Supplementary Information

Physics-driven discovery and bandgap engineering of hybrid perovskites

Sheryl L. Sanchez,[1] Elham Foadian[1], Maxim Ziatdinov,[2,3] Jonghee Yang[1], Sergei V. Kalinin,[1] Yongtao Liu[*2,] and Mahshid Ahmadi[1*]

[1] *Institute for Advanced Materials and Manufacturing, Department of Materials Science and Engineering, The University of Tennessee Knoxville, Knoxville, Tennessee, 37996, United States*

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

[3] *Computer Science and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*

**Corresponding author emails:*

mahmadi3@utk.edu

liuy3@ornl.gov

Materials

Guanidinium Iodide (Sigma-Aldrich, ReagentPlus 99%), Methylammonium Iodide (Sigma-Aldrich, ReagentPlus 99%), Lead Iodide (Sigma-Aldrich, ReagentPlus 99%), Lead bromide (Sigma-Aldrich, 99.999% trace metals basis), Toluene (Sigma-Aldrich, anhydrous, 99.8%), N-dimethylformamide (DMF), and anhydrous dimethylsulfoxide (DMSO) are used without any further purification.

Precursor Synthesis

First, 0.5 M of MAI and 0.5 M of PbI_2 were dissolved in a 9:1 volume ratio mixture of DMF: DMSO. The precursor solution was then stirred continuously for 5 hours at room temperature to produce 0.5 M of completely dissolved $MAPbI_3$ precursor. At the same time, 0.5 M of $GAPbBr_3$ precursor solution was prepared in the same way with dissolving $PbBr_2$ and GABr in DMF: DMO. The solutions were then mixed in different volume ratios to obtain $GA_xMA_{1-x}Pb(Br_xI_{1-x})_3$.

Thin Film Fabrication

Thin films were cast on glass substrates which are ultrasonically cleaned in deionized water, acetone, and isopropanol respectively in an ultrasonic cleaning bath for 15 mins each. The substrates were dried by blowing dry nitrogen and treated with UV-Ozone for 20 min before spin coating the precursor solution.

The spin-coated films were fabricated using a two-step process with antisolvent treatment. The spin coating process uses 1000 rpm for 10 s followed by 3500 rpm for 25 s. 10 s prior to the end of the spin coating cycle, the antisolvent (Toluene) was added to the film. Then, the films were annealed at 100 °C for 20 min.

Photoluminescence and absorption spectroscopy

Photoluminescence spectroscopy was conducted utilizing the Biotek Cytation 5 Hybrid Multi-Mode Reader. The excitation wavelength was adjusted to 430 nm, while the emitted light was detected within the 500 to 850 nm range, with an increment of 1 nm. The measurements were taken 7 mm below the well plate and employed the sweep mode.

Data Availability

GPax: https://github.com/ziatdinovmax/gpax

Colab notebooks can be found at Github Repository: https://github.com/SLKS99/-Physics-driven-discovery-and-optimization-of-hybrid-perovskite-films/tree/main

