements are also needed to determine whether the distribution of elements (e.g., cesium, carbon, nitrogen and halides in cesium-formamidinium lead iodide bromide perovskites) in the material is uniform and consistent with a single phase. Some of the most common composition measurement techniques are listed in Table 1. Using these techniques, the composition of polycrystalline thin films, single crystals or precursor solutions could be determined. The operation of these is summarized, but readers are encouraged to refer to the references listed for each technique for details. Designing the experiment to accurately measure composition requires knowledge of the interaction volume and resolution of the techniques available, as well as their common artifacts. The surface/bulk sensitivity of the techniques listed in Table 1 is summarized in Figure 3. Secondary ion mass spectroscopy (SIMS), atom probe tomography (APT) and electron energy loss spectroscopy (EELS) can be considered as both surface and bulk sensitive. In SIMS, ionized particles sputtered from the surface are characterized, but the sputtering process can proceed to the bulk of the material. APT and EELS involve characterization of a nano-sized specimen taken from the bulk or close to the surface of a sample. Also, X-ray fluorescence (XRF) typically has an analysis depth on the order of single to hundreds of microns, depending on the atomic number of the constituent elements, but this can be reduced by lowering the incident angle of the X-ray beam to the sample.

Two common sources of artifacts that are of particular importance for lead-halide perovskites and PIMs are beam damage and degradation under vacuum. Most of the methods in Table 1 involve the measurement of particles resulting from the irradiation of the sample with a probe beam of X-ray photons (X-ray photoemission spectroscopy or XPS), electrons (Auger electron spectroscopy, XRF, energy dispersive X-ray spectroscopy, electron energy loss spectroscopy) or ions (Rutherford backscattering spectroscopy or RBS). During these processes, beam damage can occur, in which the composition, oxidation state or bonding environment are changed during the measurement.

![Figure 3. Common techniques for measuring composition, classified based on whether they are surface sensitive, measure the bulk of the film or either. Abbreviations are defined in Table 1.](image-url)

A prominent example of beam damage is RBS measurements of methylammonium bismuth iodide (MBI) and MAPbI$_3$. In both cases, it was found that exposure to a high-energy ion beam resulted in a reduction in the nitrogen and iodine content, likely due to the removal of the volatile methylammonium iodide component. The sensitivity of the material to the probe beam under measurement conditions should be investigated prior to characterizing the composition. An example of this is shown in Figure 4, where we measure two inorganic (BiI$_3$ and PbI$_2$) and two hybrid materials ((CH$_3$NH$_3$)$_2$BiI$_5$ and CH$_3$NH$_3$PbI$_3$) by XRF repeatedly over a three minute period. By doing so, we show that the XRF peak intensity at two fluorescence energies remain constant, indicating that the X-ray probe beam we used does not change the composition of these materials (details in Experimental section). Additionally, XRF may be a suitable technique to assist in detecting changes in composition resulting from radiolysis, such as during RBS or XPS measurements. Raman spectroscopy has also been put forward as a non-destructive method to measure the ratios of different elements. Other methods to detect beam damage exist, such as analyzing the low-loss region in EELS measurements. A variety of methods are available to minimize beam damage in the techniques listed in Table 1. These
include controlling the acceleration voltage of the probe beam, reducing the flux of the beam by defocusing, reducing beam exposure time through rastering, using different particles for the probe beam, adjusting specimen orientation and thickness, or adjusting the wavelength of the probe beam. In all cases, evidence should be given that the composition measurements are reflective of the original sample.

