Chemical Robotics Enabled Exploration of Stability and Photoluminescent Behavior in Multicomponent Hybrid Perovskites via Machine Learning

Kate Higgins[1], Sai Mani Valleti[2], Maxim Ziatdinov[3],

Sergei V. Kalinin*[,2,3], and Mahshid Ahmadi[1]*

[1] *Joint Institute for Advanced Materials, Department of Materials Science and Engineering, University of Tennessee, Knoxville, TN 37996, USA*

[2] *The Bredesen Center, University of Tennessee, Knoxville, TN 37996, USA*

[3] *The Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*

*Corresponding authors email:

mahmadi3@utk.edu, sergei2@ornl.gov

Abstract:

Hybrid organic-inorganic perovskites have attracted immense interest as a promising material for a variety of optoelectronic and sensing applications. However, issues regarding long-term stability have emerged as the key bottleneck for applications and still require further study. Here, we develop automated experimental workflow based on combinatorial synthesis and rapid throughput characterization to explore long-term stability of these materials in ambient conditions, and apply it to four model perovskite systems: $MA_xFA_yCs_{1-x-y}PbBr_3$, $MA_xFA_yCs_{1-x-y}PbI_3$, $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ and $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$. We have both established a new workflow and found out the main tendencies in the mixed cation and anion systems, which led to the discovery of non-trivial composition regions with high stability. The Non-negative Matrix Factorization and Gaussian Process regression are used *to interpolate the photoluminescent behavior of vast compositional space* and *to study the overall behavior of the phase diagram*. This interpolative regression analysis helps to distinguish mixtures that form solid solutions from those that segregate into multiple materials, pointing out the most stable regions of the phase diagram. We find the stability dependence on composition to be extremely non-uniform within the composition space, suggesting the presence of potential preferential compositional regions. This proposed workflow

is universal and can be applied to other perovskite systems and solution-processable materials. Furthermore, incorporation of experimental optimization methods, e.g., those based on Gaussian Processes, will enable the transition from combinatorial synthesis to guide materials research and optimization.

Hybrid organic-inorganic perovskites (HOIPs) have received immense research attention due to the combination of outstanding optoelectronic properties and low fabrication cost, making them uniquely attractive for numerous applications, including photovoltaics[1], light-emitting diodes[2,3] photodetectors[4], and ionizing radiation sensors[5,6]. In general, three-dimensional (3D) HOIPs adopt an ABX_3 structure and are typically comprised of a monovalent cation, A = ($CH_3NH_3^+$ (MA); $CH(NH_2)_2^+$ (FA); Cs^+), a divalent metal cation, B = (Pb^{2+}; Sn^{2+}), and halide anion X = (Cl^-; Br^-; I^-). More than a thousand perovskite-inspired candidate compounds have been theoretically predicted during the last few years,[7-10] a number that can be expanded further via doping and alloying to generate a virtually unlimited range of possible solid solutions.

Despite the extensive theoretical effort, the synthesis of these materials preponderantly relied on Edisonian trial-and-error approach. This synthesis usually entailed modifying a single compositional or synthesis variable and observing the changes in structure and functionalities. Combined with lengthy processing and optimization times, this approach has been inefficient in its ability to explore vast design spaces. Furthermore, the capability to predict the outcomes quantitatively, offer a reasonable understanding of variable correlations, factors challenging reproducibility, and parameters necessary to facilitate the commercialization of the material of interest has been lacking. Consequently, despite theoretical predictions, only a small fraction of predicted HOIP compounds has been experimentally realized in the form amenable to materials property (high-quality thin film) and single-crystal device level studies.

Currently, the primary concern at the forefront of scientific and application-driven research is the stability of these materials in the pure or device-integrated form[11,12]. Ultimately, the application potential of HOIPs will be determined equally by the intrinsic properties of pure HOIP materials and by material stability and responses to environmental conditions such as light, humidity, and heat. Recently, it has been demonstrated that the alloying or solid solution formation can alleviate many of these factors. Alloying to form solid solutions between $FAPbI_3$, $MAPbBr_3$, and $CsPbI_3$ has shown to be the pathway to combat the instability issue and allow for the development of more stable HOIP photovoltaics[13-15]. However, solid solutions in HOIP families that derive from the alloying and substitution on monovalent and divalent cations and anion sites can lead to the endmembers belonging to different crystallographic groups and possessing dissimilar structural and optoelectronic properties and non-monotonic stability over compositional space. The interplay of the variability of endmember structures in conjunction with the classical

thermodynamic factors such as limited mutual solubilities gives rise to complex phase spaces with multiple solid solutions and two- and three-phase regions, etc. A large number of possible solid solutions, the fact that different aspects of the figure of merit are optimized in the other areas of the phase diagram, and the need for optimization of synthesis conditions for each specific composition make the classical search and optimization of these materials extremely tedious. To date, only a few tens of solid solution systems have been explored[16-18] compared to over thousands of possible variants analyzed theoretically[7-9] and hundreds of thousands of possible 3- and 4-composition systems.

Recently, several groups have demonstrated the applications of the automated experiment to materials discovery, including synthesis[19-21] and full device-preparation workflows[22]. For example, MacLeod et al. developed an autonomous research system that produces and characterizes a thin film sample[22]. It then utilizes software that applies a Bayesian optimization algorithm to determine the next sample to synthesize. Such approach of iterating between automated experimentation and machine learning-based experiment planning has resulted in early successes in addressing high-dimensional problems in experimental physics[23], chemistry[24], life-sciences[25], and recently starting to be implemented in the materials sciences[26-30]. High throughput pipelines for the discovery of new perovskites for solar cells have also been reported. Sun et al. fabricates and characterizes 75 different perovskite solar cells over two months[18]. A deep neural network diagnostic was utilized for accelerated X-ray diffraction analysis of the fabricated perovskite thin films. These works are mostly focused on the optimization of optoelectronic devices. The combinatorial approaches in materials synthesis using techniques as similar to pulsed laser, physical vapor, and chemical vapor depositions, etc., have been actively explored for over 20 years[31].

However, until now, solution-based combinatorial synthesis has been limited to custom design microfluidic or robotic manipulation systems[16,19,20,32]. Also, until recently, the limiting factors for combinatorial methods were the availability of high-throughput characterization techniques and machine learning tools that allow the characterization results to be integrated into the correlative models and further reduced to predictive physical laws. Note that continuous combinatorial libraries like those used in pulsed laser deposition also require composition characterization, since the local composition can significantly differ from the expected due to the specifics of the deposition process. Here, the use of the 96-well plate-based libraries addresses

4

both issues. First, the compositions, in this case, are controlled by the pipetting/dilution process and endmember concentrations and, as such, are determined precisely (or at least as good as macroscopic synthesis). Second, using the 96-well plate allows for simple characterization via the optical spectrometry enabled by the optical plate reader. In addition, the synthesis and characterization of up to 96 different compositions can be done at one time. The defining characteristic of this study is the automated exploration, characterization of vast compositional spaces and utilizing the material knowledge to determine the optimal compositional regions to exploit for specific optoelectronic properties. An added perk to the developed workflow is that it is also cost-effective as compared to other autonomous research systems described in the literature.

In this work, we establish the workflow for the rapid synthesis and characterization of HOIPs via combinatorial synthesis combined with rapid throughput photoluminescent (PL) screenings. To gain insight into the variability of the optical band gap and PL properties across the compositional series, we adopt an approach based on the multivariate statistical analysis. We map composition-optical property (luminescence) and stability relationships simultaneously across the ternary phase diagram and use the Gaussian Processing framework to determine associated uncertainties. This study establishes the utilization of automated synthesis and enables the creation of data-driven machine learning (ML) models for accelerated discovery of large compositional space in HOIPs with optimized properties for multifunctional optoelectronics. This approach can be extended to higher-dimensional spaces via the evolutionary experiment optimization in the future studies.

Results and Discussion

To effectively explore the photoluminescent properties over multidimensional compositional spaces, we develop a workflow combining robotic synthesis, automated characterization, and machine learning techniques, as shown in **Figure 1**. Here, robotic synthesis is used to create a combinatorial library all at once in a chosen multidimensional compositional space. Optoelectronic properties and their evolution with time are explored using automated optical methods, providing insight into the bandgap energy and the stability of the resulting materials. The Gaussian Process optimization, the machine learning technique employed here, is used to extrapolate the observed behaviors across the compositional space, and the associated uncertainties can be further used to guide the selection of new compositional points. Overall, the

1   developed workflow is an exploratory one that prioritizes materials discovery but can be used for

2   the optimization of optoelectronic devices. Here, we performed only one iteration.

3

4
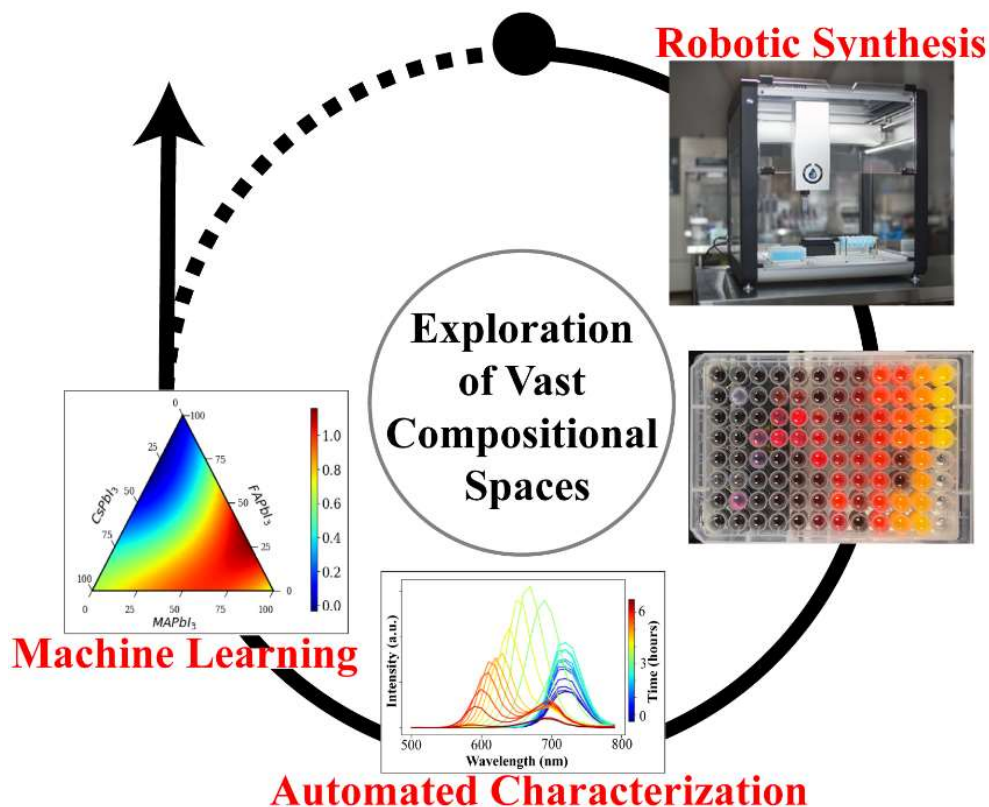
5   **Figure 1.** A schematic of the automated workflow for the combinatorial discovery of HOIP
6   materials. The combinatorial library is first created in a chosen multidimensional compositional
7   space using the pipetting robot (robotic synthesis). The precipitation of 92 microcrystals of HOIP
8   is indicated by the change in color in a well-plate. Photoluminescent properties are measured over
9   the combinatorial library as a function of time (automated characterization). This data is analyzed
10   via machine learning methods to yield information on the time-dependent behaviors in the system
11   that can be further used to navigate the compositional space.

12

13   In this study, we chose to explore several 3D ABX_3 ternary and quasi-ternary systems of

14   mixed A cations and X anions. As a model, we decided on Cs-MA-FA as the cation solutions and

15   I^- and Br^- as anions to verify our observations relative to previously explored endmembers and

16   explore optoelectronic properties and stability in the individual systems. For the combinatorial

17   libraries of triple (monovalent) cations, 92 combinations of $(Cs/MA/FA)PbI_3$ and 92 combinations

18   of $(Cs/MA/FA)PbBr_3$ solid solutions are synthesized in two separate well plates. For the

combinatorial library of mixed cations and anions, 184 combinations of $(Cs/MA/FA)Pb(I/Br)_3$ are synthesized in two independent 96 well plates. The photos of these plates are shown in **Figure S2 (a-b)**. It is important to note that this approach is universal and can be applied to other perovskite systems.