Data Analysis

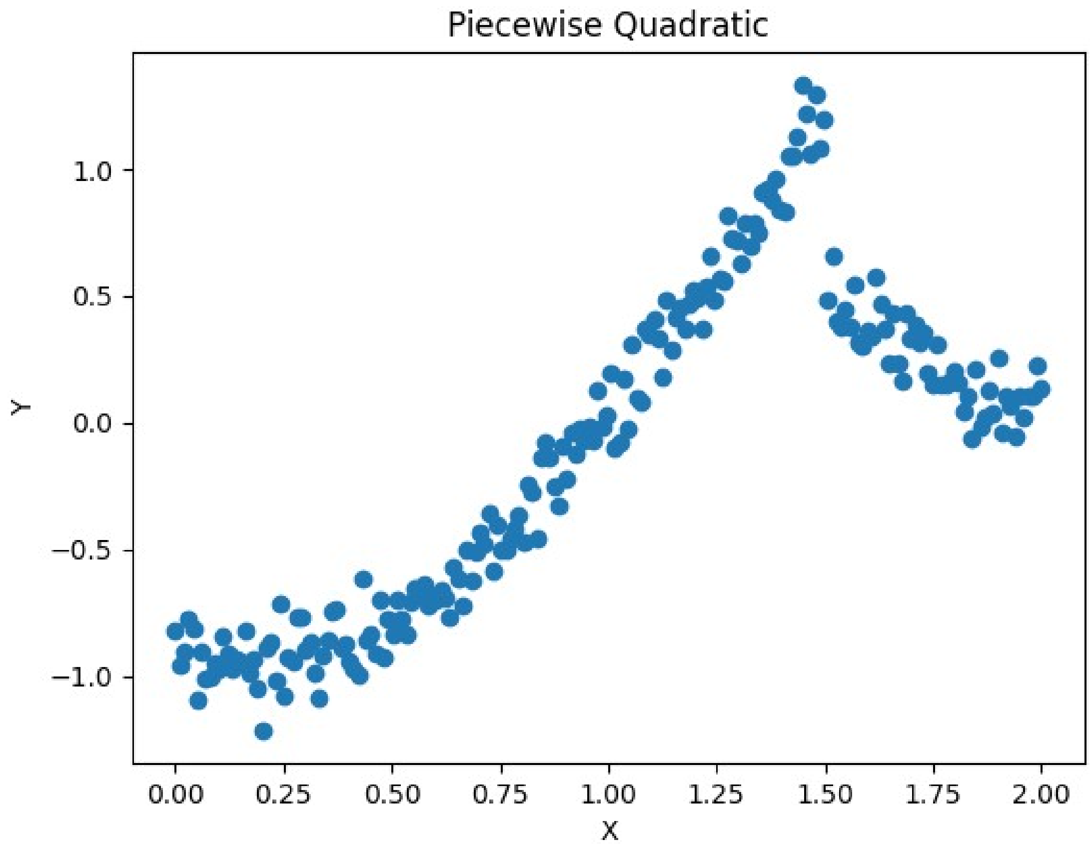

Figure S1: *Toy model that is piecewise with a quadratic and linear interval and high noise (0.1)*

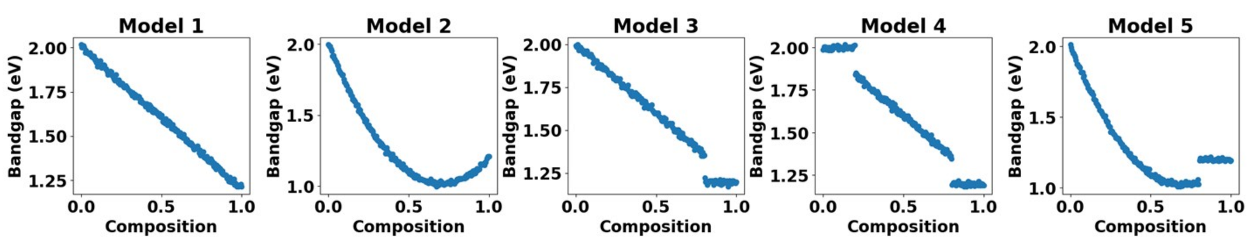

Figure S2: *Comparative Illustration of Five Bandgap Toy Models for* AB_xC_{1-X}

We have developed five toy models to represent the bandgap change of the mixed material with AB_xC_{1-x} chemical formula, where AB has a bandgap of 1.2 eV and AC has a bandgap of 2 eV, as a function of x .

1. Linear Model: The first model assumes a linear relationship between the bandgap and x. The equation is given as $E_{abc} = E_{ab} \times x + E_{ac} \times (1 - x)$, where E_{ab} and E_{ac} represents the bandgap energies of AB and AC, respectively.

2. Bowing Model: In the second model, we introduce the concept of bandgap bowing. The equation becomes $E_{abc} = E_{ab} \times x + E_{ac} \times (1 - x) + k \times x \times (1 - x)$, where k represents the bowing coefficient.

3. Alloy Formation Model with Threshold (x < 20%): The third model incorporates the condition of alloy formation. When the ratio of AC is smaller than 20%, no alloy is formed, and the bandgap is solely determined by AB. Therefore, we assume the bandgap of $ABxC_{1-x}$ in this range is constant and equal to the bandgap of AB. The equation is given as $E_{abc} = E_{ab}$ for x > 80% and $E_{ab} \times x + E_{ac} \times (1 - x)$ for x < 80%.