Figure 4. X-ray fluorescence (XRF) measurements of BiI₃, (CH₃NH₃)₃Bi₂I₉, PbI₂ and CH₃NH₃PbI₃ thin films on quartz (QSI Scientific, Inc.). (a, c, e, g) XRF spectra and (b, d, f, h) repeated measurements of XRF intensity at selected binding energies over a 3 min period. Rh is from the X-ray source target and therefore appears in the measurement. These measurements were taken by the authors.
<table>
<thead>
<tr>
<th>Technique</th>
<th>How measurement is conducted</th>
<th>Resolution/detection limits</th>
<th>Common artifacts</th>
</tr>
</thead>
</table>
| X-ray photoemission spectroscopy (XPS)        | X-ray photons (typically 1.25 keV for Mg Kα or 1.49 keV for Al Kα) incident on sample remove bound electrons. Kinetic energy of photoemissions measured and binding energy can be calculated. Binding energy is specific to the element and binding environment, enabling element identification and chemical analysis. Quantification through the areas of the fitted peaks of XPS core level spectra, with consideration given to instrument- and element-specific sensitivity factors. | - Energy resolution: ≤0.3 eV (monochromated Al Kα X-rays, commercial system)  
- Energy resolution: a few meV (synchrotron radiation)  
- Detection limit: 0.05 at.%  
- H and He not detectable | - Damage and decomposition of soft and metastable materials  
- Preferential sputtering during depth-profiling can remove more volatile components and change composition  
- Beam-induced changes in oxidation state  
- Binding energy shift due to charge accumulation, as well as junction and surface photovoltage  
- Gas-phase surface contaminants can occur |
| Auger electron spectroscopy (AES)             | Sample irradiated with 3–20 keV electrons, ejecting core-level electrons in atoms and creating vacancies. Relaxation of outer electron to vacancy results in emission of Auger electron. Kinetic energies of Auger electrons are characteristic of the elements present and their bonding environments. Quantification through spectra fitting. | - Relative energy resolution: 0.05% (SCA/CHA analyzer)  
- Lateral resolution: <10 nm  
- Detection limit: 0.1 at.%  
- H and He not detectable | - More beam damage than XPS, and can result in the creation of defects, changes in crystal structure and changes in oxidation state  
- Gas-phase surface contaminants can occur |
| X-ray fluorescence (XRF)                      | Sample irradiated with high-energy X-rays or gamma rays, ejecting electrons and creating vacancies. Outer electrons relax to vacancies, emitting X-rays that are measured. X-rays either collected simultaneously (energy-dispersive) or separated through diffraction (wavelength-dispersive). Quantification of XRF spectrum through calibration with standards or models. | - Binding energy resolution (energy-dispersive): 150–600 eV  
- Binding energy resolution (wavelength-dispersive): 5–20 eV  
- Detection limit: ≤0.01 at.% but element specific  
- Spot size: 100 μm to 1 mm, but can be reduced in μXRF  | - Accurate quantification can be limited by self-absorption effects in films with high concentrations of the element analyzed  
- Inaccurate quantification can occur if the sample is insufficiently thick to attenuate the primary X-ray beam or if the sample is not homogeneous  
- Typically limited to heavier elements for standard Be detector windows: Z ≥ u (Na)  
- Escape peaks can be generated |
| Energy dispersive X-ray spectroscopy (EDX or EDS) | Electron beam incident on sample generates X-rays. These X-rays are detected, e.g., by Si(Li) crystal. The voltage generated by the detector is proportional to the X-ray photon energy. A | - Energy resolution: 80–180 eV (SEM)  
- 100 eV (TEM)  
- Detection limit: 0.1 wt.% | - Spectral artifacts can occur due to the detector, detector window, pulse-processing or microscope environment (e.g., icing in the chamber)  
- EDX spectral peaks can be |
<table>
<thead>
<tr>
<th>Method</th>
<th>Details</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rutherford backscattering spectroscopy (RBS)</strong></td>
<td>Sample bombarded with 0.5–4 MeV ions, which are typically protons, He(^+) ions, Li(^+) ions or N(^+) ions. The energy and number of elastically backscattered particles are measured. Composition and quantification obtained by fitting the RBS spectrum with simulations based on target structure.</td>
<td>- Energy resolution: 10 keV (protons)(^{142}) 12 keV (He(^+) ions)(^{142}) 20 keV (Li(^+) ions)(^{142}) 80 keV (N(^+) ions)(^{43})  - Sensitivity lower for low atomic mass elements. N(^+)-RBS can measure (10^{10}) atom cm(^{-2}) for Au and (10^{12}) atom cm(^{-2}) for As, Cr or Fe.(^{143})  - Inaccurate models for stopping power of compounds leads to inaccurate quantification(^{42})  - Volatile components can be removed by high energy ions, which changes the composition during the measurement(^{21,143})  - Sample can also become amorphous due to exposure to ion beam.</td>
</tr>
<tr>
<td><strong>Inductively coupled plasma mass spectroscopy (ICP-MS)</strong></td>
<td>Thin film samples are usually dissolved. Solution is nebulized to create an aerosol, which is mixed with an inert gas. A plasma is generated and the ionized particles are detected by a mass spectrometer. The mass spectrometer distinguishes between particles by the mass to charge ratio. The concentration of each element is determined from the mass spectrum. Accurate quantification requires calibration with known concentration element standards.</td>
<td>- Resolution: 0.7–1.0 amu (for quadrupole mass spectrometer)(^{146})  - Detection limit: nanogram per liter to microgram per liter, depending on the element(^{144,145})  - Some species cannot be resolved because they have the same mass to charge ratio.(^{144}) This could be overcome by using inductively coupled optical emission spectroscopy (ICP-OES).(^{147,148})  - Contaminants can be introduced if solvents with inadequate purity are used, e.g., common laboratory-grade solvents. Contaminants can also be introduced from inadequately clean laboratory equipment used to prepare the solutions for analysis.(^{144}) High purity solvents, often with ppt level of metal contaminants, are required.(^{149})  - Detectors can be coated with the elements analyzed, which limits future measurements.</td>
</tr>
<tr>
<td><strong>Secondary ion mass spectroscopy (SIMS)</strong></td>
<td>Sample sputtered with primary energetic ions (0.25–30 keV). Charged atoms/molecules removed by sputtering guided to mass spectrometer. Modelling the mass spectrum enables quantification of the composition. Standards are usually required for full depth resolution.</td>
<td>- Depth resolution depends on the incident angle of primary ions. &lt;1 nm depth resolution has been demonstrated.(^{107})  - Lateral resolution: 5 nm(^{107})  - Different species with the same mass to charge ratio cannot be distinguished by the mass spectrometer(^{97})  - Contaminant species in the vacuum chamber will be analyzed by the mass spectrometer(^{97})  - Primary ions can chemically ...</td>
</tr>
<tr>
<td>Technique</td>
<td>Description</td>
<td>Examples</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Atom probe tomography (APT)     | Sample is milled to a needle shape. Surface atoms are ionized and removed from the apex of the needle by an electric field or a combination of pulsed-laser heating and electric field. An electric field accelerates the ions to the detector in a mass spectrometer. The x-y and z positions of the elements in the original sample are determined. | - Detection limit depends on the analytical volume. It has been demonstrated that <1 ppb K in Si can be measured by SIMS.  
- Artifacts can occur in depth profiles if the surface is rough. Similarly, measurement of buried interfaces can be inaccurate because a crater needs to be sputtered through the overlayers, which can be rough due to repeated sputtering.  
- Reconstruction artifacts can occur due to the quantum nature of the atom, field penetration at the apex of the needle, an aspherical apex, changes in the evaporation order or different evaporation rates of different elements.  
- In heterogeneous structures, surface electric field variations can lead to local magnification effects and the trajectory of different elements overlapping.  
- Thermal artifacts can occur under high laser fluence.  
- Atom probe specimens have large surface area to volume ratios, making them sensitive to the processing environment. |}
| Electron energy loss spectroscopy (EELS) | STEM-EELS uses a 100–300 keV kinetic energy electron beam, which is inelastically scattered when interacting with electron clouds surrounding atoms. Energy distribution of transmitted electron beam measured using a magnetic prism to separate electrons by energy, and the energy spectrum is recorded. Concentration of elements estimated from the intensities of the energy loss spectrum and partial cross-section for core-loss scattering. | - Spatial resolution:  
  x-y plane: 0.2 nm  
  z direction: 0.05 nm  
- Single atom sensitivity  
- Energy resolution:  
  0.1-1 eV  
- Sub-Angstrom spatial resolution possible in STEM-EELS  
- STEM-EELS is more sensitive than STEM-EDX for elements up to Ne. For atomic numbers between 10 and 25, STEM-EELS has higher sensitivity if the L_{31} edge is analyzed.  
- Multiple scattering events can occur if the sample is thicker than the mean free path of electrons (typically 70 nm, but depends on the atomic number and energy range). This restricts elemental mapping.  
- Accuracy of quantification limited by approximations used (e.g., for the partial cross-section of core-loss scattering) and diffraction contrast artifacts due to different amounts of inelastic scattering by different elements.  
- Beam-energy-induced damage of organic and some inorganic materials can occur. |}

Many of the techniques listed in Table 1 are performed under ultrahigh vacuum (10^{-9} Torr) to avoid introducing gas-phase surface contaminants. However, extra caution must be exercised when working with materials con-
Photoemission spectroscopy (PES) for electronic structure determination. In the development of new photovoltaic materials, the design and optimization of absorbers and charge transport layers calls for a direct experimental assessment of electronic structure at functional interfaces. These include the valence and conduction band onsets with respect to the Fermi level, as well as the vacuum level position. Photoemission spectroscopy is a powerful tool to investigate the band structure of novel semiconductor materials and the position of the Fermi level in the bandgap. The techniques are applicable not only to thin films but also single crystals, for which band positions need to be precisely known for electrical characterization and device applications. XPS, for instance, is not only useful for measuring the chemical composition and composition of materials by evaluating the core level area intensities, it can also be employed to determine the valence band density of states with respect to the Fermi level. More precisely, the position of the valence band maximum with respect to the Fermi level can be determined from the leading edge of the XPS spectrum. Ultraviolet photoemission spectroscopy (UPS) operates in an analogous manner, but uses lower energy photons in the spectral range of a few eV and 150 eV. This allows the electronic density of states in the outermost valence band region to be analyzed with higher accuracy and spectral resolution compared to XPS measurements.

Avoiding common artifacts in photoemission spectroscopy measurements

To ensure accuracy in these PES measurements, the experiments need to be performed in ultrahigh vacuum (10⁻¹⁰⁻¹⁰⁻⁹ Torr). This is primarily because of the high surface sensitivity, as explained above in the Composition section. The information depth is limited to the surface 1-10 nm due to the small mean free path of electrons in solids in the energy range of typical PES measurements, although the development of hard X-ray sources is overcoming this limit.

Proper calibration of the energy scale for both PES techniques is both important and prone to artifacts. This requires measuring a conductive sample with a known spectrum, such as the core levels of metals published by Seah in Ref. 167. It is particularly important to calibrate the binding energy scale to more than one core peak in a broad range of binding energies, because the analyzer electron optics exhibit an energy-dependent transmission function and non-linear scaling. Eventually, the position of the Fermi edge of a clean metal surface should be used to finalize the calibration. The determined position of the Fermi level serves as a reference point for the binding energy scale because the conductive sample is in direct electrical contact with the spectrometer. The measured kinetic energy of photoelectrons emitted \( E_{KE} \) can then be used to calculate the binding energy relative to the Fermi level \( E_{B} \) from Eq. 1, where \( h\nu \) is the energy of the incident photons.

\[
E_{KE} = h\nu - E_{BE}
\]

In a properly-calibrated instrument, the binding energies of core levels as well as the valence band onset of a semiconductor can then be measured relative to this Fermi level position. The work function of the sample surface, \( \phi \), and hence the vacuum level position can be determined from inelastically scattered secondary electrons with \( E_{BE} = h\nu - \phi \). The signal from these secondary electrons forms a cut-off in the photoemission spectra (secondary electron cut-off), because at higher binding energies, electrons have insufficient energy to overcome the work function of the sample. For both valence band and work function measurements, it is necessary to ensure that the metal surfaces used for calibration are clean of contaminants. This can be
achieved by repeated sputter cleaning and annealing cycles.