Firstly, utilizing a programmable pipetting robot and employing classical 96-well microplates shown in **Figure 1**, we create combinatorial libraries of ternary $MA_xFA_yCs_{1-x-y}PbBr_3$ and $MA_xFA_yCs_{1-x-y}PbI_3$ and quasi-ternary $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ and $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$ systems. The composition of each well is controlled by pipetting the desired quantity of endmember solutions. Precipitation of microcrystals occurs via an antisolvent approach[6,16]. Each system, consisting of 92 samples, is synthesized in a separate well plate. Details of synthesis and photos of each microplate are provided in the supplementary information. We utilize an automated Multi-Mode well plate reader with the capability to perform photoluminescence (PL) spectroscopy to characterize compositional dependent optoelectronic properties. Details of these measurements are provided in the supplementary information. Recent studies have demonstrated that the dynamical processes related to HOIPs formation, ionic movement, degradation, and phase changes can be understood by studying the evolution of their luminescent properties[33-35]. Although a relatively simple measurement to perform, characterizing a combinatorial library with a conventional PL set-up would be laborious. By utilizing the automated reader, we are effectively able to measure the optoelectronic properties of 96 compositions in roughly five minutes. Repeating this process for several hours yields the set of time- and composition-dependent data on the optoelectronic properties and stability of the chosen system. It should be noted that in this study, the intrinsic stability of each composition in ambient conditions is explored. Here, the focus is not on the optimization of the electronic properties, but the exploration of vast compositional spaces and time evolution of PL properties, i.e., PL aging. Thus, we are able to discover the compositions or compositional regions in the phase diagram that might be suitable for optoelectronic applications, such as solar cells.

Firstly, we explore the PL properties of the four materials systems before prolonged exposure in ambient conditions. Upon further inspection of this exploration, we can provide conclusions about what region of the phase diagram would be optimal for certain applications. **Figure S3** in supplementary information shows the corresponding PL spectra of the yielded microcrystals immediately after precipitation by chloroform. Each panel illustrates the

composition-PL emission from the mixture of three precursors with 92 different concentrations and the formation of three-component solid solutions. As can be seen, the shift in the emission wavelength in quasi-ternary systems with varying concentrations of halide is significant. Also, it is particularly challenging to understand the trend in alloying behavior. To get insight into the variability of the optical band gap and PL properties across the compositional series, we adopt an approach based on the multivariate statistical analysis.

For ternary systems, this synthesis and characterization yield a set of spectra as $A(x, y, \omega, t)$, where A is the PL emission, x and y define the composition, ω is the wavelength, and t is the time since synthesis. As a first step towards gaining insight into the compositional dependence of optoelectronic properties, we choose the initial PL spectra $L_0(x, y, \omega) = L(x, y, \omega, 0)$ and analyze the spectral dataset using unsupervised linear unmixing. This set of methods generally includes techniques such as Principal Component Analysis (PCA), Non-negative Matrix Factorization (NMF), and Independent Component Analysis (ICA), with the choice of decomposition method dictated by the physical constraints on endmembers and loading maps. We choose NMF as the simplest method that gives rise to non-negative components and decomposes the initial time spectra as

$$L_0(x, y, \omega) = \sum_{i=1}^{N} L_i(x, y) w_i(\omega) \tag{1}$$

where $L_i(x, y)$ are the loading maps that represent the variability of spectral behaviors across compositional space and $w_i(\omega)$ are the endmembers that determine characteristic spectral behaviors. The number of components, N, is set at the beginning of the analysis and can be chosen based on the quality of decomposition, anticipated physics of the systems, etc..

It is important to note that blind linear separation methods explicitly assume that the nature of the characteristic behaviors, $w_i(\omega)$, is unknown, but the system is linear in these components. If the endmembers are assumed to be known, e.g., comport to the spectra of pure endmembers, then Eq. (1) will become the linear regression model, and additional unknown components can be introduced to account for chemical reactivity and formation of new compounds. Such an approach is expected to be beneficial for spectrophotometric data (i.e., optical absorption). However, this is not the case for PL data that can show complex behavior as a function of composition, crystallinity, defect properties, and environmental stability. Hence, here we use unconstrained linear blind

separation to allow for possible chemical reactions and morphological evolution in the system. At the same time, we limit the analysis to the linear unmixing as a natural first approximation.

To interpolate between the NMF components, we use the Gaussian Processing (GP) regression.[36,37] GP is a general approach for determining the unknown function, f, from a set of noisy observations $D = (x_1, y_1), \ldots (x_N, y_N)$. The constraints of the GP model are very general, and it is only assumed that the observational noise ε has zero mean. Given the observations, the function f is constructed via Bayesian inference in a function space. In this analysis, it is assumed that the function f has a prior distribution $f \sim \mathcal{GP}(0, K_f(x, x'))$, where K_f is a kernel function. The kernel describes the correlations between the values of the function across the parameter space. The functional form of the kernel is postulated at the beginning of the process, whereas the kernel hyperparameters are optimized during the regression, $y = f(x) + \varepsilon$. The learned kernel is then used for predicting function values at new points. The choice of the kernel can be dictated by the known physical constraints on the system; in the cases when such prior knowledge is absent, the typical option is the Gaussian (or Radial Basis Function, RBF) kernel, which corresponds to a Bayesian linear regression model with an infinite number of basis functions[36]. The output of the GP regression is the predicted function values and uncertainty maps representing the quality of prediction. Additionally, the learned kernel parameters can yield information on the physics of the system. We note that because GP is formulated in a Bayesian setting, the uncertainty quantification is inherent in GP-based methods and this has been the primary motivation for choosing this method as compared to readily available simpler interpolation schemes since the uncertainty allows the formulation of the strategies for exploratory and exploration-exploitation searches through the composition space as will be explored in the future.

Here, GP is used as a universal interpolator to convert the set of measurements in the chosen compositional space into the dense maps of functional responses. The GP interpolation is implemented using the GPim library[32] in Python. The visualization of functional behaviors in ternary diagrams is performed using the ternary library.
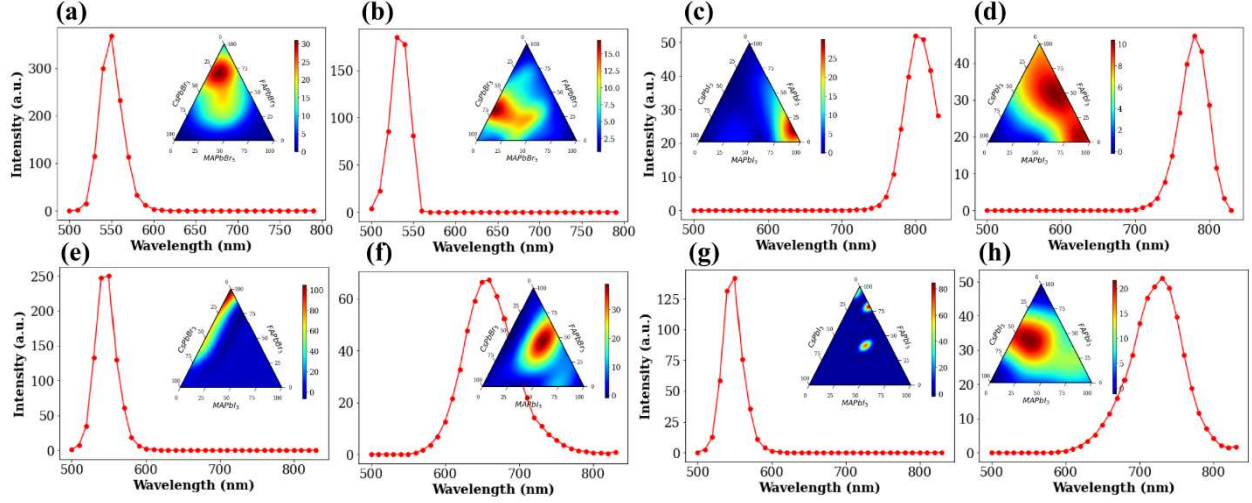
Figure 2. Multivariate analysis of the PL data immediately after synthesis for the explored compositional spaces: **(a,b)** $MA_xFA_yCs_{1-x-y}PbBr_3$, **(c,d)** $MA_xFA_yCs_{1-x-y}PbI_3$, **(e,f)** $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ and **(g,h)** $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$. The PL spectrum shows the characteristic behavior for each composition series and the inset ternary diagrams are the compositional dependencies of the PL intensities. The ternary inset plots are the loading maps that represent the variability of the spectral behaviors across the compositional space. Specifically, the intensity of the color map indicates how the characteristic spectrum describes that composition. The corresponding triangle maps before GP interpolation is found in supplementary information, Figure S4.

Illustrated in **Figure 2** is this combined NMF-GP analysis on the explored ternary phase diagrams. Here, we chose $N = 2$ as the simplest representation of variability in the system, as explained in **Figure S5**. The ternary phase diagrams represent $L_i(x, y)$ where x and y are the compositions of the precursors and L_i is the weight of this characteristic spectrum for each composition. The density of experimental data points in the triangle plots is shown in **Figure S1** to guide how the sample points were chosen in the phase diagram. In the ternary phase systems, $MA_xFA_yCs_{1-x-y}PbBr_3$ and $MA_xFA_yCs_{1-x-y}PbI_3$, we directly observe the effect of cation doping on photoluminescent behavior. In the $MA_xFA_yCs_{1-x-y}PbBr_3$ system, the first component shown in **Figure 2(a)** demonstrates that $FAPbBr_3$-rich compositions, represented by a red color in the inset, have a higher peak wavelength, and therefore lower bandgap energy than compositions, which are primarily solid solutions of $CsPbBr_3$ and $FAPbBr_3$. Highlighted in red of the second component in **Figures 2(b)** are the solid solutions mainly between Cs and MA region of the phase diagram with smaller emission wavelength and higher energy bandgaps. Similar cation effects are also apparent

in the $MA_xFA_yCs_{1-x-y}PbI_3$ system. MAPbI$_3$-rich compositions shown in **Figures 2(c)** have a higher peak wavelength than other solid solutions shown in **Figures 2(d)**.

In the quasi-ternary phase systems, $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ and $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$, halide effects become more apparent. In the $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ system, we observe a distinct contribution of the Br$^-$ as shown in the first component in **Figure 2(e)**. For this system, the second component shown in **Figure 2(f)** demonstrates much of the phase diagram is a solid solution with a broad emission having a peak wavelength of around 650 nm. In the $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$ system, the first component in **Figures 2(g)** illustrates specific regions where the endmember corresponds to a wavelength close to 540 nm, clearly identifying the compositions that have undergone halide segregation upon synthesis and would not be suitable for optoelectronic devices. As shown in **Figure 2(g)**, the majority of the phase diagram is dark blue, representing that the weight of this characteristic spectrum is approximately zero. Conversely, the red dots within the same phase diagram represent where this characteristic spectrum is highly weighted. If the GP were not perfomed along with NMF, there would be sample points within the phase diagram instead of the continuous image (**Figure S4**) because GP interpolates the weights between each sample point. Conversely, the second component shown in **Figure 2(h)** establishes that a majority of the phase diagram is a solid solution with a peak wavelength of around 720 nm. This analysis can provide crucial information about the optimal compositions for certain applications.
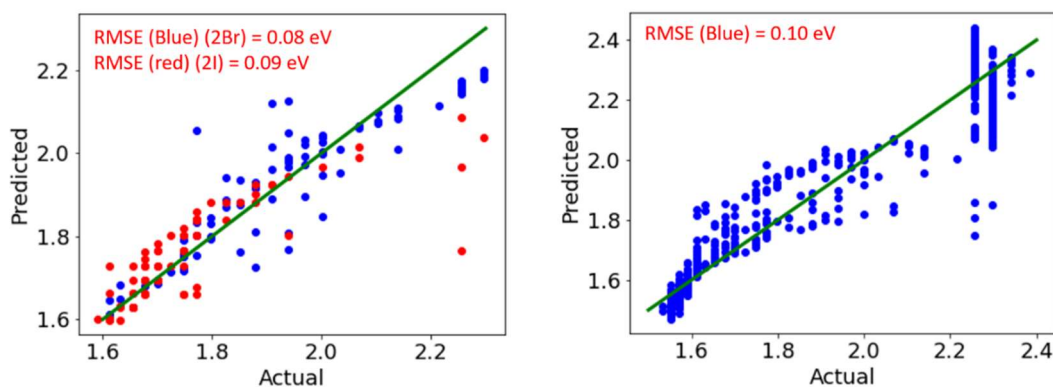
Figure 3. (a) Comparison of the linear prediction and actual bandgap for the $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$ (red) and $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ (blue) systems. **(b)** Comparison of linear prediction and actual bandgap of the interpolated quaternary system. The green line is plotted to show the deviation from slope 1 and the root mean squared error (RMSE) is calculated for evaluating the quality of the surrogate model.