4. Alloy Formation Model with Thresholds (x < 20% and x > 80%): The fourth model extends the alloy formation condition to include extremely small values of x . In this case, no alloy is formed when x is below 20% or above 80%. The equation becomes $E_{abc} = E_{ab}$ for $x \geq 80\%$, E_{ac} for $x \leq 20\%$, and $E_{ab} \times x + E_{ac} \times (1 - x)$ for $20\% < x < 80\%$.

5. Alloy Formation Model with Bowing (x < 80%): The fifth model combines the alloy formation condition with the bowing effect. No alloy is formed when x is greater than 80%, and the bandgap follows the bowing law when an alloy is formed. The equation is given as $E_{abc} = E_{ab}$ for $x \geq 80\%$ and $E_{ab} \times x + E_{ac} \times (1 - x) + k \times x \times (1 - x)$ for $20\% < x < 80\%$.

These toy models allow us to explore different scenarios and capture the variations in the bandgap of AB_xC_{1-x} as a function of x. They provide insights into the relationship between composition and bandgap, taking into account factors such as linearity, bowing, and alloy formation conditions.

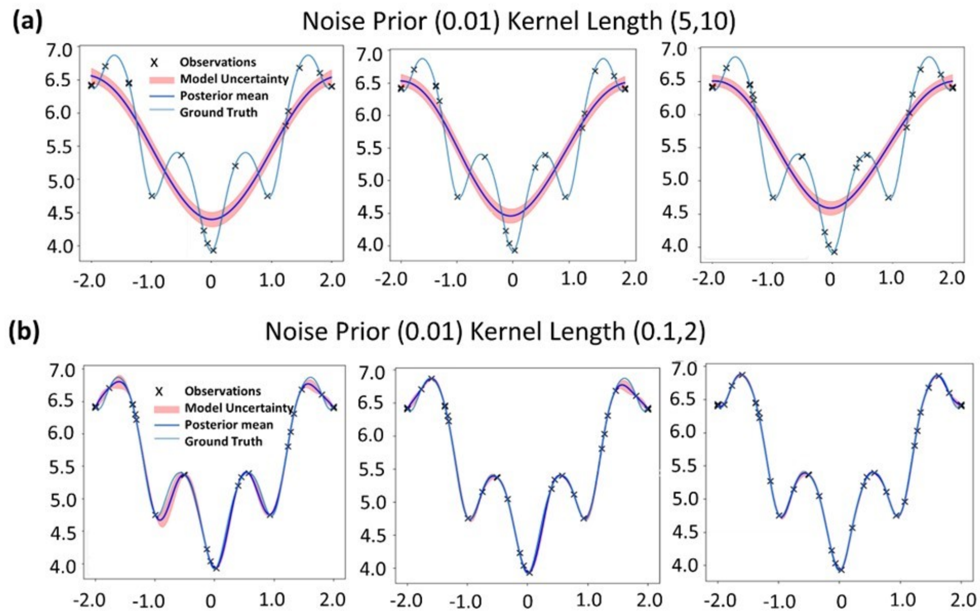

Figure S3: *Effect of priors on ground truth discovery in GP (a) Noise prior (0.01) and kernel length (5,10) (b) Noise prior (0.01) and kernel length (0.1,2)*