Another important consideration for the measurement of electronic energy levels using XPS/UPS is the effect of surface band bending resulting from junction photovoltage, surface photovoltage or trapping/de-trapping of surface defect states. These effects are common in the semiconductors used in photovoltaic applications. In particular, the excitation energies used for XPS/UPS measurements may produce a sufficient amount of free charge carriers to result in a surface photovoltaic effect and change the surface band bending. This may be pertinent to long-diffusion length materials, in which photogenerated charge carriers could easily diffuse and accumulate at surfaces if inadequately grounded. To account for this effect, measurements at higher temperatures, with/without illumination and with varying source intensity should be performed. Shifts in measured band positions may also occur due to surface charging, which particularly needs to be considered for materials with low background carrier concentrations (e.g., MAPbI$_3$ or BiI$_3$). UPS/XPS spectra taken on the same spot over time should be acquired to determine if spectral shifts occur due to charging and/or composition degradation as described later in this section.

**Fitting the leading edge in photoemission spectroscopy measurements**

Inaccuracies in the determination of energy levels by PES can also arise from the way the leading edge of the photoemission onset and the secondary electron cut-off are fitted. One of the most physically relevant and accurate methods currently available to obtain the band edge is by fitting the leading edge using the calculated electronic density of states from density functional theory. An early method was developed by Kraut et al. for Ge and GaAs. In this study, the valence band density of states was calculated, broadened by instrumental parameters and fit to the leading edge of the XPS spectrum to enable more precise location of the valence band edge.

A limitation of the Kraut method is the requirement for accurate density functional theory calculations. Currently, simpler methods are used to approximate the band positions. One of the most common approximations is to use a linear fit to the edge and the background of the spectrum. In this method, the intersection of the leading edge with the background level determines the position of the band offset. However, this approximation can lead to erroneous fits in the case of band edges that are not abrupt. In particular, inaccuracies may arise from this method if the density of states at the leading edge is low and there is no distinct tangent to take, as occurs for many PIMs reported. A method to visualize gap states is to plot the leading edge on a semilogarithmic scale, since band tails can be described by exponential functions. This is illustrated in Figure 6a for organic small molecules. By using a semilogarithmic plot, the gap states and valence band density of states may be distinguished. Similarly, an early investigation of hybrid organic-inorganic lead-halide perovskites showed that displaying the spectra in a semilogarithmic plot yields a better approximation of the valence maximum, as depicted in Figure 6b. It should be noted that the background fitted by this method can be highly susceptible to the measurement conditions and more accurate measurements are only possible if monochromatic UV sources are used. Best practice in this case is to evaluate and compare both the linear and semilogarithmic representations of the measured spectra.

![Figure 5](image-url)
The case of PbS quantum dots is a prominent example of the errors in fitting PES spectra to determine the band edges. A simple linear fit to the leading edge of the valence band density of states yields unrealistically high valence band maximum offsets exceeding the electronic bandgap of the material. This is due to a large tail in the density of states extending from the band edge to the bandgap and because the density of states at the valence band is below 1% of the maximum density of states. A method analogous to the original paper by Kraut et al. was therefore adopted. The density of states of PbS was calculated using GW calculations with spin-orbit coupling (Figure 7a), broadened with the instrument response and fit to the leading edge (Figure 7b). The procedure developed for PbS quantum dots is as follows:

1. From the $E$ vs. $k$ band diagram, the points in reciprocal space that would contribute to the valence band density of states were determined. For the case of PbS, these were the $L$ and $\Sigma$ points (Figure 7a).

2. The instrument broadening was approximated from the width of the Fermi edge of a clean metal surface.

3. The leading edge of the XPS/UPS spectrum was fit with a parabolic density of state model according to Eq. 2:

$$\text{DOS}_{\text{VB}} = (A(\text{DOS}_L + h\text{DOS}_\Sigma)) \otimes g$$

(2)

In Eq. 2, $\text{DOS}_{\text{VB}}$ is the calculated density of states for the leading edge of the valence band, $A$ is a scaling factor, $h$ is a scaling factor between the parabolic models to the density of states at the $L$ and $\Sigma$ points calculated using $k \cdot p$ theory, $\otimes$ is the convolution function, $g$ is a Gaussian line width with a measured FWHM to account for instrument broadening. As seen from Figure 7c, the parabolic model can be applied to approximate the density of states at the $L$ and $\Sigma$ points. $\text{DOS}_L$ and $\text{DOS}_\Sigma$ can therefore be calculated from Eq. 3.

$$\text{DOS} = 2(m^*_{\text{e}})^{3/2} (E - E_{\text{VBM}})^{3/2}$$

(3)

In Eq. 3, $m^*_{\text{e}}$ is the electron effective mass at the $L$ or $\Sigma$ point, $E$ the electron binding energy, $E_{\text{VBM}}$ the valence band maximum energy. In fitting $\text{DOS}_{\text{VB}}$ to the leading edge, $A$ and $E_{\text{VBM}}$ are the parameters that are varied.

Similar approaches are highly applicable to many of the PIMs discussed earlier in this section. For instance, a parabolic model has been employed to fit angle-resolved photoemission data of MAPbBr$_3$ single crystals. In particular, since many of the investigations into new PIMs involves heavy use of theory, fitting the leading edge of XPS/UPS measurements with the instrument-broadened density of states has realistic potential to become more widely adopted. The ambiguities in determining the valence band maximum can be further reduced by performing the measurements at various excitation energies and verifying that the band onsets are the same. In laboratory scale systems, various discharge lines of UV sources (e.g., He I and He II lines) can be used. Monochromated X-ray radiation can also be used, as can synchrotron radiation sources, which also allows for a large range of excitation energies.
Figure 7. (a) Electronic band structure and density of states of bulk PbS calculated with the GW approximation. (b) Comparing the GW-calculated density of states, fitted instrument-broadened valence band density of states and XPS spectrum of bulk PbS. (c) Comparison of the XPS measurement of 0.95 eV bandgap PbS quantum dots and the fitted parabolic density of states model. Reproduced with permission from Ref. 164. Copyright 2016 American Chemical Society.

Complementary techniques to photoemission spectroscopy
An effective way to confirm the measured band positions is to perform complementary analysis with additional techniques. The conduction band position can be estimated by measuring the valence band positions using the leading edge in UPS/XPS spectra and adding the optical bandgap measured using optical spectroscopy. It is important to keep in mind that at the high excitation energies typically employed in photoemission spectroscopy, the acquired spectral information is reasonably well described in a single particle picture. However, this is not true for most of the optical experiments where additional correction factors, such as the excitation binding energy, need to be considered. A direct way to measure the positions of the conduction band minimum is to use inverse photoemission spectroscopy (IPES), which samples the unoccupied states. As shown in Figure 8, the onset of the IPES spectrum can be fit to determine the bottom of the unoccupied states and therefore the conduction band onset relative to the Fermi level. As an example, the electronic single particle gap of MAPbI₃ has been measured as 1.7±0.1 eV, compared to an optical bandgap of 1.65 eV and 9-17 meV exciton binding energy matching the predictions from GW calculations with corrections for spin-orbit coupling (1.67 eV). An improvement to this evaluation can be achieved by carefully lining up the experimentally obtained spectra from PES and IPES measurements with simulated spectra from DFT calculations, as described above and demonstrated in Figure 8. A significant refinement of the band onset determination for a range of hybrid perovskite films was reported by Endres et al. Therein, the electrical bandgap for MAPbI₃ was determined to be 1.6±0.1 eV, in agreement with optical measurements, and demonstrating greater accuracy than previous electronic bandgap measurements that did not make use of DOS fitting of the IPES spectra. In the same way, the bandgaps and band offsets have been determined for MAPbBr₃ and CsPbBr₃ compounds.