1   This data can further be used to analyze the compositional dependence of the bandgap

2   energy. In this analysis, the position of the photoluminescence maximum is used as a measure of

3   the bandgap in the material, and the concentration dependence of the latter can be analyzed. Here,

4   we use the simple linear estimator to explore the cation and anion effects on the bandgap. For the

5   pure Br system, the linear regression $Eg = a_{Br}x + b_{Br}y + c_{Br}$ where x and y correspond to

6   concentrations in $MA_xFA_yCs_{1-x-y}PbBr_3$, yielding a_{Br} = -0.046, b_{Br} = -0.1, and c_{Br} = 2.33.

7   Similarly, for $MA_xFA_yCs_{1-x-y}PbI_3$, the corresponding parameters are a_I = -0.041, b_I = 0.021, and

8   c_I = 1.58 satisfying the Vegard's law for bandgap energy as function of cation concentration in the

9   HOIPs.

10   In the quasi-ternary systems $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$ and $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-}$

11   $_y)_3$ the composition dependence of the bandgap energy can be explored similarly, albeit for specific

12   parameter combinations. Shown in **Figure 3(a)** is the comparison of the linear prediction and

13   actual bandgap energies for the two systems. As can be seen, there is a general agreement between

14   the predicted and the actual bandgap energy from the measurements, with an evidence of small

15   deviation from Vegard's law and a near-linear behavior in the bandgap energy in these systems.

16   At the same time, the analysis of the quaternary diagram formed by aggregation of the data for all

17   four plates illustrates significant deviation form linear model, **Figure 3(b)**, indicative of the

18   formation of more complex solid solutions and potential chemical segregation in higher

19   dimensional composition space.

20
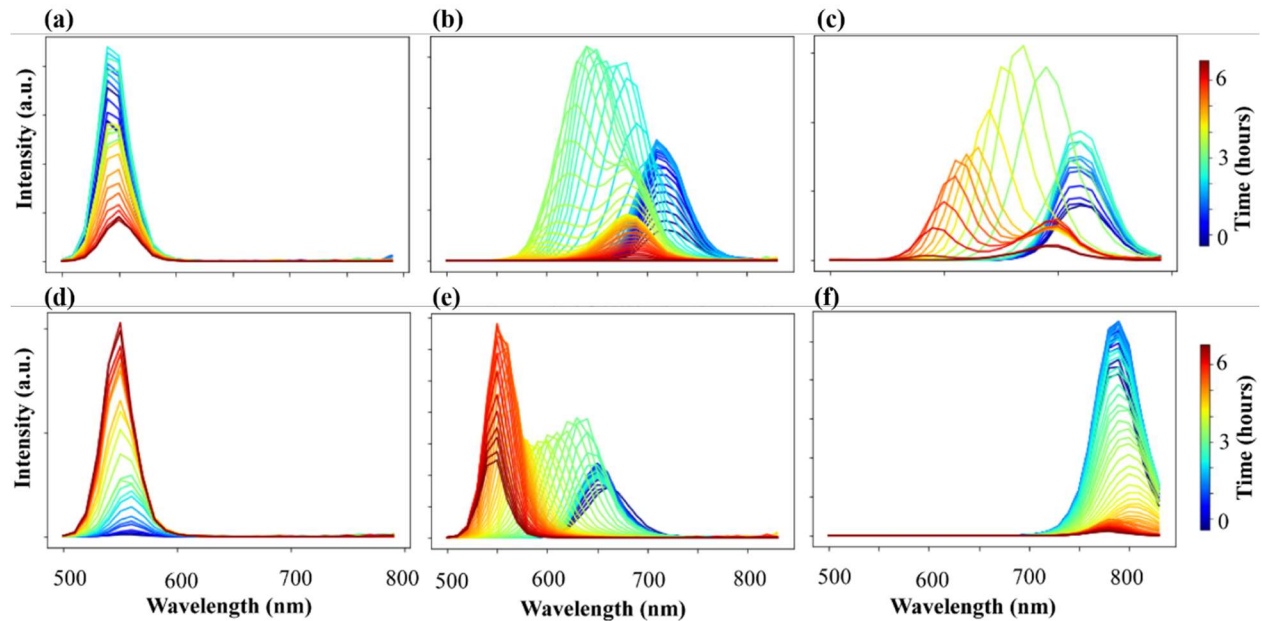
Figure 4. Time dependent PL spectra for several compositions in ambient condition: **(a)** $MA_{0.84}FA_{0.16}PbBr_3$, **(b)** $MA_{0.60}FA_{0.32}Cs_{0.08}PbI_{2.04}Br_{0.96}$, **(c)** $MA_{0.68}FA_{0.28}Cs_{0.04}PbI_{2.04}Br_{0.96}$, **(d)** $MA_{0.20}FA_{0.72}Cs_{0.08}PbI_{0.60}Br_{2.40}$, **(e)** $MA_{0.44}FA_{0.48}Cs_{0.08}PbI_{1.32}Br_{1.68}$ and **(f)** $MA_{0.36}FA_{0.56}Cs_{0.08}PbI_3$.

To gain insight into the compositional dependence of stability in these systems, we explore how the PL behavior of each composition changes as a function of time, providing insight into the stability of these microcrystals in ambient conditions. Shown in **Figure 4** are several representative time-dependent photoluminescent behaviors. For the binary cation composition, $MA_{0.84}FA_{0.16}PbBr_3$, shown in **Figure 4(a)**, we observe a PL enhancement within the first 3 hours, indicating potential traps states are being passivated as time progresses under ambient conditions in agreement with previous PL studies.[33] The PL intensity gradually begins to decrease over the remaining 3 hours. However, even after 6 hours of exposure in air, the PL peak does not entirely disappear, indicating that this binary composition did not fully degrade. Conversely, as shown in **Figure 4(b)**, the PL intensity for the quasi-ternary composition, $MA_{0.6}FA_{0.32}Cs_{0.08}PbI_{2.04}Br_{0.96}$, exhibits a more complicated behavior, in which there is a combination of peak intensity enhancement, peak position shift, peak separation, peak intensity decrease and, finally, the complete disappearance of the peak after approximately 6 hours in ambient conditions, demonstrating full degradation of this perovskite composition. Notably, this peak shifting and peak separation as time progress are indicative of phase separation in this iodide rich system.[33] In the same phase diagram, a composition richer in MA and less in FA and Cs than in **Figure 4(b)**, exhibits similar PL behavior shown in **Figure 4(c)**; however, at the end of 6 hours under ambient conditions, the peaks have not entirely disappeared. While it is interesting to note that such small concentration variation can change the mechanism of degradation, it is out of the scope of this study to explore such a mechanism.

From analyzing the time-dependent PL behavior of quasi-ternary phase systems shown in **Figures 4(d)-(e)**, we observe the effect of mixed halides on the instability of such compositions. Focusing on the composition, $MA_{0.2}FA_{0.72}Cs_{0.08}PbI_{0.6}Br_{2.4}$, in the bromide rich mixed halide system, we find a slight peak shift towards lower wavelengths as the peak intensity gradually increases during the 6 hours of measurements as shown in **Figure 4(d)**. The final composition remains stable. In the same quasi-ternary system, the composition, $MA_{0.44}FA_{0.48}Cs_{0.08}PbI_{1.32}Br_{1.68}$, exhibits different behavior, as shown in **Figure 4(e)**. The PL peak gradually shifts 100 nm towards lower wavelengths without apparent peak splitting, while the peak intensity changes. The final

decomposed composition also remains stable. Finally, for the ternary cation composition of $MA_{0.36}FA_{0.56}Cs_{0.08}PbI_3$, we observe an initial broad PL peak that gradually decreases in intensity, as shown in **Figure 4(f)**. A small peak shift follows this to higher wavelengths as the peak splits into two and finally disappears.

The results illustrated in **Figure 4** suggest that the time dynamics of the PL in the multicomponent system can be very complex, as the result of simultaneous processes of crystallization, defect healing, compositional changes induced by the, e.g., preferential I^- oxidation in ambient, and chemical degradation. Correspondingly, it is difficult and time consuming to study the PL behavior of each composition in a vast compositional space separately. To explore the stability by time dependent-PL behavior across the compositional space systematically, we extend the analysis in Eq. (1) towards the joint analysis of spectral and time domain. Here, we analyze the full $L(x, y, \omega, t)$ dataset using NMF via the decomposition:

$$L(x, y, \omega, t) = \sum_{i=1}^{N} L_i(x, y) g_i(\omega, t) \tag{2}$$

where $L_i(x, y)$ are the loading maps that represent the variability in behaviors across compositional space and $g_i(\omega, t)$ are the endmembers that determine characteristic behaviors in the time-spectral domain. Similar to the analysis of the initial state of PL behavior, the loading maps can be represented via the ternary phase diagram representation after GP interpolation. In contrast, the endmembers can be visualized via classical 2D representations.