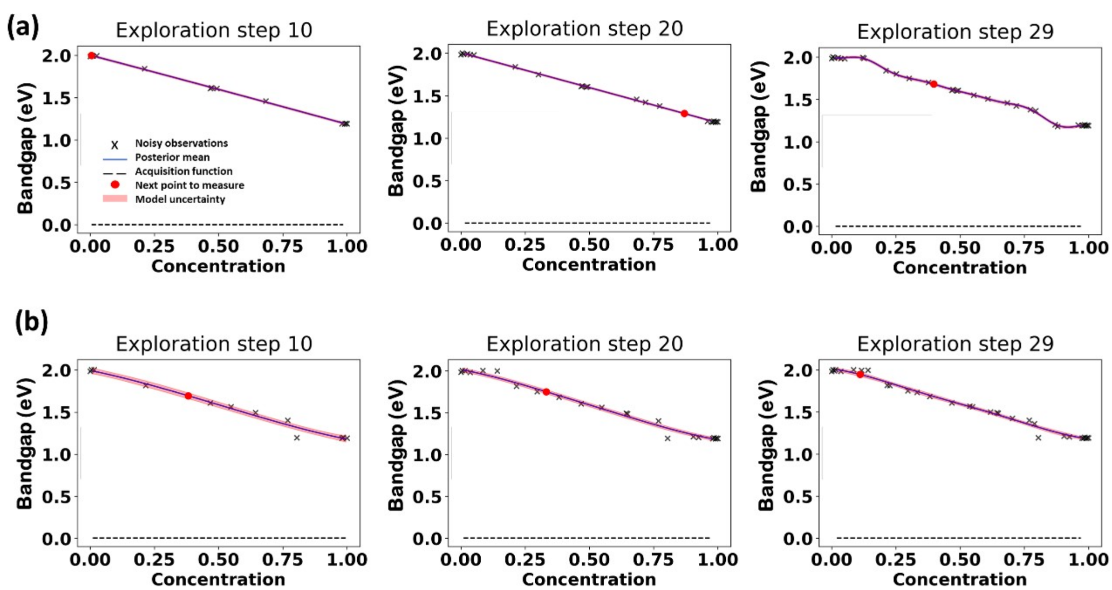

Figure S4: *Effect of priors on ground truth discovery in Model 4 (a) Slope prior a= Normal (0,1000), Intercept b= Normal (0,1000) (b) Slope prior a= LogNormal (0,1), Intercept b= Normal (0,2).*

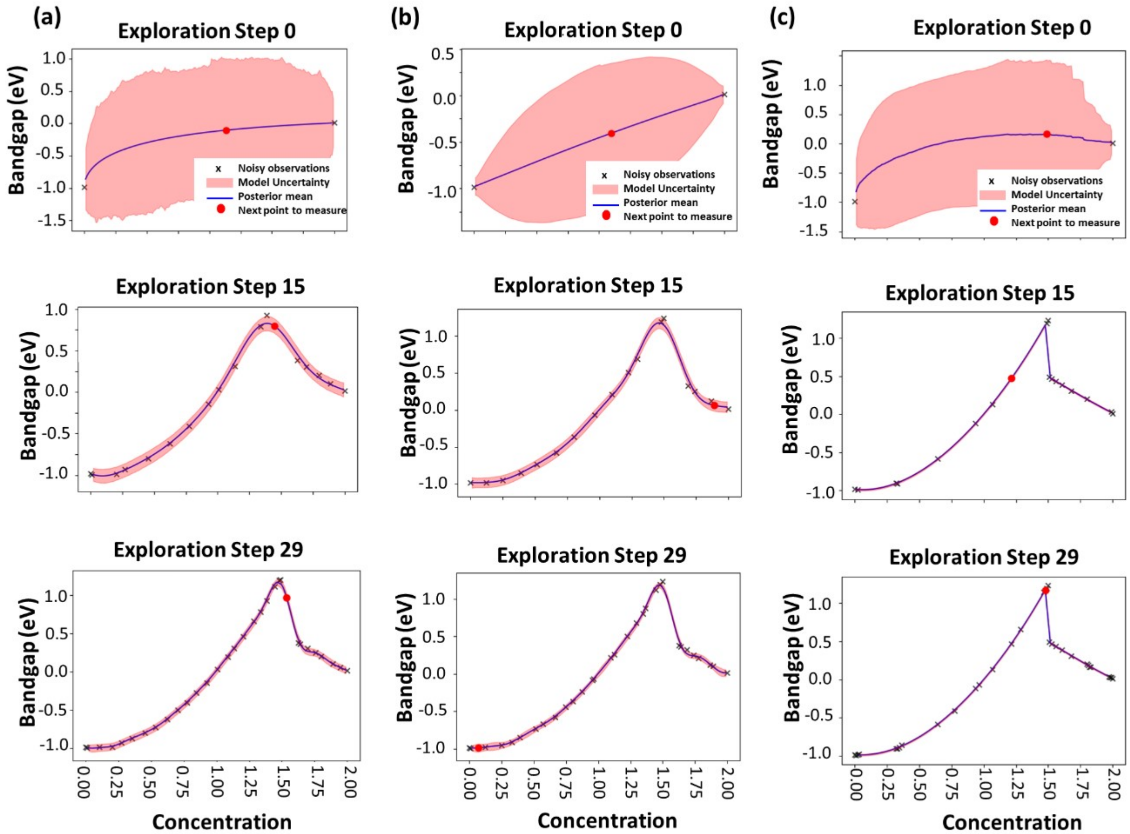

Figure S5: *(a) GP (b) d-sGP with linear mean function c) c-sGP with a piecewise quadratic mean function for piecewise quadratic toy model for 30 exploration steps with noise of (0.01) and noise prior of (0.1)*