It is important to note that the optical bandgap determined from optical characterization is prone to its own set of measurement and analysis errors. These could lead to inaccurate estimates of the conduction band minimum position. Tauc plots are commonly misused to determine the bandgap from the optical absorption spectra of MAPbI₃ samples that violate the multiple assumptions made in the original work on amorphous Si and Ge. These include diffuse light scattering on large sample grains, excitonic effects, sub-bandgap defect states and other effects that make the fitting of a tangent to the onset of absorption ambiguous. High quality optical measurements, or at least semi-logarithmic data visualization are important to achieve best practice for experimental bandgap determination.

Figure 8. UPS and IPES spectra of MAPbI₃ in (a) linear and (b) semilogarithmic plots. These spectra are compared to the density of states calculated by DFT. Reproduced with permission from Ref. 177.
Influence of substrate on band positions

Another important consideration when performing PES to determine band structure is that the band positions are influenced by the substrate the material analyzed is grown on. In materials with low background carrier concentration and few mid-gap defect states, as is the case for MAPbI, and is expected in some PIMs, the Fermi level is strongly influenced by the Fermi level position of the substrate. For example, it has been found that MAPbI seems n-type on TiO₂, but p-type on NiO₃ (Figure 9).163,178 Up to now, a clear assessment of the origin of this Fermi level shift within the perovskite film has been lacking. However, a similar magnitude in the Fermi level shift was observed for MAPbI films with thin MoO₃ layers deposited on top.254 For this particular case, the use of a MoO₃ interlayer between the perovskite and hole transport material drove a shift in the Fermi level of the underlying MAPbI film. These results have been corroborated by studies on MAPbI on a variety of substrate materials, which show the perovskite to be n-type on F-doped SnO₂, Al₂O₃, ZnO, TiO₂ and ZrO₂. On p-type substrates, such as PEDOT:PSS, NiO and CuO, the Fermi level of MAPbI shifts closer to mid-gap.79 These examples show the importance of characterizing the absorbers on the charge transport layers they would be paired with and to report the substrate with the PES measurements.

Sample degradation during photoemission spectroscopy measurements

Evaluating the degree of degradation during the measurements when characterizing these soft materials has been demonstrated to be crucial for a robust and accurate analysis of the properties described above. In a study by Emara et al., it has been pointed out that the ionization energy of MAPbI determined by PES varies across a broad range of literature and can be connected to the composition of the perovskite film.184 Prior to this study, it has been found that the core levels of the elemental components in MAPbI shift substantially with prolonged X-ray exposure during XPS measurements. This X-ray induced degradation can lead to a shift in the Fermi level by more than 0.7 eV, such that initially p-type MAPbI on NiO becomes n-type.178 X-ray photon induced degradation of MAPbI on TiO₂ under an ultrahigh vacuum of 2×10⁻¹⁰ Torr was recently studied in detail by Steirer et al.187 It was found that under a constant X-ray photon flux density of 1.5×10⁻¹⁰ photons cm⁻² s⁻¹, the methylammonium cation deprotonated and was lost, as shown in Figure 10. At the same time, the iodide content reduced exponentially over time, leading to the MAPbI converting back to PbI₂. When the I:Pb ratio was below 2.5, the valence band onset from the Fermi level reduced due to the appearance of the PbI₂ phase.187 This is particularly relevant to degradation studies of perovskites, in which prolonged exposure to the photon beam occurs.188 But even for studies involving photon exposure times of under 2 h, it is best practice to carefully keep track of transient changes in the XPS peaks to rule out degradation-related measurement artifacts. Sequential acquisition of UPS and XPS spectra can be used to further evaluate the influence of radiation exposure and beam damage from both techniques. This is particularly critical for many PIMs which exhibit soft bonds or have volatile components, as discussed earlier. The measurement conditions (e.g., photon flux and exposure time) should also be optimized to collect data with sufficient accuracy and minimal beam damage. The photon flux density, vacuum, exposure time, duration of the measurements, and spectra at the start and end of the measurements should all be reported.

Figure 9. Comparison of the Fermi level position in MAPbI thin films on different substrates. (a) UPS and IPES spectra used for band structure determination. Reproduced with permission from Ref. 178. (b) The energy level diagram of MAPbI adjacent to different charge transport layers.178
Crystal Structure.

Atomic structure model

Consistent first principles calculations rely on the underlying atomic structure of the material being accurately known, particularly the lattice parameters and atomic coordinates. To illustrate this, we calculated the bandgaps and effective masses at the DFT level (generalized gradient approximation with spin-orbit coupling, or GGA+SOC) using different structure models of BiI₃. Even though the absolute magnitude of the bandgap is generally underestimated by DFT, this comparison illustrates the degree of sensitivity to the structure model. Table 2 shows the results calculated using four different models. (1) GGA: the atomic structure was fully relaxed in GGA. (2) Scaled GGA: the volume was scaled to the experimental volume while maintaining the cell shape and positions from GGA. (3) optIB86b: the structure was fully relaxed using the optIB86b van der Waals functional. (4) ICSD: the experimentally reported structure was taken from the inorganic crystal structure database (ICSD, no.: 78791).²⁸⁸ It can be seen from Table 2 that there are significant variations in the predicted bandgaps and effective masses. Accurate computations cannot be expected without accurate crystal structure.

Table 2. Variation in the electronic properties of BiI₃-SG148 in DFT (GGA+SOC) for different crystal structure models.

<table>
<thead>
<tr>
<th>Structure model</th>
<th>$E_g$ (eV)</th>
<th>$E_g$,Direct (eV)</th>
<th>$m_e$ (m₀)</th>
<th>$m_h$ (m₀)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGA</td>
<td>1.47</td>
<td>1.56</td>
<td>1.9</td>
<td>8.1</td>
</tr>
<tr>
<td>Scaled GGA</td>
<td>1.70</td>
<td>1.83</td>
<td>1.5</td>
<td>6.6</td>
</tr>
<tr>
<td>optIB86b-vdW</td>
<td>1.24</td>
<td>1.29</td>
<td>1.5</td>
<td>5.2</td>
</tr>
<tr>
<td>ICSD</td>
<td>1.27</td>
<td>1.33</td>
<td>1.4</td>
<td>5.3</td>
</tr>
</tbody>
</table>

Beyond the need to base electronic structure calculations on precise lattice parameters of the ground state structure, finite temperature effects can be pronounced in PIMs. Large thermal expansion coefficients and significant dynamic motion have been measured in methylammonium lead iodide, approximately an order of magnitude larger than inorganic materials.²⁸⁹ The large thermal expansion coefficient could arise from the destabilization of the covalent bond due to the lone pair that leads to a small elastic modulus, and could therefore be present in other hybrid PIMs. Temperature-driven polymorphism can also lead to changes in lattice parameters at different temperatures.²⁹⁰ More effort should therefore be dedicated by theorists in the community to understand finite temperature effects in computations on perovskites as well as the hybrid PIMs.
Polymorphism