14
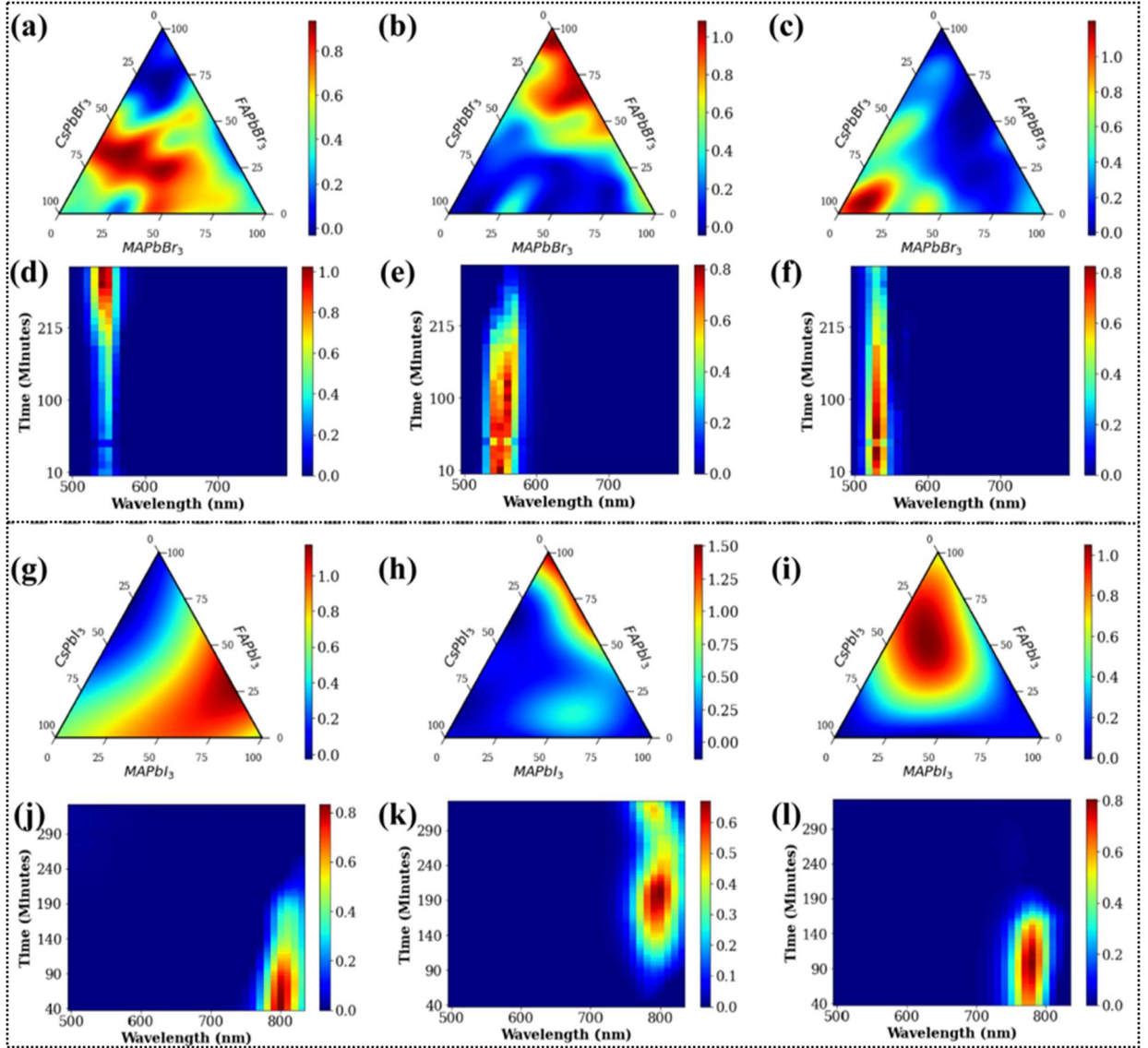
Figure 5. The multivariate analysis of the stability in **(a)-(f)** $MA_xFA_yCs_{1-x-y}PbBr_3$ system and **(g)-(l)** $MA_xFA_yCs_{1-x-y}PbI_3$ system. Shown as ternary diagrams are the compositional dependencies of the component intensities (top rows) and 2D NMF components (bottom rows). The 2D plots describe the PL behavior of the corresponding phase diagram (represented by a non-zero weight) as a function of wavelength and time. For the ternary phase diagrams, the color is indicative of how the characteristic PL behavior over time describes that compositional region. For the 2D plots, the color describes the intensity of the PL peak. The corresponding triangle maps before GP interpolation is found in supplementary information, **Figure S7**.

Employing this strategy, we begin by examining the time evolution of PL properties of an entire bromine ternary phase system, $MA_xFA_yCs_{1-x-y}PbBr_3$. As illustrated in **Figure 5,** this analysis is performed using $N = 3$ in Eq. 2 as it more accurately describes the dynamics of the system,

details can be found in **Figure S6**. Shown in **Figures 5(a)** and **(d)**, the first component demonstrates that solid solutions between $MAPbBr_3$, $FAPbBr_3$, and $CsPbBr_3$ having PL emission wavelength between 540-550 nm experience an increase in PL intensity in the air as time progresses. Increasing PL intensity in ambient conditions has been previously attributed to a boost in the carrier radiative recombination rate.[38] The PL properties of $FAPbBr_3$ rich compositions exhibit different behavior, as shown in the second component in **Figures 5(b)** and **(e)**. Firstly, we observe a broad PL emission between 530 and 570 nm. The PL intensity gradually decreases after approximately 160 minutes, an indication of compositional degradation. Finally, as shown in the third component **Figures 5(c)** and **(f)**, the $CsPbBr_3$ rich compositions exhibit a much narrower PL emission around 530 nm, signifying this compositional space with a higher bandgap emission and possibly higher crystallinity[39]. Although the PL intensity begins to decline, the compositions in this region remain stable under ambient conditions during the study.

Next, we explore the time evolution of PL behavior in the iodide ternary phase system of $MA_xFA_yCs_{1-x-y}PbI_3$. As shown in the first component in **Figures 5(g)** and **(j)**, $MAPbI_3$ rich compositions and a region of solid solutions between $MAPbI_3$ and $CsPbI_3$ experience high PL intensity at wavelengths around 800 nm that begin to decrease over time, indicating full degradation. The second component shown in **Figures 5(h)** and **(k)** details how $FAPbI_3$ and solid solutions between $FAPbI_3$ and $MAPbI_3$ exhibit an increase in PL intensity that eventually decreases as time progresses. Finally, the third component shown in **Figures 5(i)** and **(l)** focuses on solid solutions with FA-rich compositions that show an increase in PL intensity, followed by an abrupt decrease. Overall, this analysis confirms that compositions in the ternary bromide system are more stable than compositions in the ternary iodide system in agreement with the literature[40]. Furthermore, this analysis offers the strategy in cation or anion management, which can successfully guide the design of compositions with higher stability or crystallizations in such ternary systems.

To assess how the GP model fits the experimental data, we have estimated the Bayesian uncertainty model. The uncertainty maps for each three components in **Figure 5 (a-c)** is shown in supplementary information, **Figure S8**. The resultant RMSE values indicate the high quality of the GP prediction.
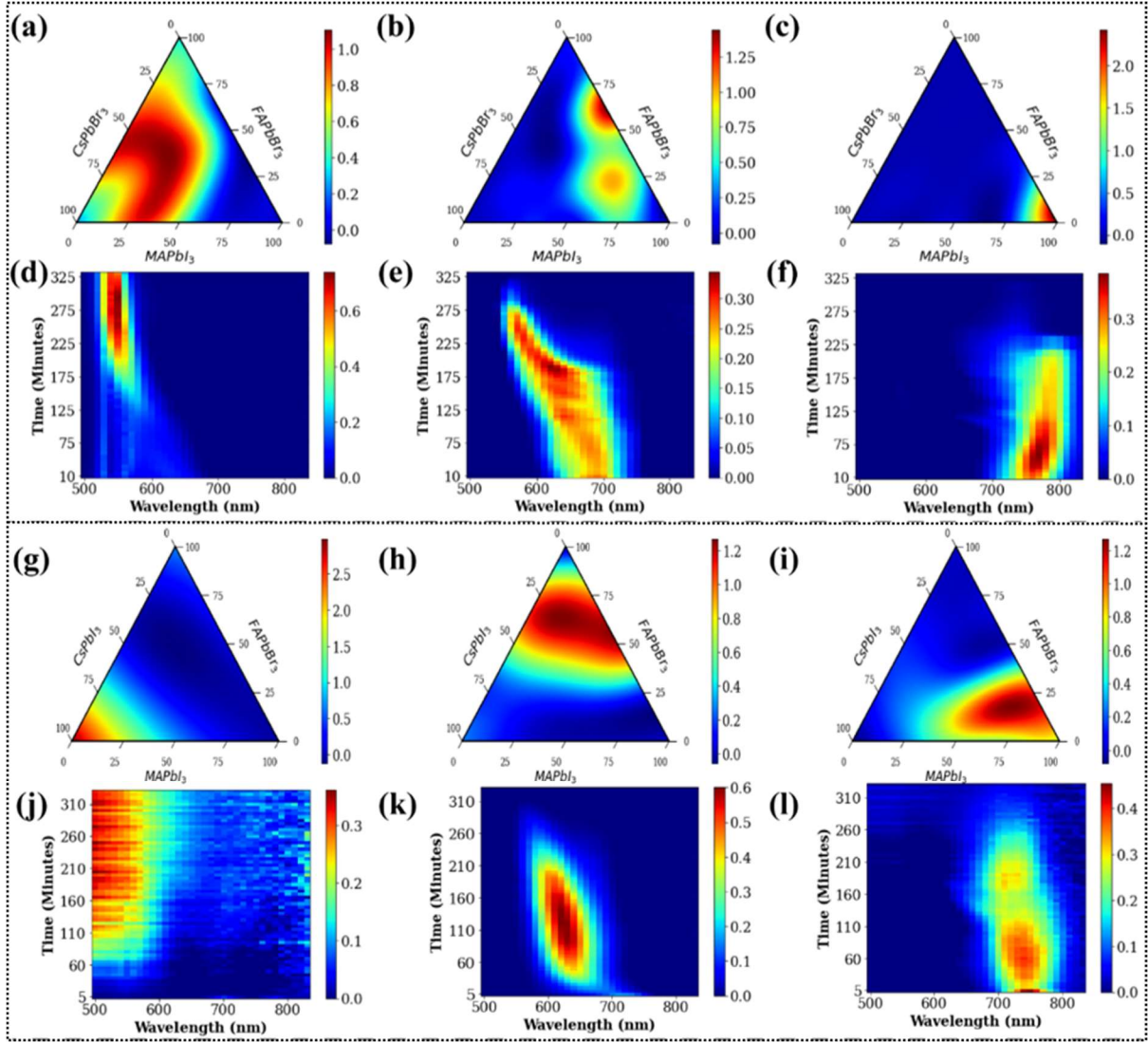
Figure 6. The multivariate analysis of the stability in quasi-ternary **(a)-(f)** $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ and **(g)-(l)** $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$ systems. Shown as ternary diagrams are the compositional dependencies of the component intensities (top rows) and 2D NMF components (bottom rows), describing the PL behavior of the corresponding phase diagram (represented by a non-zero weight) as a function of wavelength and time. The color of the ternary phase diagram is indicative of how the characteristic PL behavior over time describes that compositional region. In the 2D plots, the color describes the intensity of the PL peak. The corresponding triangle maps before GP interpolation is found in supplementary information, **Figure S9**.

We proceed further by exploring the evolution of the PL behavior over time of the two quasi-ternary systems: bromide rich $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ and iodide rich $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$. The first component of the $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ system shown in **Figure 6(a)** and **(d)** demonstrates a shift in the energy bandgap of mixed iodide-bromide solid solutions

17

towards lower wavelengths with higher intensity of the PL emission as time progresses. We also observe a similar PL enhancement as exhibited in the $MA_xFA_yCs_{1-x-y}PbBr_3$ earlier with a final stable composition. We notice for a selected region with a solid solution between $MAPbI_3$ and $FAPbBr_3$ shown in **Figure 6(b)**, the 2D NMF components in **Figure 6(e)** shows a significant blue shift towards lower wavelengths; a clear indication of phase separation as the time progresses which ultimately degrades. The third component in **Figure 6(c)** and **(f)** highlights a small region rich in $MAPbI_3$ with emission wavelength around 760 nm. The PL emission slightly redshifts and finally disappears. An excellent example of this is the quasi-ternary system, $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$, shown in **Figures 6 (a-f)**. We conclude that in this system, $CsPbBr_3$- and $FAPbBr_3$-rich regions are more stable because not only the PL wavelength does not change in ambient as a function of time but also the PL intensity increases as their time in ambient conditions increases. We can further confirm this by looking at an individual composition, $MA_{0.20}FA_{0.72}Cs_{0.08}PbI_{0.60}Br_{2.40}$, in this region that is represented in **Figure 4(d).** On the other hand, the remainder of the phase diagram experiences peak shifts along with a decrease in photoluminescent intensity, indicating that these compositions are less stable in ambient conditions. Furthermore, we can use the analysis demonstrated in **Figure 3** as a material-property guide for exploring different optoelectronic applications.

For the $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$ system, a small $CsPbI_3$-rich region can be seen in the first component in **Figure 6 (g)** and **(j)**. This region of the phase diagram shows a broad PL emission between 500-600 nm, indicative of the degradation into PbI_2 of these microcrystals. We conclude this from the yellow precipitates that formed along with the PL spectrum seen in **Figure S10**. As shown in **Figure 6(j)**, this broad peak is indicative of PbI_2. We observe the effect of halide segregation in solid solutions of $MAPbI_3$ and $FAPbBr_3$ as a function of time in the second component, as shown in **Figure 6(h)** and **(k)**. As can be seen in the 2D NMF component, the PL emission shifts toward lower wavelengths. For both the second and third components shown in **Figures 6(h)** and **(i)**, we see the degradation of iodide-rich compositions as time progresses. The third components in **Figure (i)** and **(l)** represent a solid solution between $CsPbI_3$ and $MAPbI_3$. The analysis of quasi-ternary systems demonstrates that most solid solutions in mixed halides experience compositional changes induced by possibly preferential I^- oxidation in ambient, phase separation towards more stable components and ultimately experience chemical degradation.