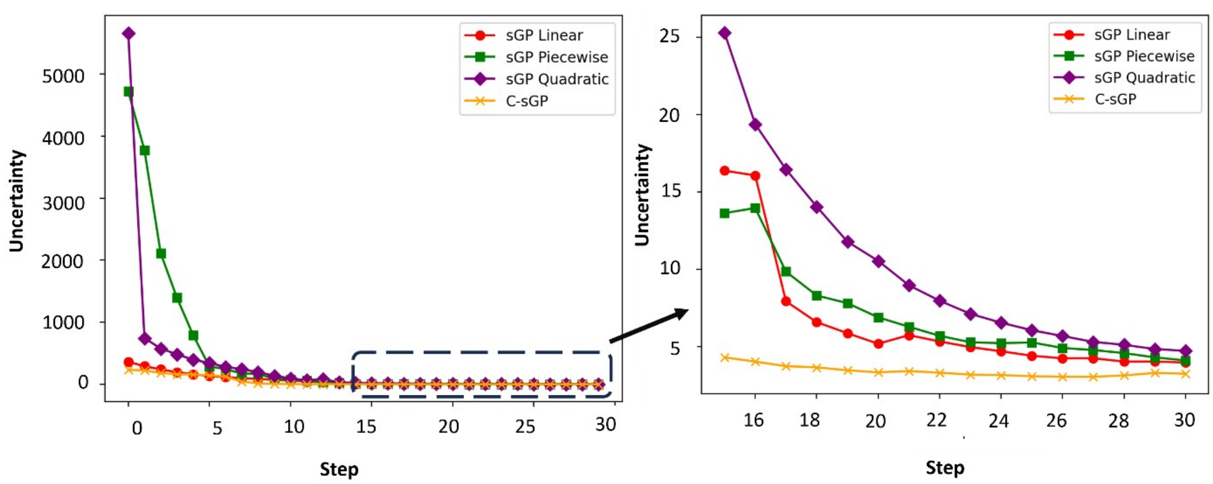

Figure S6: *Uncertainty d-sGP and c-sGP with mean functions linear, quadratic, piecewise and custom bandgap models.*