Knowing the atomic structure of a material can be challenging because many compounds exhibit polymorphism, in which the materials can exist in different crystal structures.\textsuperscript{99,90} It is important that the relative energies of the polymorphs are accurately described by the DFT exchange-correlation functionals, so that the correct ground-state structure is used for property predictions. DFT is generally considered to accurately calculate the total energy but can, in some cases, fail to predict the correct polymorph. For example, anatase TiO\textsubscript{2} has been found to have a lower energy than the rutile phase,\textsuperscript{90} which is the experimental ground-state.\textsuperscript{91} Similarly, MnO has been described as the tetrahedral wurtzite structure instead of the octahedral rock-salt.\textsuperscript{92} These shortcomings have recently been addressed using the computationally-intensive calculations in the random phase approximation (RPA).\textsuperscript{93,94} which yield the correct ground-state for MnO\textsubscript{6,98} and TiO\textsubscript{2}.\textsuperscript{94}

Weak dispersion interactions

Another factor in accurately calculating atomic structure is the presence of dispersion interactions in many PIMs, which are poorly described with standard electron exchange-correlation functionals. Such interactions influence the interlayer separation in layered crystal structures, such as PbI\textsubscript{2} and BiI\textsubscript{3}. In addition, many PIMs contain low valence cations with an ns\textsuperscript{2} electronic configuration, such as Pb\textsuperscript{2+} or Bi\textsuperscript{3+}, which often occur in asymmetric local environments with low coordination numbers.\textsuperscript{99,100} The directional accumulation of electron density due to the lone-pair formation in BiI\textsubscript{3} is illustrated in Figure 11. Dispersion-corrected functionals have been shown to improve the crystal structure parameters and atomization energies in layered materials.\textsuperscript{101} However, it is unclear how such functionals perform for the polymorph energy ordering. As a many-body theory, the RPA naturally includes dispersion interactions in an \textit{ab initio} fashion, and is used here as a benchmark.

In Tables 3–5 we report the structural parameters and polymorph energies of SnO, SnS, and BiI\textsubscript{3} PIMs. To emphasize the influence of weak dispersion interactions, we include as a reference in these tables hypothetical high-symmetry structures that exclude the symmetry breaking due to the lone pair. These high symmetry structures are the NaCl structure (space group 225) for SnS and SnO, and the ReO\textsubscript{3} structure (space group 221) for BiI\textsubscript{3}. Many computational searches for new materials routinely use the popular PBE version of the GGA functional\textsuperscript{102} for obtaining optimized crystal structures and energies. However, often overlooked is the tendency of PBE to be more accurate in predicting the structural and energetic properties of materials composed of light atoms as compared to the heavy elements in PIMs.\textsuperscript{103} This effect is evident in the computed lattice parameters and volume per atom reported in Tables 3–5. The overestimation of these quantities by PBE, as compared to accurate RPA or experimentally reported values, is pronounced for materials containing heavy elements. The failure of PBE for describing the structural properties of PIMs can be overcome by employing a DFT that explicitly treats vdW interactions. We show in Tables 3–5 that adding a semi-empirical term describing the dispersion forces to the DFT total energy, implemented using the popular DFT-D\textsubscript{2} functional of Grimme,\textsuperscript{203} improves the overestimation in lattice parameters of SnO, SnS, and BiI\textsubscript{3} by PBE. These improvements are achieved with little additional computational cost. Further quantitative improvement of both lattice parameters and polymorph energies of the PIMs studied here is found by explicitly describing the non-local electron-electron correlations omitted by semi-local DFT (e.g., local density and generalized gradient approximations) in the exchange-correlation functional, computed in Tables 3–5 using the optB86b van der Waals functional.\textsuperscript{204} These results demonstrate the need for explicit consideration of nonlocal interactions when computing the structural properties and polymorph energies of PIMs.

![Figure 11. Layered crystal structure of BiI\textsubscript{3} (space group 148) depicting the localized Bi cation 6s lone pair by the red spheres (0.06 e Å\textsuperscript{3} with respect to the superposed atomic densities).](image-url)
Table 3. Structural parameters and polymorph energies of SnO computed using various DFT methods, compared to total energies obtained using RPA and experimentally reported lattice parameters. SG is space group.

<table>
<thead>
<tr>
<th>SnO</th>
<th>SnO-SG225</th>
<th>a (Å)</th>
<th>b</th>
<th>c</th>
<th>V (Å³)</th>
<th>E_{hull} (meV/Å³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE</td>
<td>--</td>
<td>5.12</td>
<td>--</td>
<td>--</td>
<td>16.76</td>
<td>449</td>
</tr>
<tr>
<td>PBE-D2-vdW</td>
<td>5.12</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>16.80</td>
<td>446</td>
</tr>
<tr>
<td>optB86b-vdW</td>
<td>5.08</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>16.40</td>
<td>433</td>
</tr>
<tr>
<td>RPA</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>16.56</td>
<td>454</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SnO-SG14</th>
<th>PBE</th>
<th>5.98</th>
<th>5.09</th>
<th>5.71</th>
<th>21.29</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE-D2-vdW</td>
<td>5.79</td>
<td>4.92</td>
<td>5.49</td>
<td>19.19</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>optB86b-vdW</td>
<td>5.91</td>
<td>4.87</td>
<td>5.57</td>
<td>19.54</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>RPA</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>19.16</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SnO-SG29</th>
<th>PBE</th>
<th>5.60</th>
<th>5.77</th>
<th>5.28</th>
<th>21.33</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE-D2-vdW</td>
<td>5.30</td>
<td>5.41</td>
<td>5.15</td>
<td>18.46</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>optB86b-vdW</td>
<td>5.36</td>
<td>5.63</td>
<td>5.07</td>
<td>19.14</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>RPA</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>18.54</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SnO-SG29</th>
<th>PBE</th>
<th>3.87</th>
<th>3.87</th>
<th>5.02</th>
<th>18.79</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE-D2-vdW</td>
<td>3.84</td>
<td>3.84</td>
<td>4.83</td>
<td>17.80</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>optB86b-vdW</td>
<td>3.84</td>
<td>3.84</td>
<td>4.80</td>
<td>17.67</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>RPA</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>17.40</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Expt</td>
<td>3.80</td>
<td>3.80</td>
<td>4.84</td>
<td>17.49</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Structural parameters and polymorph energies of SnS computed using various DFT methods, compared to total energies obtained using RPA and experimentally reported lattice parameters. SG is space group.

<table>
<thead>
<tr>
<th>SnS</th>
<th>SnS-SG225</th>
<th>a (Å)</th>
<th>b</th>
<th>c</th>
<th>V (Å³)</th>
<th>E_{hull} (meV/Å³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE</td>
<td>--</td>
<td>5.85</td>
<td>--</td>
<td>--</td>
<td>25.02</td>
<td>46</td>
</tr>
<tr>
<td>PBE-D2-vdW</td>
<td>5.74</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>23.69</td>
<td>53</td>
</tr>
<tr>
<td>optB86b-vdW</td>
<td>5.81</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>24.47</td>
<td>44</td>
</tr>
<tr>
<td>RPA</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>24.32</td>
<td>78</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SnS-SG63</th>
<th>PBE</th>
<th>4.12</th>
<th>4.12</th>
<th>11.80</th>
<th>25.01</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE-D2-vdW</td>
<td>4.05</td>
<td>4.08</td>
<td>11.48</td>
<td>23.70</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>optB86b-vdW</td>
<td>4.08</td>
<td>4.09</td>
<td>11.47</td>
<td>23.93</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>RPA</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>23.62</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Expt</td>
<td>4.15</td>
<td>4.18</td>
<td>11.48</td>
<td>24.86</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SnS-SG62</th>
<th>PBE</th>
<th>4.02</th>
<th>4.46</th>
<th>11.41</th>
<th>25.57</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE-D2-vdW</td>
<td>3.99</td>
<td>4.33</td>
<td>11.37</td>
<td>24.57</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>optB86b-vdW</td>
<td>4.01</td>
<td>4.32</td>
<td>11.26</td>
<td>24.36</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>RPA</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>24.64</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Expt</td>
<td>3.99</td>
<td>4.33</td>
<td>11.20</td>
<td>24.18</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Structural parameters and polymorph energies of BiI₃ computed using various DFT methods, compared to total energies obtained using RPA and experimentally reported lattice parameters. SG is space group.