This study demonstrated how robotics-assisted materials synthesis, combined with machine learning techniques, can explore vast compositional spaces in HOIPs. We accomplished this by first utilizing a pipetting robot and a Multi-Mode well plate reader to automate the synthesis and characterization of a large number of compositions. We developed a workflow to explore the PL properties and stability in ambient conditions and applied it to four perovskite materials systems: $MA_xFA_yCs_{1-x-y}PbBr_3$, $MA_xFA_yCs_{1-x-y}PbI_3$, $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ and $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$. In total, 368 compositions were synthesized, and their time-dependent PL behavior was explored within 4 days. It is well established that monitoring PL properties is an effective strategy for obtaining information on the optoelectronic quality of HOIPs[41]. However, PL properties are difficult to predict in a sizeable compositional space. Here, they were efficiently mapped as the entire optical characterization of a 96-well plate takes approximately five minutes for each system. The results allowed us to establish a composition-properties relationship for a large number of compositions across a specific phase diagram. In combination with machine learning techniques, specifically NMF and Gaussian Process, we gained insight into the variability of the optical bandgap and PL properties across the compositional series. Overall, this workflow is an exploratory process; however, combined with our knowledge and the knowledge in the literature, we are able to draw conclusions about the characteristics of our chosen perovskite systems. The interpolative regression analysis of PL properties helps to distinguish mixtures that form solid solutions from those that segregate into multiple materials in ambient conditions. The trend in PL emission wavelengths, and thus perovskite bandgaps, is found to be near-linear along some compositional axes.

This experimental information is a necessary step in material's design for specific applications. Clearly, the stability dependence on composition is extremely non-uniform within the composition space – with gradual evolution within solid solution spaces and rapid changes at the phase boundaries. The relationship of the degradation rates on the composition provides information on the kinetic mechanisms – e.g., transport or reaction limited kinetics. The observed results here can be further fed into the ML model to allow the physics-based generalization[42], and hence increase the efficiency of the evolutionary search. This exploration necessitates the knowledge of relevant HOIPs physics and incorporation of these formulating the future search strategies. Notably, these sets of problems are virtually unexplored in the theory-based ML approaches since physical priors are usually very poorly known. In particular, stabilities are

extremely difficult to predict theoretically and rationalize, necessitating a combined approach of experiment, theory, and theory-informed intuition.

Conclusion:

Commonly, significant scientific progress is slowed by the utilization of Edisonian trial and error approaches, where one variable is changed at a time and result is then measured. Automated combinatorial synthesis provides the opportunity of accelerating the production of materials with large compositional spaces. In this article, we further advance this field of study by demonstrating a proof of concept for the development of an automated workflow that incorporates robotics synthesis and machine learning to explore the changes in photoluminescent behavior of vast compositional spaces in HOIPs. With this workflow, we synthesized and characterized 368 different microcrystals within 4 days. It was demonstrated that a combination of laboratory automation and machine learning can be used for the rapid mapping of both the concentration-dependent physical properties and long-term stability in broad concentration spaces of hybrid organic-inorganic perovskites, verifying the presence of the regions with high stability. This workflow is a necessary step to guide the material synthesis and to gain an insight into the overall behavior of the complex phase diagram of mixed perovskites. Therefore, this approach opens the pathway for the discovery of new HOIPs with optimized properties for multifunctional optoelectronics.

References

(1)	Park, N.-G.; Grätzel, M.; Miyasaka, T.; Zhu, K.; Emery, K. Towards stable and commercially available perovskite solar cells. *Nature Energy* **2016**, *1*, 16152.

(2)	Zhao, B.; Bai, S.; Kim, V.; Lamboll, R.; Shivanna, R.; Auras, F.; Richter, J. M.; Yang, L.; Dai, L.; Alsari, M.; She, X.-J.; Liang, L.; Zhang, J.; Lilliu, S.; Gao, P.; Snaith, H. J.; Wang, J.; Greenham, N. C.; Friend, R. H.; Di, D. High-efficiency perovskite–polymer bulk heterostructure light-emitting diodes. *Nature Photonics* **2018**, *12*, 783-789.

(3)	Zou, Y.; Yuan, Z.; Bai, S.; Gao, F.; Sun, B. Recent progress toward perovskite light-emitting diodes with enhanced spectral and operational stability. *Materials Today Nano* **2019**, *5*, 100028.

(4)	Ahmadi, M.; Wu, T.; Hu, B. A Review on Organic–Inorganic Halide Perovskite Photodetectors: Device Engineering and Fundamental Physics. **2017**, *29*, 1605242.

(5)	Lukosi, E.; Smith, T.; Tisdale, J.; Hamm, D.; Seal, C.; Hu, B.; Ahmadi, M. Methylammonium lead tribromide semiconductors: Ionizing radiation detection and electronic properties. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **2019**, *927*, 401-406.

(6)	Shrestha, S.; Fischer, R.; Matt, G. J.; Feldner, P.; Michel, T.; Osvet, A.; Levchuk, I.; Merle, B.; Golkar, S.; Chen, H.; Tedde, S. F.; Schmidt, O.; Hock, R.; Rührig, M.; Göken, M.; Heiss, W.; Anton, G.; Brabec, C. J. High-performance direct conversion X-ray detectors based on sintered hybrid lead triiodide perovskite wafers. *Nature Photonics* **2017**, *11*, 436.

(7)	Jain, A.; Voznyy, O.; Sargent, E. H. High-Throughput Screening of Lead-Free Perovskite-like Materials for Optoelectronic Applications. *The Journal of Physical Chemistry C* **2017**, *121*, 7183-7187.

(8)	Nakajima, T.; Sawada, K. Discovery of Pb-Free Perovskite Solar Cells via High-Throughput Simulation on the K Computer. *The Journal of Physical Chemistry Letters* **2017**, *8*, 4826-4831.

(9)	Zhao, X.-G.; Yang, D.; Ren, J.-C.; Sun, Y.; Xiao, Z.; Zhang, L. Rational Design of Halide Double Perovskites for Optoelectronic Applications. *Joule* **2018**, *2*, 1662-1673.

(10)	Saliba, M. Polyelemental, Multicomponent Perovskite Semiconductor Libraries through Combinatorial Screening. *Advanced Energy Materials* **2019**, *9*, 1803754.

(11)	Liu, W.-W.; Wu, T.-H.; Liu, M.-C.; Niu, W.-J.; Chueh, Y.-L. Recent Challenges in Perovskite Solar Cells Toward Enhanced Stability, Less Toxicity, and Large-Area Mass Production. *Advanced Materials Interfaces* **2019**, *6*, 1801758.

(12)	Salhi, B.; Wudil, Y. S.; Hossain, M. K.; Al-Ahmed, A.; Al-Sulaiman, F. A. Review of recent developments and persistent challenges in stability of perovskite solar cells. *Renewable and Sustainable Energy Reviews* **2018**, *90*, 210-222.

(13)	Ono, L. K.; Juarez-Perez, E. J.; Qi, Y. Progress on Perovskite Materials and Solar Cells with Mixed Cations and Halide Anions. *ACS Applied Materials & Interfaces* **2017**, *9*, 30197-30246.

(14)	Wang, Z.; Shi, Z.; Li, T.; Chen, Y.; Huang, W. Stability of Perovskite Solar Cells: A Prospective on the Substitution of the A Cation and X Anion. **2017**, *56*, 1190-1212.

(15)	Xu, F.; Zhang, T.; Li, G.; Zhao, Y. Mixed cation hybrid lead halide perovskites with enhanced performance and stability. *Journal of Materials Chemistry A* **2017**, *5*, 11450-11461.

(16)	Chen, S.; Hou, Y.; Chen, H.; Tang, X.; Langner, S.; Li, N.; Stubhan, T.; Levchuk, I.; Gu, E.; Osvet, A.; Brabec, C. J. Exploring the Stability of Novel Wide Bandgap Perovskites by a Robot Based High Throughput Approach. *Advanced Energy Materials* **2018**, *8*, 1701543.

(17)	Jesper Jacobsson, T.; Correa-Baena, J.-P.; Pazoki, M.; Saliba, M.; Schenk, K.; Grätzel, M.; Hagfeldt, A. Exploration of the compositional space for mixed lead halogen perovskites for high efficiency solar cells. *Energy & Environmental Science* **2016**, *9*, 1706-1724.

(18)	Sun, S.; Hartono, N. T. P.; Ren, Z. D.; Oviedo, F.; Buscemi, A. M.; Layurova, M.; Chen, D. X.; Ogunfunmi, T.; Thapa, J.; Ramasamy, S.; Settens, C.; DeCost, B. L.; Kusne, A. G.; Liu, Z.; Tian, S. I. P.; Peters, I. M.; Correa-Baena, J.-P.; Buonassisi, T. Accelerated Development of Perovskite-Inspired Materials via High-Throughput Synthesis and Machine-Learning Diagnosis. *Joule* **2019**, *3*, 1437-1451.

(19)   Epps, R. W.; Felton, K. C.; Coley, C. W.; Abolhasani, M. Automated microfluidic platform for systematic studies of colloidal perovskite nanocrystals: towards continuous nano-manufacturing. *Lab on a Chip* **2017**, *17*, 4040-4047.

(20)   Epps, R. W.; Bowen, M. S.; Volk, A. A.; Abdel-Latif, K.; Han, S.; Reyes, K. G.; Amassian, A.; Abolhasani, M. Artificial Chemist: An Autonomous Quantum Dot Synthesis Bot. *Advanced Materials* **2020**, *32*, 2001626.

(21)   Gu, E.; Tang, X.; Langner, S.; Duchstein, P.; Zhao, Y.; Levchuk, I.; Kalancha, V.; Stubhan, T.; Hauch, J.; Egelhaaf, H. J.; Zahn, D.; Osvet, A.; Brabec, C. J. Robot-Based High-Throughput Screening of Antisolvents for Lead Halide Perovskites. *Joule* **2020**.

(22)   MacLeod, B. P.; Parlane, F. G. L.; Morrissey, T. D.; Häse, F.; Roch, L. M.; Dettelbach, K. E.; Moreira, R.; Yunker, L. P. E.; Rooney, M. B.; Deeth, J. R.; Lai, V.; Ng, G. J.; Situ, H.; Zhang, R. H.; Elliott, M. S.; Haley, T. H.; Dvorak, D. J.; Aspuru-Guzik, A.; Hein, J. E.; Berlinguette, C. P. Self-driving laboratory for accelerated discovery of thin-film materials. *Science Advances* **2020**, *6*, eaaz8867.

(23)   Gopalaswamy, V.; Betti, R.; Knauer, J. P.; Luciani, N.; Patel, D.; Woo, K. M.; Bose, A.; Igumenshchev, I. V.; Campbell, E. M.; Anderson, K. S.; Bauer, K. A.; Bonino, M. J.; Cao, D.; Christopherson, A. R.; Collins, G. W.; Collins, T. J. B.; Davies, J. R.; Delettrez, J. A.; Edgell, D. H.; Epstein, R.; Forrest, C. J.; Froula, D. H.; Glebov, V. Y.; Goncharov, V. N.; Harding, D. R.; Hu, S. X.; Jacobs-Perkins, D. W.; Janezic, R. T.; Kelly, J. H.; Mannion, O. M.; Maximov, A.; Marshall, F. J.; Michel, D. T.; Miller, S.; Morse, S. F. B.; Palastro, J.; Peebles, J.; Radha, P. B.; Regan, S. P.; Sampat, S.; Sangster, T. C.; Sefkow, A. B.; Seka, W.; Shah, R. C.; Shmyada, W. T.; Shvydky, A.; Stoeckl, C.; Solodov, A. A.; Theobald, W.; Zuegel, J. D.; Johnson, M. G.; Petrasso, R. D.; Li, C. K.; Frenje, J. A. Tripled yield in direct-drive laser fusion through statistical modelling. *Nature* **2019**, *565*, 581-586.