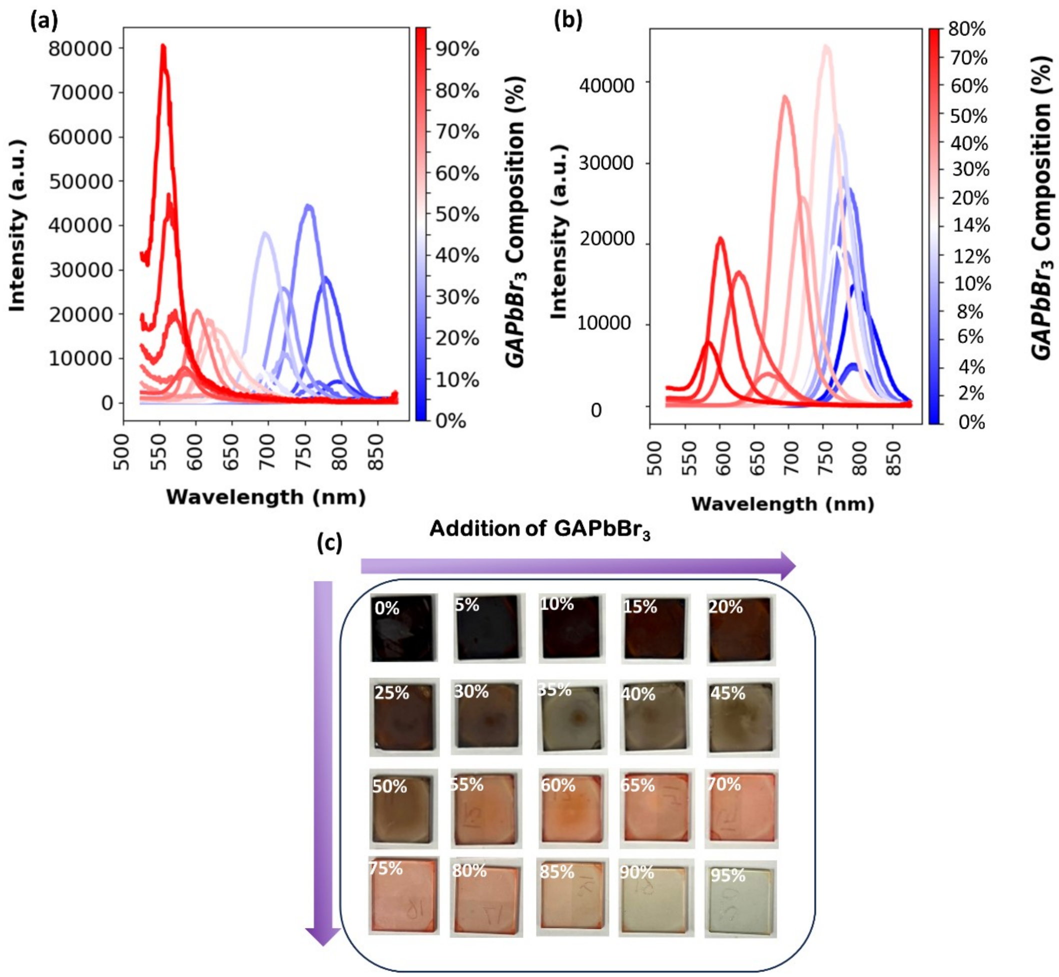

Figure S7: *PL of mixed $GA_xMA_{1-x}Pb(Br_xI_{1-x})_3$ thin films (a) with 5% increment of $GAPbBr_3$ addition (b) 2% increment up 14% then 10% increment up to 80% (c) thin films at 5% increments.*

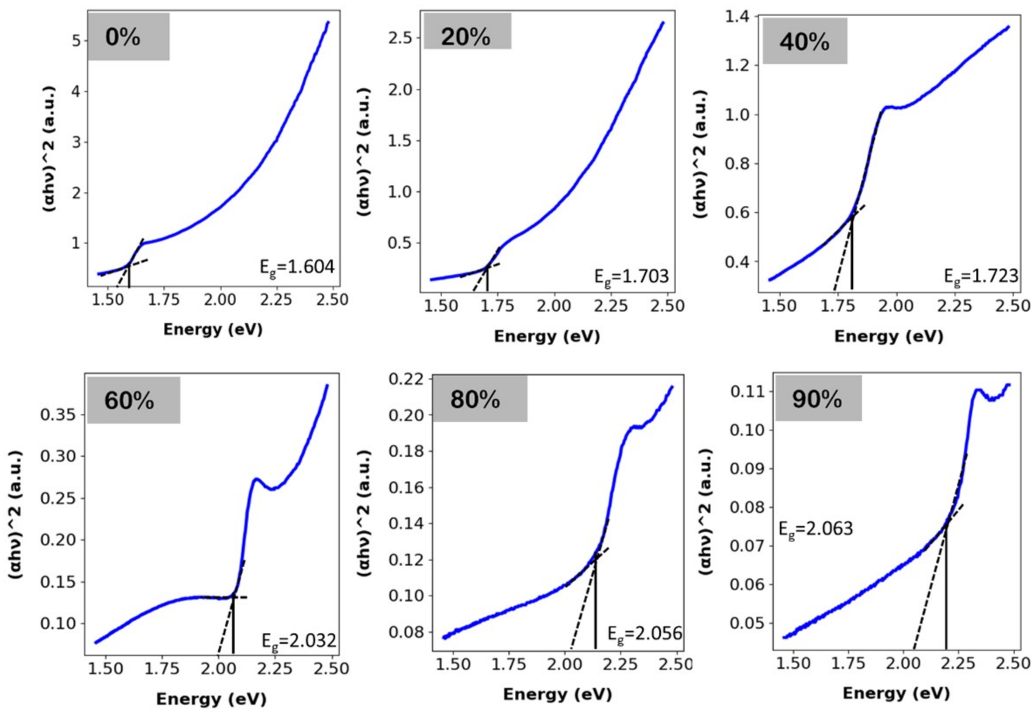

Figure S8: *Tauc Plot of Measured Bandgap of various addition of* $GAPbBr_3$.