<table>
<thead>
<tr>
<th>BiI₃</th>
<th>BiI₃-SG221</th>
<th>a (Å)</th>
<th>b</th>
<th>c</th>
<th>V (Å³)</th>
<th>E_{hull} (meV/Å³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE</td>
<td>6.17</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>58.80</td>
<td>185</td>
</tr>
<tr>
<td>PBE-D2-vdW</td>
<td>6.10</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>56.78</td>
<td>287</td>
</tr>
<tr>
<td>optB86b-vdW</td>
<td>6.11</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>57.09</td>
<td>266</td>
</tr>
<tr>
<td>RPA</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>55.33</td>
<td>280</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BiI₃-SG162</th>
<th>PBE</th>
<th>7.83</th>
<th>7.83</th>
<th>7.95</th>
<th>52.76</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE-D2-vdW</td>
<td>7.41</td>
<td>7.41</td>
<td>6.92</td>
<td>41.07</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>optB86b-vdW</td>
<td>7.53</td>
<td>7.53</td>
<td>6.96</td>
<td>42.75</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>RPA</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>41.71</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Expt</td>
<td>7.50</td>
<td>7.50</td>
<td>6.90</td>
<td>42.02</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BiI₃-SG148</th>
<th>PBE</th>
<th>7.84</th>
<th>7.84</th>
<th>23.24</th>
<th>51.53</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBE-D2-vdW</td>
<td>7.41</td>
<td>7.41</td>
<td>20.67</td>
<td>49.96</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>optB86b-vdW</td>
<td>7.55</td>
<td>7.55</td>
<td>20.75</td>
<td>42.64</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>RPA</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>41.69</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Expt</td>
<td>7.52</td>
<td>7.52</td>
<td>20.72</td>
<td>42.24</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

Electronic Structure. A common feature of hybrid organic-inorganic perovskites and PIMs is the composition of a heavy metal cation. The presence of heavy atoms necessitates the proper treatment of relativistic effects. In particular, spin-orbit coupling (SOC) scales approximately with the square of the atomic mass. As has been observed in calculations on MAPbI₃, if spin-orbit effects are neglected then fortuitous agreement with the experimental value of the gap can be seen at the DFT level, but this agreement disappears upon the use of more accurate methods such as GW. Given the well-known deficiencies of DFT for calculating electronic structure properties, such as the bandgap, it is best to include spin-orbit coupling into an accurate electronic structure theory such as GW or hybrid functionals. For example, it has been shown that the addition of SOC to calculations on MAPbI₃ reduces the predicted bandgap by 1 eV, resulting in a split-spin conduction band that is associated with slow charge-carrier recombination.

The important optoelectronic properties of semiconductors that can be computed include the strength of the light absorption and emission, their ionization potentials, and their electron affinities. Computationally expedient DFT calculations in LDA (local density approximation) or GGA are well known to strongly underestimate bandgaps. The GW approximation is considered to be the gold standard, but it is computationally-intensive and often requires careful optimization of convergence parameters, which makes it difficult to use, e.g., for high-throughput screening. Hybrid functionals have emerged
over the past decade as a compromise between these methods. An intuitive way to understand hybrid functionals is to make the observation that density functional theory underestimates bandgaps, while Hartree-Fock calculations, which include only exact exchange and no correlation effects, overestimate the bandgaps due to a lack of screening. Hence, these hybrid functionals mix in a fraction of exact exchange into the DFT functionals for electronic interactions within a given length scale. However, a disadvantage of hybrid functionals is that a greater number of external parameters are required. Depending on the implementation, these parameters determine the fraction of the exact exchange and the screening length for the Coulomb interaction.

To illustrate the influence of the fraction of the exact exchange (α) on hybrid functionals, we calculate the bandgap and density of states effective masses for electrons (m_e) and holes (m_h) at different α values for two polymorphs of BiI_3: space group 148 (R3) and 162 (P31m). We use HSE, the most popular hybrid functional for solid-state systems for these calculations, which are shown in Table 6. The bandgaps of two experimentally realized polymorphs of BiI_3 show an expected linear trend with increasing α. These bandgaps vary from 1.71 eV (α=0.20) to 2.17 eV (α=0.40) for space group 148, and from 1.47 eV (α=0.20) to 1.91 eV (α=0.40) for space group 162. The variations in the bandgaps span about 0.5 eV depending on the α value. The HSE-calculated bandgaps match the GW-calculated bandgaps (Table 7) for α values of 0.35-0.40 (space group 148), and 0.30-0.35 (space group 162). These α values are relatively high compared to the standard α=0.25 for the HSE06 functional.

Table 6. Electronic properties of BiI_3 computed using a hybrid HSE-type functional with spin-orbit coupling and varying exact exchange mixing. SG is space group.

<table>
<thead>
<tr>
<th>BiI_3</th>
<th>E_g (eV)</th>
<th>E_g(Direct) (eV)</th>
<th>m_e (m_0)</th>
<th>m_h (m_0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BiI_3-SG148</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α=0.20</td>
<td>1.71</td>
<td>1.79</td>
<td>1.34</td>
<td>5.30</td>
</tr>
<tr>
<td>α=0.25</td>
<td>1.83</td>
<td>1.91</td>
<td>1.33</td>
<td>5.03</td>
</tr>
<tr>
<td>α=0.30</td>
<td>1.94</td>
<td>2.04</td>
<td>1.33</td>
<td>4.80</td>
</tr>
<tr>
<td>α=0.35</td>
<td>2.06</td>
<td>2.15</td>
<td>1.32</td>
<td>4.60</td>
</tr>
<tr>
<td>α=0.40</td>
<td>2.17</td>
<td>2.27</td>
<td>1.32</td>
<td>4.44</td>
</tr>
<tr>
<td>BiI_3-SG162</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α=0.20</td>
<td>1.47</td>
<td>1.47</td>
<td>0.50</td>
<td>2.78</td>
</tr>
<tr>
<td>α=0.25</td>
<td>1.57</td>
<td>1.57</td>
<td>0.50</td>
<td>2.57</td>
</tr>
<tr>
<td>α=0.30</td>
<td>1.68</td>
<td>1.68</td>
<td>0.50</td>
<td>2.40</td>
</tr>
<tr>
<td>α=0.35</td>
<td>1.79</td>
<td>1.79</td>
<td>0.49</td>
<td>2.25</td>
</tr>
<tr>
<td>α=0.40</td>
<td>1.91</td>
<td>1.91</td>
<td>0.49</td>
<td>2.12</td>
</tr>
</tbody>
</table>

Table 7. Electronic properties of BiI_3 computed using GW with spin-orbit coupling.

<table>
<thead>
<tr>
<th>BiI_3</th>
<th>E_g (eV)</th>
<th>E_g(Direct) (eV)</th>
<th>m_e (m_0)</th>
<th>m_h (m_0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BiI_3-SG221</td>
<td>0.63</td>
<td>0.69</td>
<td>3.72</td>
<td>3.38</td>
</tr>
<tr>
<td>BiI_3-SG148</td>
<td>2.05</td>
<td>2.09</td>
<td>1.70</td>
<td>5.15</td>
</tr>
<tr>
<td>BiI_3-SG162</td>
<td>1.82</td>
<td>1.82</td>
<td>0.76</td>
<td>2.49</td>
</tr>
</tbody>
</table>

Figure 12. Comparison of GW (solid lines) and HSE (points) bandstructures for BiI_3 (a) space group 148 and (b) space group 162. Both calculations were performed with spin-orbit coupling included. Dashed lines for space group 148 indicate the locations of the valence band maximum and conduction band minimum.