(24)   Milo, A.; Bess, E. N.; Sigman, M. S. Interrogating selectivity in catalysis using molecular vibrations. *Nature* **2014**, *507*, 210.

(25)   King, R. D.; Whelan, K. E.; Jones, F. M.; Reiser, P. G. K.; Bryant, C. H.; Muggleton, S. H.; Kell, D. B.; Oliver, S. G. Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* **2004**, *427*, 247-252.

(26)   Aspuru-Guzik, A.; Persson, K.: *Materials Acceleration Platform: Accelerating Advanced Energy Materials Discovery by Integrating High-Throughput Methods and Artificial Intelligence*, 2018.

(27)   Nikolaev, P.; Hooper, D.; Perea-López, N.; Terrones, M.; Maruyama, B. Discovery of Wall-Selective Carbon Nanotube Growth Conditions via Automated Experimentation. *ACS Nano* **2014**, *8*, 10214-10222.

(28)   Ren, F.; Ward, L.; Williams, T.; Laws, K. J.; Wolverton, C.; Hattrick-Simpers, J.; Mehta, A. Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments. **2018**, *4*, eaaq1566.

(29)   Li, J.; Lu, Y.; Xu, Y.; Liu, C.; Tu, Y.; Ye, S.; Liu, H.; Xie, Y.; Qian, H.; Zhu, X. AIR-Chem: Authentic Intelligent Robotics for Chemistry. *The Journal of Physical Chemistry A* **2018**, *122*, 9142-9148.

(30)   Aldair E. Gongora, B. X., Wyatt Perry, Chika Okoye, Patrick Riley, Kristofer G. Reyes, Elise F. Morgan, Keith A. Brown. A Bayesian Experimental Autonomous Researcher for Mechanical Design. *Sci Adv* **2020**, *6*.

(31)   McGinn, P. J. Thin-Film Processing Routes for Combinatorial Materials Investigations—A Review. *ACS Combinatorial Science* **2019**, *21*, 501-515.

(32)   Caramelli, D.; Salley, D.; Henson, A.; Camarasa, G. A.; Sharabi, S.; Keenan, G.; Cronin, L. Networking chemical robots for reaction multitasking. *Nature Communications* **2018**, *9*, 3406.

(33)   Knight, A. J.; Wright, A. D.; Patel, J. B.; McMeekin, D. P.; Snaith, H. J.; Johnston, M. B.; Herz, L. M. Electronic Traps and Phase Segregation in Lead Mixed-Halide Perovskite. *ACS Energy Letters* **2019**, *4*, 75-84.

(34)   Babbe, F.; Sutter-Fella, C. M. Optical Absorption-Based In Situ Characterization of Halide Perovskites. *Advanced Materials*, *n/a*, 1903587.

(35)   Ruan, S.; Surmiak, M.-A.; Ruan, Y.; McMeekin, D. P.; Ebendorff-Heidepriem, H.; Cheng, Y.-B.; Lu, J.; McNeill, C. R. Light induced degradation in mixed-halide perovskites. *Journal of Materials Chemistry C* **2019**, *7*, 9326-9334.

(36)   Rasmussen, C. E.; Williams, C. K. I.: *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*; The MIT Press, 2005.

(37)   Williams, C. K. I.: Prediction with Gaussian Processes: From Linear Regression to Linear Prediction and Beyond. In *Learning in Graphical Models*; Jordan, M. I., Ed.; Springer Netherlands: Dordrecht, 1998; pp 599-621.

(38)   Howard, J. M.; Tennyson, E. M.; Barik, S.; Szostak, R.; Waks, E.; Toney, M. F.; Nogueira, A. F.; Neves, B. R. A.; Leite, M. S. Humidity-Induced Photoluminescence Hysteresis in Variable Cs/Br Ratio Hybrid Perovskites. *J Phys Chem Lett* **2018**, *9*, 3463-3469.

(39)   Si, H.; Zhang, Z.; Liao, Q.; Zhang, G.; Ou, Y.; Zhang, S.; Wu, H.; Wu, J.; Kang, Z.; Zhang, Y. A-Site Management for Highly Crystalline Perovskites. *Advanced Science* **2020**, *32*, 1904702.

(40)   Noh, J. H.; Im, S. H.; Heo, J. H.; Mandal, T. N.; Seok, S. I. Chemical management for colorful, efficient, and stable inorganic-organic hybrid nanostructured solar cells. *Nano Lett* **2013**, *13*, 1764-1769.

(41)   Motti, S. G.; Meggiolaro, D.; Barker, A. J.; Mosconi, E.; Perini, C. A. R.; Ball, J. M.; Gandini, M.; Kim, M.; De Angelis, F.; Petrozza, A. Controlling competing photochemical reactions stabilizes perovskite solar cells. *Nature Photonics* **2019**, *13*, 532-539.

(42)   Ling, J.; Antono, E.; Bajaj, S.; Paradiso, S.; Hutchinson, M.; Meredig, B.; Gibbons, B. M. Machine Learning for Alloy Composition and Process Optimization. **2018**, V006T024A005.

Supplementary Materials

Chemical Robotics Enabled Exploration of Stability and Photoluminescent Behavior in Multicomponent Hybrid Perovskites via Machine Learning

Kate Higgins[1], Sai Mani Valleti[2], Maxim Ziatdinov[3],

Sergei V. Kalinin*[2,3], and Mahshid Ahmadi[1]*

[1] Joint Institute for Advanced Materials, Department of Materials Science and Engineering, University of Tennessee, Knoxville, TN 37996, USA

[2] The Bredesen Center, University of Tennessee, Knoxville, TN 37996, USA

[3] The Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

*Corresponding authors email:

mahmadi3@utk.edu, sergei2@ornl.gov

Materials and methods

*Synthesis of Precursor Solutions*

The $MAPbBr_3$ 0.3 M precursor solution was made by dissolving 0.0672 g of MABr and 0.2202 g of $PbBr_2$ in 2 mL of *n'n*-dimethylformamide (DMF). The $MAPbI_3$ 0.3 M precursor solution was made by dissolving 0.2861 g of MAI and 0.8298 g of PbI_2 in 6 mL of γ-butyrolactone (GBL). The $FAPbBr_3$ 0.3 M precursor solution was made by dissolving 0.2249 g of FABr and 0.6606 g of $PbBr_2$ in 6 mL of *n'n*-dimethylformamide (DMF). The $FAPbI_3$ 0.3 M precursor solution was made by dissolving 0.1032 g of FAI and 0.2766 g of PbI_2 in 6 mL of γ-butyrolactone (GBL). The $CsPbBr_3$ 0.3 M solution was made by dissolving 0.2554 g of CsBr and 0.4404 g of $PbBr_2$ in 4 mL of dimethyl sulfoxide (DMSO). The $CsPbI_3$ 0.3 M solution was made by dissolving 0.3118 g CsI and 0.5532 g of PbI_2 in 4 mL of *n'n*-dimethylformamide (DMF). Each solution was stirred at 500 rpm for 1 hour until the precursors were completely dissolved. The vial was covered with aluminum foil to protect it from possible light effects. All steps mentioned were performed under

1 a nitrogen atmosphere. Approximately after one hour, the precursor solutions were used for the

2 synthesis of microcrystals.

3 *Automated Synthesis of Microcrystals*

4 Automated synthesis of microcrystals was performed using Opentrons OT-2 pipetting robot. The

5 protocol was written using their python API. In total, four microplates were synthesized and used

6 for photoluminescence measurements. The exact volumes of the precursors deposited in each well

7 plate is shown in Table 1. Once each of the precursors was deposited in their respective wells, 250

8 μL of chloroform was added. Immediately after the addition of antisolvent, the microcrystals

9 precipitated, and the solutions changed color indicating the presence of microcrystals.

10
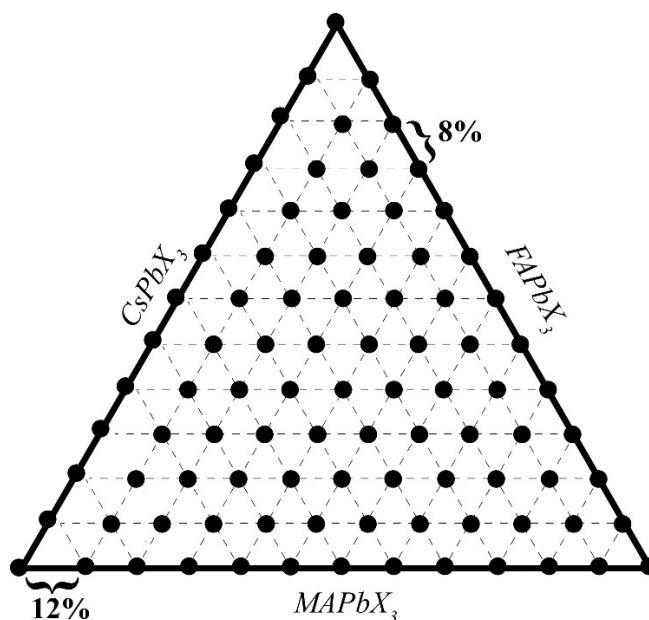
11 **Figure S1.** Schematic of the sample point distribution in the phase diagram for each microplate.

12 We sampled compositions roughly equally spaced, approximately 8 *at* % apart, (except the initial

13 data points from the triangle side which are 12 *at* % apart) to characterize the entirety of the phase

14 diagram.

15

16

17

18

Table 1. Volumes of precursor solution in each well of the microplate. Precursor 1 is shown in black, precursor 2 in red, and precursor 3 in blue.

|	1	2	3	4	5	6	7	8	9	10	11	12
A	50 0 0	38 4 8	30 16 4	26 12 12	22 12 16	18 16 16	14 24 12	10 36 4	10 4 36	6 20 24	2 40 8	2 8 40
B	46 4 0	38 0 12	30 12 8	26 8 16	22 8 20	18 12 20	14 20 16	10 32 8	10 0 40	6 16 28	2 36 12	2 4 44
C	46 0 4	34 16 0	30 8 12	26 4 20	22 4 24	18 8 24	14 16 20	10 28 12	6 44 0	6 12 32	2 32 16	2 0 48
D	42 8 0	34 12 4	30 4 16	26 0 24	22 0 28	18 4 28	14 12 24	10 24 16	6 40 4	6 8 36	2 28 20	
E	42 4 4	34 8 8	30 0 20	22 28 0	18 32 0	18 0 32	14 8 28	10 20 20	6 36 8	6 4 40	2 24 24	
F	42 0 8	34 4 12	26 24 0	22 24 4	18 28 4	14 36 0	14 4 32	10 16 24	6 32 12	6 0 44	2 20 28	
G	38 12 0	34 0 16	26 20 4	22 20 8	18 24 8	14 32 4	14 0 36	10 12 28	6 28 16	2 48 0	2 16 32	
H	38 8 4	30 20 0	26 16 8	22 16 12	18 20 12	14 28 8	10 40 0	10 8 32	6 24 20	2 44 4	2 12 36	
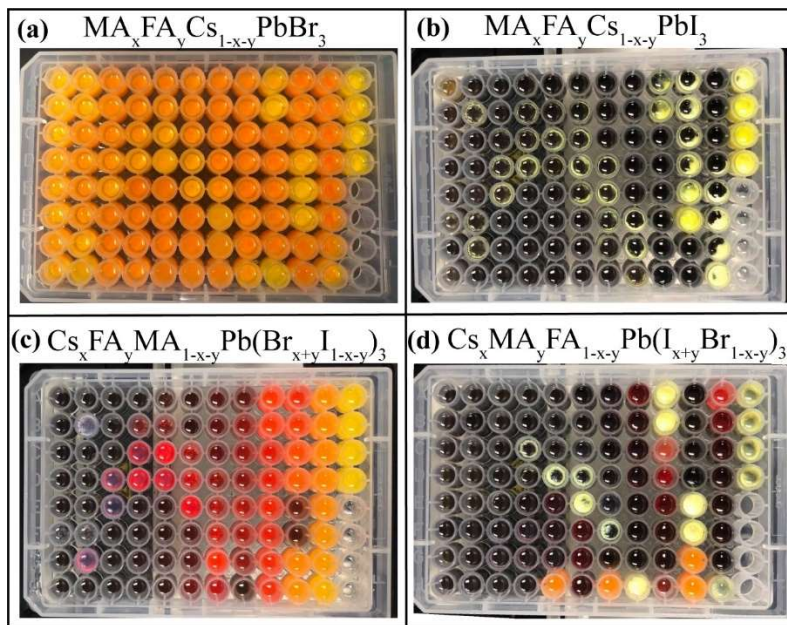
Figure S2. Photos of the four microplates synthesized in this study. Compositions of these plates are (a) $MA_xFA_yCs_{1-x-y}PbBr_3$, (b) $MA_xFA_yCs_{1-x-y}PbI_3$, (c) $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ and (d) $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$. The precipitated perovskite microcrystals can be clearly seen from the colors. In the plates containing iodide, we observe yellow precipitates, indicating the presence of PbI_2, incomplete crystallization or non-perovskites phases.