Some of the limitations of using hybrid functionals can be found by analyzing the agreement with the more accurate, but computationally more expensive GW calculations. In Figure 12 it can be seen that if a good value for α is used for the HSE calculations, good agreement between the two calculations can be reached. In this case, the HSE calculations were made using α=0.35. Comparing Tables 6 and 7 shows that there is good agreement between the bandgaps and effective masses predicted by HSE and GW. It is tempting to conclude that hybrid functionals might provide a computationally less expensive alternative to GW. However, this is only true if a good reference for gauging the fraction of the exact exchange exists. Otherwise, special care needs to be taken when interpreting HSE results in light of the dependence on the choice of α.
Table 8 summarizes the detailed discussion in this article, and we also present our suggestions on protocols for minimum reporting. Following this, we discuss the implications of implementing minimum best practice.

Table 8. Summary of important considerations to avoid artifacts and suggested minimum reporting required

<table>
<thead>
<tr>
<th>Technique</th>
<th>Important considerations/best practices</th>
<th>Suggested minimum reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Crystallography, composition measurements or PES</strong></td>
<td>• Analyze control samples first before performing measurements to understand sample damage (e.g., beam damage, sputter damage) and optimize measurement conditions to minimize sample damage. These can include dosing studies</td>
<td>• Report the composition of the growth and measurement atmosphere (e.g., grown in N₂, but measured in 30 % relative humidity air)</td>
</tr>
</tbody>
</table>
| **Crystallography** | • Twinning and disorder can be difficult to model in single crystal diffraction, especially if light elements are present.  
• Generally, 8-12 data points over the full width at half maximum of the sharpest diffraction peak is needed for Rietveld refinement. The resolution of the diffraction patterns should be to a d-spacing close to one Angstrom or Q ~6.3 Å⁻¹  
• Use 2D diffraction pattern of thin film and powder samples to determine if preferred orientation is present. If a 2D detector is not available, the sample can be rotated and several one-dimensional line scans taken and compared  
• The instrument broadening should be measured using a standard sample before analyzing crystallite size from line broadening | • Report the flux, sample temperature, exposure time, and probe beam wavelength or energy.  
• Best practices in reporting structure determination are given in author guidelines of several chemistry journals, e.g., Ref. 85, 216, 217. The difference curve between measured and calculated data should be reported  
• Raw XRD data should be made available. If a new structure was solved, the crystallographic information file should be made available  
• In addition, for single crystals, the dimensions and appearance should be reported, as should the methods used to model disorder and twinning |
| **Composition measurements** | • Investigate whether artifacts known in the literature are present in the measurement (consult Table 1)  
• Determine whether the sample changes under vacuum during experiment | • Report whether measured composition changes over time  
• Report vacuum level, flux, exposure time, particles used for probe beam (e.g., ions, photons, electrons), wavelength of probe beam  
• Report quantification method and whether the underlying assumptions of the method are fulfilled |
| **PES** | • Calibrate instrument with critically clean (ideally sputter-cleaned) Au, Ag or Cu and compare to Ref. 167  
• Investigate whether the binding energy shifts due to the surface photovoltage effect, sample charging or sample degradation | • Report calibration method  
• Report time in vacuum, radiation flux, exposure time, energy source for probe beam and excitation energy along with the instrument resolution  
• The accuracy of the method used to
| Theory calculations | Safety considerations. In promoting best practices in the exploration of PIMs, we first emphasize the importance of evaluating the materials hazards before any work is begun. While one of the aims of exploring PIMs is to find non-toxic materials, it should not be assumed that lead-free or thallium-free compounds are non-toxic. For example, Sn$^{2+}$ is generally considered less toxic than Pb$^{2+}$,218 and some Sn-based compounds, such as SnS,219 are indeed not hazardous. However, other compounds, such as SnCl$_{3}$, have the same oral toxicity classification as PbI$_{3}$,220,221 Consultation of existing safety data sheets and literature on the toxicity of new materials, if available, and designing the experiment to minimize safety risks is an essential pre-requisite to working with new PV absorber materials, especially if handled outside the glovebox. It should also not be assumed that handling hazardous chemicals inside the glovebox is risk-free. Several solvents can easily penetrate through butyl gloves, such as diethyl ether. | • The leading edge of the valence band should be fit with the instrument-broadened density of states. Approximations through other methods can be used | • Crystal structure needs to be accurately known for consistent calculations | • Fit the leading edge should be discussed |
| | • Carefully converged RPA calculations may more accurately predict the relative energies of polymorphs. | • Report full sample structure, including film thickness and substrate the material measured was grown on | • Calculations of layered materials should include van der Waals interactions | • Report the crystal structure or space group used for performing the calculations, ideally including the optimized structure file |
| | • Inclusion of SOC is necessary for accurate electronic structures and correct description of band edge effective masses of PIMs | • Report critical computational parameters whose convergence may affect the results, e.g., the pseudopotentials, basis set expansion and $k$-point sampling of reciprocal space | • Hybrid functionals are an improved method over DFT for calculating optoelectronic properties, but GW calculations are required for parameter-free predictions. | • Report the calculation methods (LDA, GGA, GW, HSE) used |
| | • Finite temperature effects and temperature-driven polymorphism should be considered | • For crystal structure calculations, report whether non-local interactions (e.g., van der Waals interactions) were taken into account | | • Comments |

**Safety considerations.** In promoting best practices in the exploration of PIMs, we first emphasize the importance of evaluating the materials hazards before any work is begun. While one of the aims of exploring PIMs is to find non-toxic materials, it should not be assumed that lead-free or thallium-free compounds are non-toxic. For example, Sn$^{2+}$ is generally considered less toxic than Pb$^{2+}$,218 and some Sn-based compounds, such as SnS,219 are indeed not hazardous. However, other compounds, such as SnCl$_{3}$, have the same oral toxicity classification as PbI$_{3}$,220,221 Consultation of existing safety data sheets and literature on the toxicity of new materials, if available, and designing the experiment to minimize safety risks is an essential pre-requisite to working with new PV absorber materials, especially if handled outside the glovebox. It should also not be assumed that handling hazardous chemicals inside the glovebox is risk-free. Several solvents can easily penetrate through butyl gloves, such as diethyl ether.

**Best practice in materials characterization.** Before characterizing a material, it is critical to understand the common artifacts of the technique considered and whether it has a suitable interaction volume, depth resolution, detection limit and accuracy to unambiguously test the hypotheses considered. For example, if considering to measure the work function of a sample by UPS, the sample preparation method should be designed to minimize surface contamination method be designed to minimize surface contamination due to the highly surface-sensitive nature of the technique; the binding energy scale needs to be calibrated for the instrument on a clean metal surface; surface charging effects need to be considered and minimized through sample preparation. Comments on how common artifacts were addressed should be made more routinely in publications.

An artifact common to materials characterization of perovskites and PIMs is sample damage, especially if volatile or mobile components are present in the material. This may be through interaction with the probe beam or after prolonged vacuum exposure. Sample damage should be quantified through experiments with control samples. Regular checks of the sample during prolonged measurements should be made to determine whether the sample has been altered. For example, if performing PES measurements of hybrid organic-inorganic materials over an extended period of time, scans should be made at regular intervals to determine if compositional changes occur. The experiment should be designed to minimize sample damage before performing the important measurements. The optimized measurement conditions
should be reported and the minimum parameters to report are given in Table 8.

An important part of designing experiments to avoid artifacts is to understand the limitations of the technique. For example, if it is important to understand the position of light elements (e.g., methylammonium cations), neutron diffraction is more suitable than XRD on bulk powders or single crystals. As another example, if it is important to measure defect concentrations accurate to 100 ppb, then literature on oxide perovskites would suggest that temperature-dependent conductivity measurements are more suitable and sensitive than the more common stoichiometry characterization techniques. If possible, complementary techniques should be used as part of the experiment. For example, when determining the electronic structure of a material, characterization by UPS, XPS, IPES and optical measurements should be used to check that the same band positions are measured.

When analyzing data, inaccuracies can arise due to the inappropriate use of quantification methods. For example, when analyzing the phase of a thin film, Rietveld refinement is less appropriate if texture is present and results in missing peaks or a change in peak intensities. Alternatively, if the primary edge of a UPS spectrum is fit with a parabolic model without knowing the DOS of the material, errors can arise if the DOS at the valence band maximum significantly deviates from a parabola or has contributions from multiple points in the Brillouin zone. In this specific case, the DOS of a new material should always be calculated first and fit to the spectra. But in general, the validity of the assumptions underlying the analysis should be addressed and reported.

A similar approach to those detailed above should be taken for optical characterization and electrical measurements of perovskites and PIMs. These include Kelvin probe force microscopy, electroluminescence measurements, UV-visible spectrometry for bandgap determination, and time-resolved photoluminescence spectroscopy, as detailed in the introduction, with additional useful explanations in the PES section. For example, time-resolved photoluminescence measurements often show single crystals to have a fast initial decay, followed by a slower decay. While the fast decay is typically attributed to surface recombination, others note that fast non-radiative recombination processes should be considered. Degradation and changes in the optoelectronic properties of the material due to sample heating by excitation laser may also be an important factor to consider as part of best practices. The excitation wavelength and fluence used are also critical factors that can influence the measured quantity. The measurement atmosphere (whether air, N₂, or vacuum), and its influence on degradation during measurements is another critical consideration. While putting together the best practices for electrical measurements, optical characterization and spectroscopy is beyond the scope of this work, it should be undertaken because these techniques provide critical information to understand the material.

Important considerations to enable best practice in theory calculations. It is critical to have accurate crystal structure information before performing theory calculations. Best practice in preparing samples and analyzing the measurements to solve the crystal structure is given in the crystallography sub-section of the materials characterization section above. Predicting the ground-state phase of a material is challenging with standard DFT functionals. The random phase approximations (RPA) can be more accurate, but vibrational contributions to the free energy may also have to be considered.

Non-local interactions can affect the lattice parameters and energy. These include van der Waals interactions, which are prevalent weak dispersion interactions in layered crystal structures. In our analysis of layered PIMs, we found that the most accurate calculations were obtained by using the optB86b van der Waals functional.

In PIMs with heavy metal cations, it is particularly important to incorporate the effects of spin-orbit coupling in order to more accurately calculate the electronic properties. The bandgap and other optoelectronic properties can be predicted with hybrid functionals. This has the advantage of requiring less experience to use than the GW approximation. However, it is critical to understand the influence of the fraction of the exact exchange. Where possible, the hybrid functional calculations should be compared with other methods. As shown in the theory section, if hybrid functionals are used with the incorrect exact exchange, the predictions can be significantly inaccurate.

CONCLUSIONS

In summary, lead-halide perovskites and related compounds present important challenges for analysis by materials characterization and calculation that need to be addressed. When planning experiments, the artifacts discussed in this article associated with each technique should be carefully considered. We suggest protocols for the execution and reporting of crystallographic, composition, photoemission spectroscopy and computational analyses, summarized in Table 8. The authors believe that following these guidelines for minimum best practice would lead to greater reproducibility and fewer misinterpretations of data. We
hope these guidelines will lead to greater debate within the wider community on best practices.

**ASSOCIATED CONTENT**

**Supporting Information**

Raw X-ray diffraction data for Figure 1. The Supporting Information is available free of charge on the ACS Publications website.

**AUTHOR INFORMATION**

**Corresponding Author**
* T.B.: E-mail, buonassisi@mit.edu
* R.L.Z.H.: E-mail, rlzh2@cam.ac.uk

**Author Contributions**
The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript. R.L.Z.H. organized the paper outline, compiled the sections, made the samples for the composition section and co-wrote the composition section. P.S. wrote the PES section. L.T.S., K.H.S. and M.F.T. compiled the crystallography section. A.M.H., D.V.-F., S.L. and V.S. put together the DFT section. J.D.P. and S.S. took the XRF measurements, compositional analyses and co-wrote the composition section. T.B. conceived of the paper, organized the paper outline and edited the paper. D.O.S., A.Z., A.W., I.C.S., B.C.M., R.C.K., Y.W., J.S., F.C.M., J.J.B., W.T. were attendees of the bismuth seminar held at the 2015 Fall MRS workshop and helped us establish community consensus on best practices. D.O.S. and A.W. contributed to the theory section. A.Z. contributed to the crystallography and PES sections. J.J.B. and W.T. contributed to the organization, direction and tone of the paper through their edits and broad overview of the field. I.C.S. contributed to the crystallography section. B.C.M. contributed to the introduction, crystallography and composition sections. R.C.K. contributed to the crystallography and theory sections. Y.W. and J.S. contributed to the crystallography and outlook sections. F.C.M. contributed to the outlook section and edited the paper.

**Notes**
The authors declare no competing financial interests.

**ACKNOWLEDGMENT**
The authors thank Dr. Austin J. Akey, Rebecca A. Belisle, Dr. Riley E. Brandt, Jeremy R. Poindexter, Douglas Fabini, Prof. Dr. Douglas Galvão, Natália F. Coutinho and Dr. Cristiano Woellner for technical discussions. This work was supported as part of the Center for Next Generation Materials by Design (CNGMD), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Science under Contract No. DE-AC52-76SF00515. P.S. was supported by the hybrid perovskite solar cell program of the National Center for Photovoltaics funded by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Office of Solar Energy Technology under Award Number DE-AC36-08GO28308DOE with the National Renewable Energy Laboratory (NREL). A.Z. was supported by the ‘Rapid Development’ project funded by the same agency. Y.W. and J.S. were supported by a start-up fund from Rensselaer Polytechnic Institute and the National Science Foundation under grant CMMI 1509484. F.C.M. was supported by MIT-Brazil/FAPESP Grant 2012/10127-5 and FAPESP Grant 2014/50718-8.

**REFERENCES**


(25) Ganose, A. M.; Butler, K. T.; Walsh, A.; Scanlon, D. O. Relativistic Electronic Structure and Band


(44) Stranks, S. D.; Eperon, G. E.; Grancini, G.; Menelaou, C.; Alcocer, M. J. P.; Leijtens, T.;...


(66) Godby, R. W.; Needs, R. J. Metal-Insulator Transition in Kohn-Sham Theory and


(89) Zhumekov, A. A.; Saidaminov, M. I.; Haque, M. A.; Alarousu, E.; Sarmah, S. P.; Murali, B.; Dursun, I.; Miao, X.-H.; Abdelhady, A. L.; Wu,


(128) Fadley, C. S.; Shirley, D. a. Electronic Densities of States from X-Ray Photoelectron


(170) Hecht, M. Photovoltaic Effects in Photoemission Studies of Schottky Barrier


(188) Ruck, M. Darstellung Und Kristallstruktur von Fehlordnungsfreiem Bismuttriiodid. Zeitschrift


Insert Table of Contents artwork here