Photoluminescence Measurements

BioTek Cytation Hybrid Multi-Mode Reader popular in biology and cell science is used for photoluminescence spectroscopy. In this equipment the data collection is designed through Gen 5^{TM} software with capability of data processing for most complex arrays. The reader features monochromators with variable bandwidths from 10-50 nm to optimize detection conditions, with unique excitation or emission parameters and filter-based fluorescence optics for high transmission and sensitivity. The monochromator can excite from both the top and the bottom of the microplate. Photoluminescence measurements were performed with an excitation wavelength of 450 nm and measured over the range of 500 to 799 nm with steps of 10 nm. It was measured from the bottom, approximately 7 mm below the well plate. All measurements were performed in a sweep mode. Photoluminescence measurements were performed every 10 minutes for approximately 6 hours.
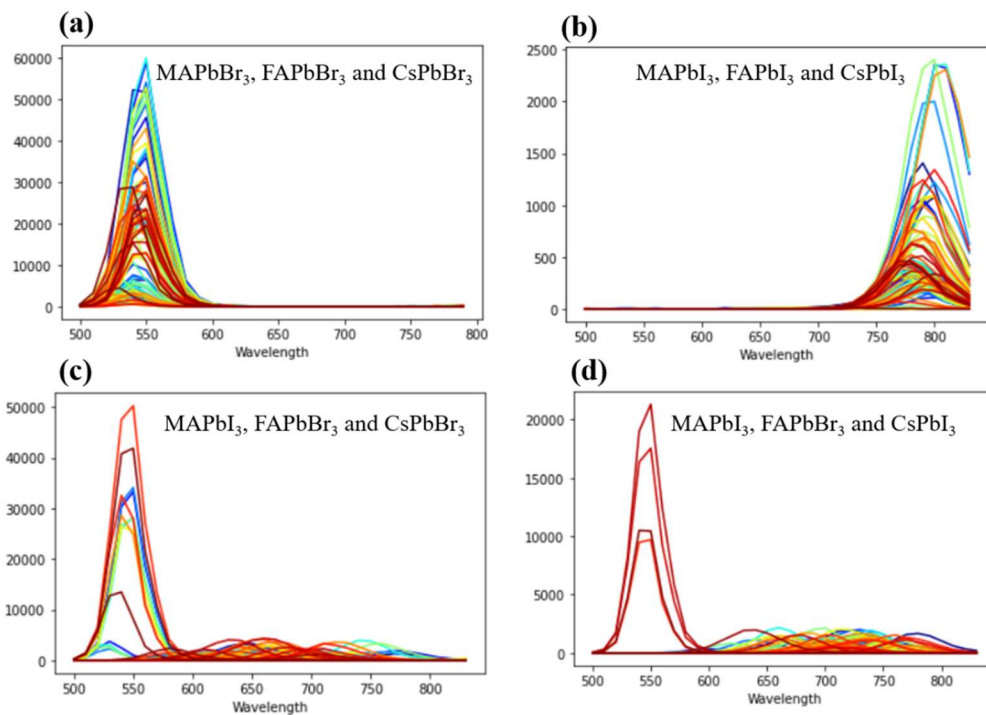
Figure S3. Characteristic PL properties of 4 different plates at time zero in **(a)** $Cs_xFA_yMA_{1-x-y}PbBr_3$, **(b)** $Cs_xFA_yMA_{1-x-y}PbI_3$, **(c)** $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ and **(d)** $Cs_xMA_yFA_{1-x-y}Pb(I_{x+y}Br_{1-x-y})_3$ systems. The color change in each panel represents different concentration of precursors.
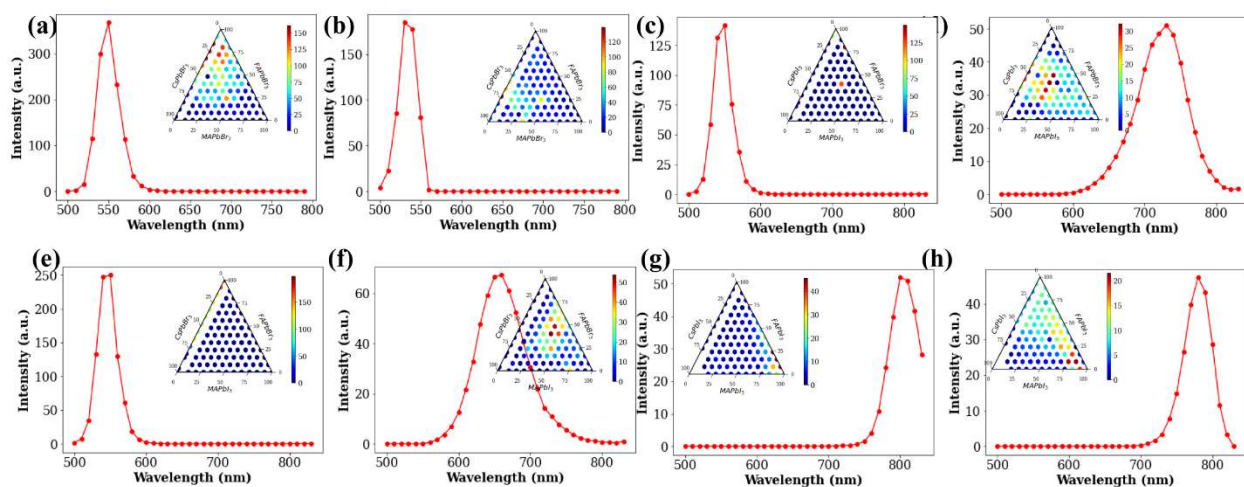
**Figure S4.** The complementary maps to Figure 2. The triangle maps demonstrate the analysis of PL responses prior to GP interpolation.
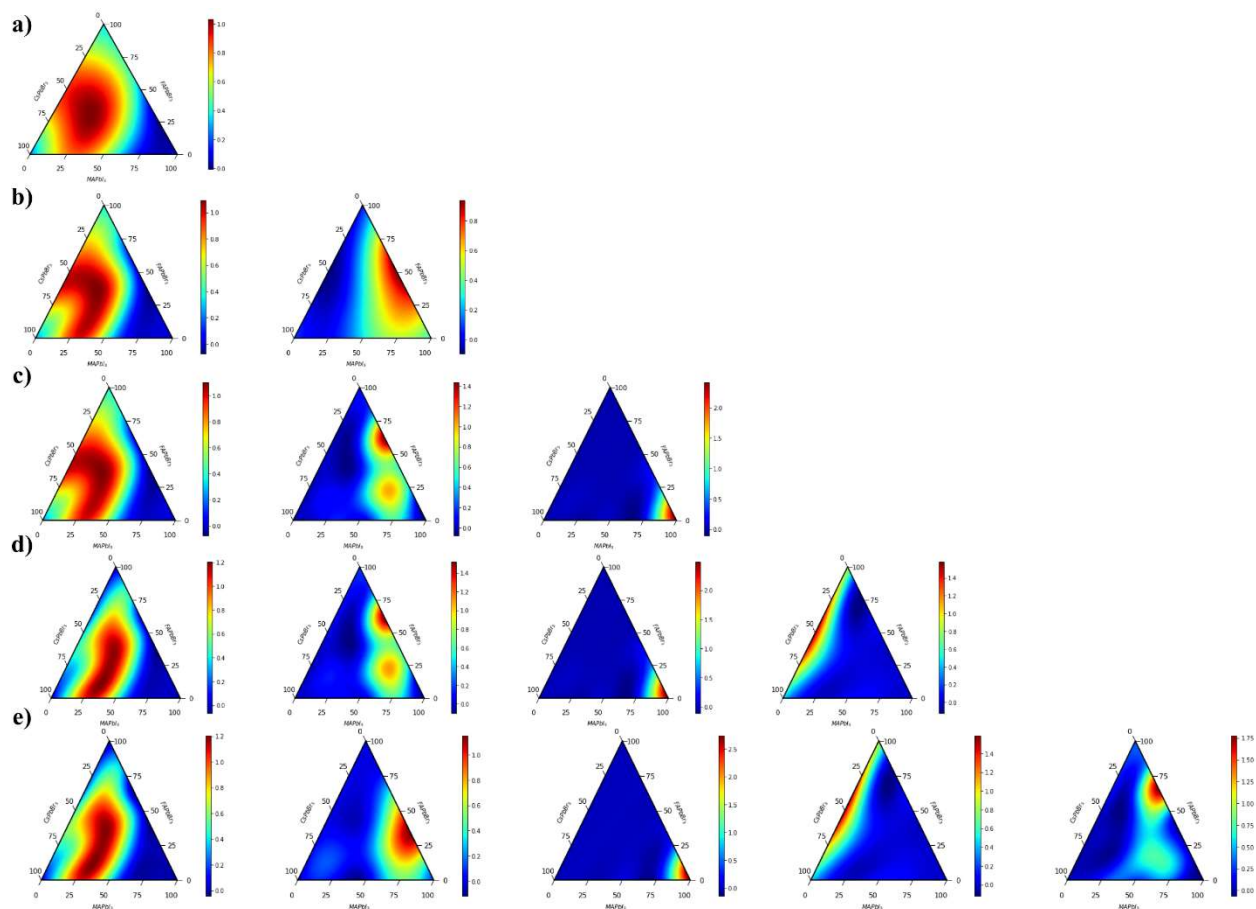
Determining the correct number of components, N, for Non-negative Matrix Factorization (NMF):

Overall, the operator chooses the number of components (N) based on the knowledge of the system. In this study, we chose the number of components subjectively by visually inspecting the loading maps as we increase the number of components. However, we can estimate the "optimal" number of components analytically by checking the error between NMF and the actual matrix. We then calculate the error difference, δ_1, by subtracting the error for $N + 1$ components by N components. We do this for $N + k$ components, building a list of delta values. Once δ_1 is not changing substantially, we assume that the approximation is not improving by adding more components.

We have included exemplary figures that help demonstrate how we chose a certain number of components for both **Figures 2**, **5**, and **6**. As an example, we compared the number of components for the model system: $MAPbI_3$, $FAPbBr_3$ and $CsPbBr_3$ at $t = 0$. We observe similar components once we increase the number of components. For example, for $N = 3$, the first component from $N = 2$ has been split into the first and third components. To determine the number of components used for our stability measurements, we followed a similar approach. For the visual determination, we noticed that after three components, the components began to look quite similar; therefore, we decided to use three components for this part of the experiment. This observation is confirmed by our analytical approach.
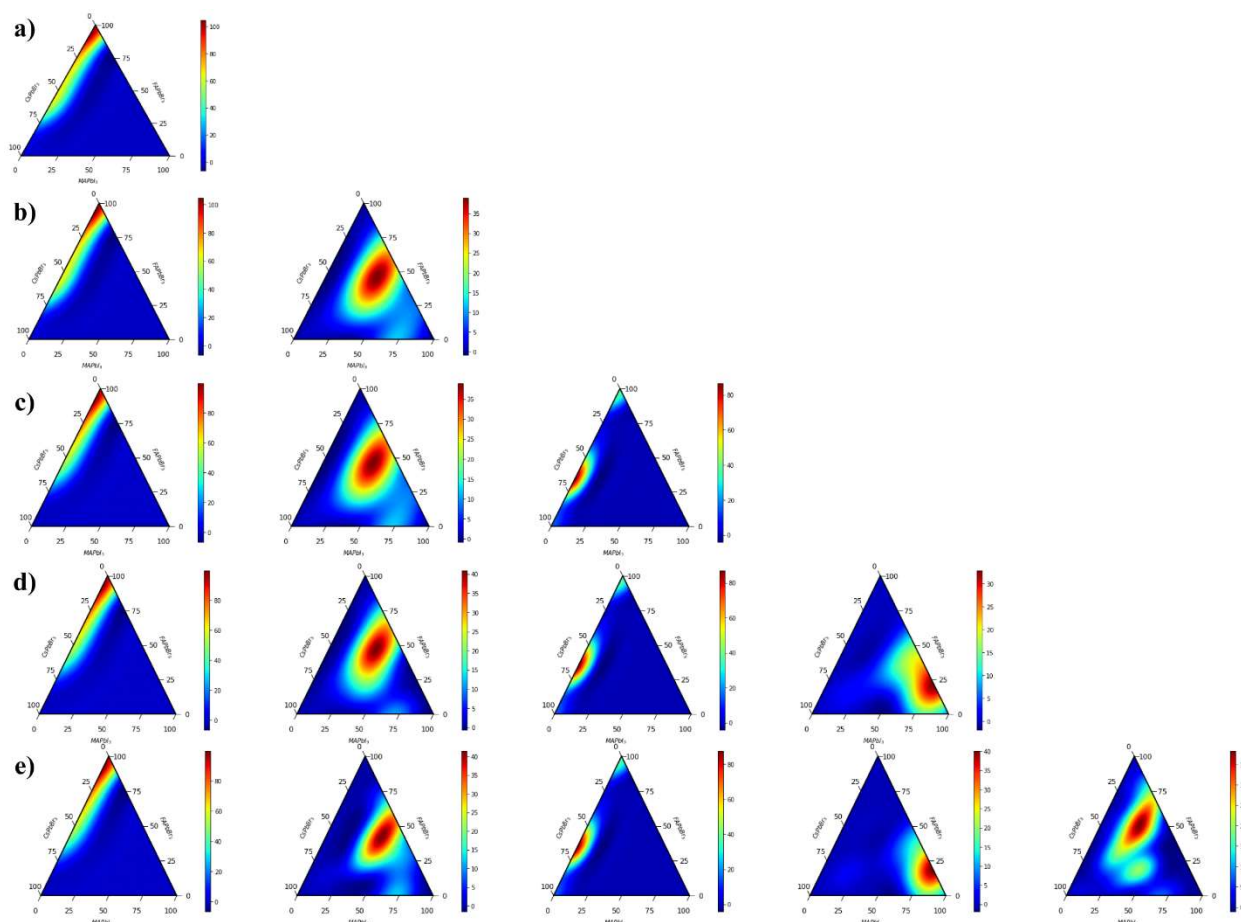
Figure S5. Visual comparison for determining the correct number of components for Non-negative Matrix Factorization (NMF) for the $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ quasi-ternary system at time zero in Figure 2.

Figure S6. Visual comparison for determining the correct number of components for Non-negative Matrix Factorization (NMF) for the $Cs_xFA_yMA_{1-x-y}Pb(Br_{x+y}I_{1-x-y})_3$ quasi-ternary system in Figure 6 as it was exposed to ambient conditions.
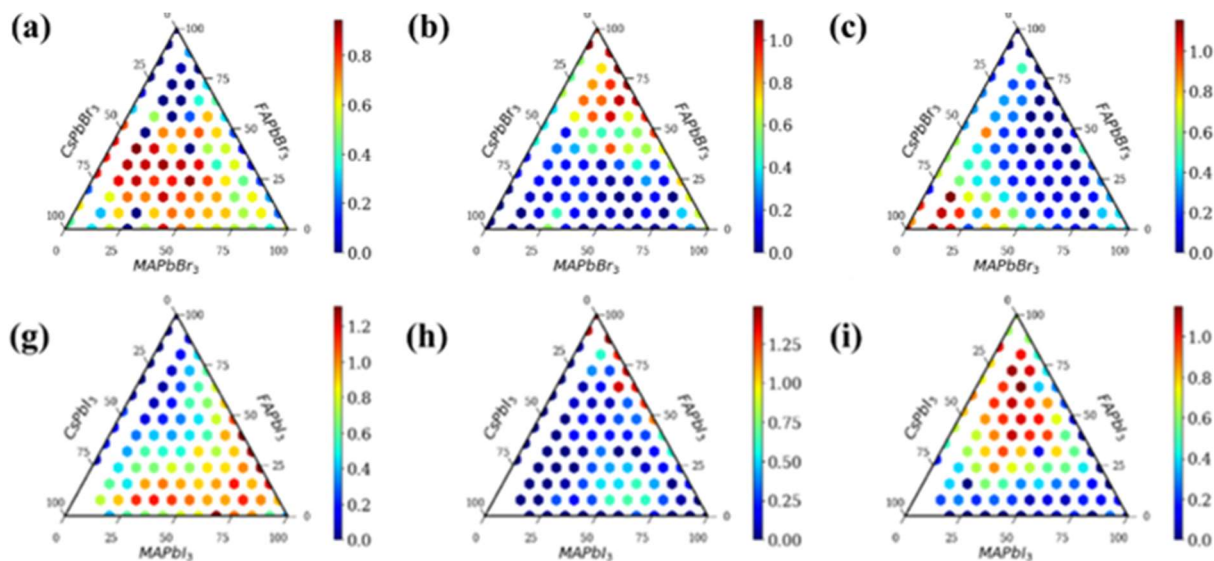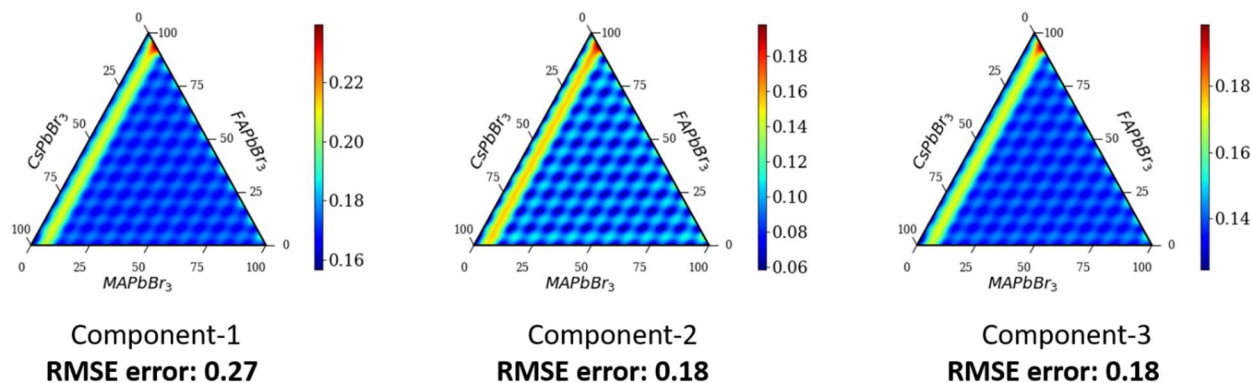
Figure S7. The corresponding triangle maps in Figure 4 showing the PL analysis performed without the GP interpolation.

The assessment of the Gaussian process (GP) model with the experimental data

The Gaussian process (GP) belongs to the category of 'lazy' machine learning algorithms, for which, there are no pure "train" and "test" phases (unlike for many other mainstream ML algorithms such as deep learning). The reason for using "train"/"test" or "fit"/"predict" terminology is simply consistency from a coding standpoint. The quality of GP prediction or, more precisely, model's (un)certainty, is usually accessed via Bayesian uncertainty estimates. The uncertainty maps for each three components is included here. We have also performed a classical model validation. Here, 80% of the labelled data is used for extrapolating the entire phase diagram. Then, the values corresponding to the other 20% of the labelled data are compared with GP predictions in those points, **Figure S8**. The resultant RMSE values are shown, indicating of high quality of the prediction.

Component-1
RMSE error: 0.27

Component-2
RMSE error: 0.18

Component-3
RMSE error: 0.18

Figure S8: The estimated Bayesian uncertainty for the assessment of GP model for the three components in **Figure 5 (a-c)**. The resultant root mean squared error (RMSE) calculated for each component is also shown. The line with large uncertainties is related to large spaing between measurement points in the parameter space.
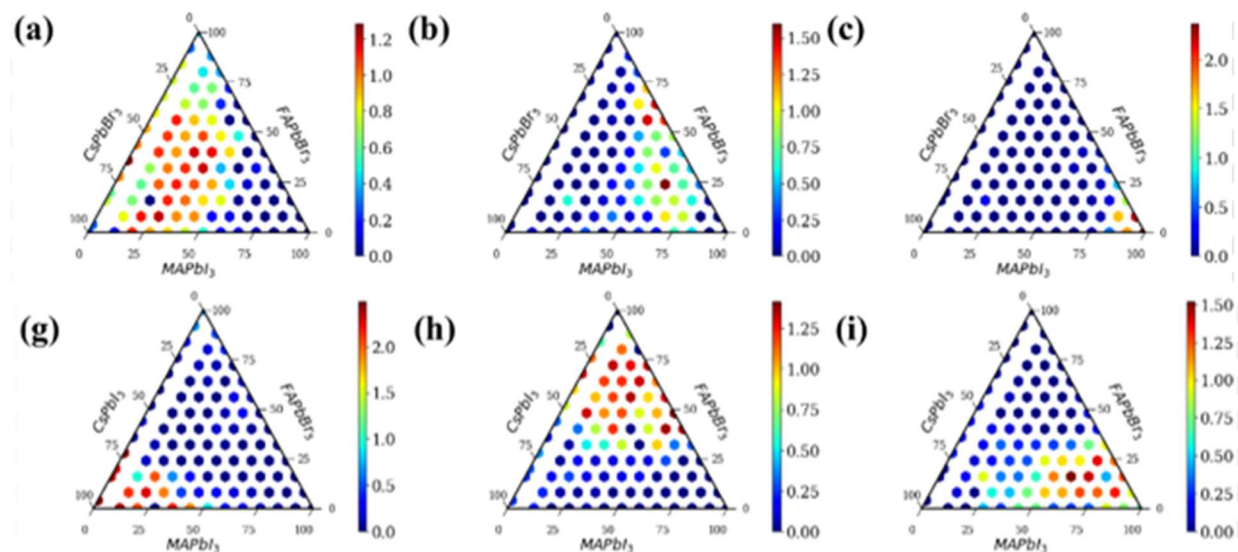
Figure S9. The corresponding triangle maps in **Figure 5** showing the PL analysis performed without the GP interpolation.

1
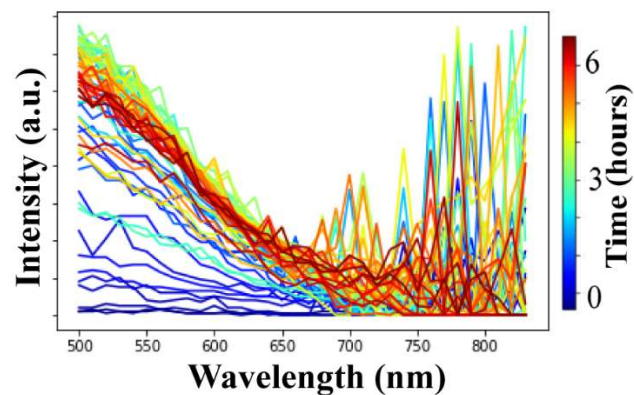
2 **Figure S10.** PL spectra of $CsPbI_3$ microcrystals as function of time. We observe a broad peak

3 around the 500-600 nm region with extraneous peaks at higher wavelengths. In combination with

4 the color of the precipitates, this spectrum indicates the formation of PbI_2 microcrystals instead of

5 $CsPbI_3$. Overall, this is indicative of this composition being unstable before exposure to ambient

6 conditions.