SUPPLEMENTARY INFORMATION

Empirical equations to estimate tetragonal to cubic phase transitions in lead iodide perovskites

Fernando B. Minussi[1,2,3*], Rogerio M. Silva Jr.[4], Thays A. Abreu[5], Jose A. Eiras[3], Eudes B. Araujo[2]

[1] Department of Mechanical Engineering, Santa Catarina State University, Joinville, SC, Brazil

[2] Department of Physics and Chemistry, São Paulo State University, Ilha Solteira, SP, Brazil

[3] Department of Physics, Federal University of São Carlos, São Carlos, SP, Brazil

[4] Department of Electrical Engineering, São Paulo State University, Ilha Solteira, SP, Brazil

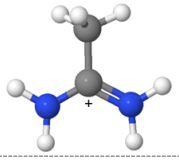
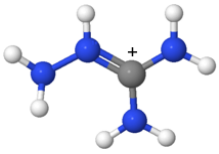
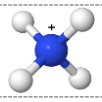
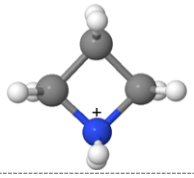
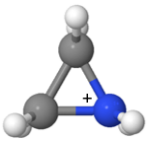
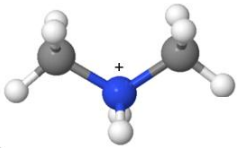
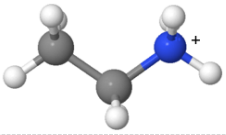
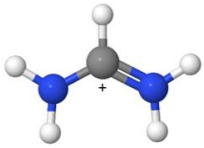
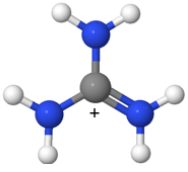
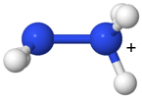
[5] Department of Mathematics, São Paulo State University, Ilha Solteira, SP, Brazil

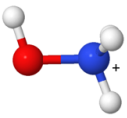
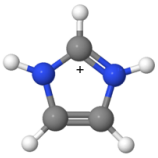
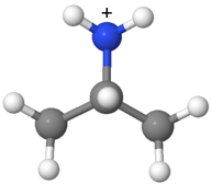
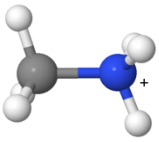
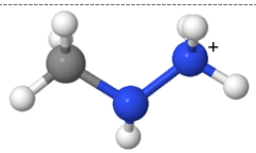
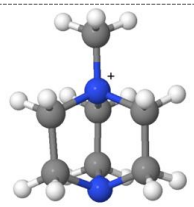
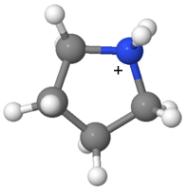
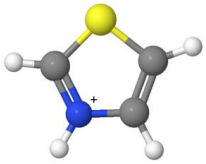
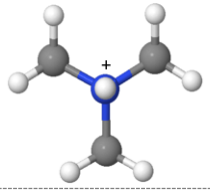
Contents

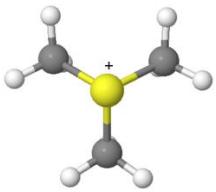
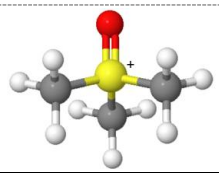
* fbminussi@gmail.com / fb.minussi@udesc.br

Supplementary Note 1: Compositions and phase data of perovskites from literature

Table S1 - Basic characteristics of all A-site cations contained in compounds of the data bank. Dipole moments estimated using MolCalc [1]. r_A, μ_A, and NH_A, respectively, the radius, dipole moment, and number of N-H bonds. Color of spheres representing the elements in molecular cations: hydrogen (white), carbon (gray), nitrogen (blue), oxygen (red), and sulfur (yellow).

Cation (abbreviation)	Structure	r_A (pm)	μ_A (D)	NH_A	Ref.	Observations
Acetamidinium (AC^+)		277	2.02	4	[2]	
Aminoguanidinium (AGA^+)		351	2.22	7		Radius estimated with the procedure given by [3].
Ammonium (NH_4^+)		146	0	4	[3]	
Azetidinium (AZE^+)		250	3.79	2	[3]	
Aziridinium (AZI^+)		227	2.77	2	[4]	
Cesium (Cs^+)		188	isotropic	0	[5]	Ionic radius for CN = 12.
Dimethylammonium (DMA^+)		272	1.93	2	[3]	
Ethylammonium (EA^+)		274	4.97	2	[3]	
Formamidinium (FA^+)		253	0.21	4	[3]	
Guanidinium (GA^+)		278	0	6	[3]	
Hydrazinium (HY^+)		217	3.5	5	[3]	

Hydroxylammonium (HOA^+)		216	4.21	4	[3]	The O-H bond was included as a N-H bond.
Imidazolium (IM^+)		258	1.99	2	[3]	
Isopropylammonium (iPA^+)		282	5.57	3		For simplicity, we considered its radius the same as trMA.
Lithium (Li^+)		125	isotropic	0		Radius linearly extrapolated for CN = 12 from values for other CNs given by [5].
Methylammonium (MA^+)		217	2.7	3	[3]	
Methylhydrazinium (MHY^+)		264	5.13	4	[6]	
N-methyldabconium ($MDAB^+$)		375	5.55	0		Radius estimated with the procedure given by [3] with data from [7].
Potassium (K^+)		164	isotropic	0	[5]	Ionic radius for CN = 12.
Pyrrolidinium (Pyr^+)		272	4.77	2		Radius was considered the same of pyrrolinium [8].
Rubidinium (Rb^+)		172	isotropic	0	[5]	Ionic radius for CN = 12.
Thiazolium (Tz^+)		320	4.06	1	[8]	
Trimethylammonium ($trMA^+$)		282	1.12	0		Radius was considered as the mean between DMA^+ and tetramethylammonium (292 pm [3]).

Trimethylsulfonium (TMS^+)		244	1.07	0	[9]
Trimethylsulfoxonium ($TMSO^+$)		289	5.77	0	[10]

Data from 68 references were not included due to unavailable or insufficient phase information for the present work. Reasons for data not being useful include information collected under unspecified experimental conditions, very restricted measurement ranges, provided figures that lead to ambiguous results, results obtained by computational methods, and synthesis methods that are difficult to control. Some articles were also not included because they contained similar or identical data to others already extensively included in the database. Additionally, some systems may have been obtained under "unusual" conditions, such as in [11], leading to results that differ from the literature. Other systems, particularly $CsPbI_3$ and $FAPbI_3$ [12] and, more recently, $EAPbI_3$ [13,14], may not generate perovskites under usual synthesis conditions, but can form perovskites if other routes are employed. A recent example can be found at [15]. Furthermore, it is important to emphasize that films and nanocrystals may behave differently compared to large crystals. An example is $FA_xCs_{1-x}PbI_3$ nanocrystals, which appear to be total solid solutions [16], while in films and bulk crystals, phase separation is observed. Finally, despite earlier propositions [17], it seems to be mostly accepted that the sole or concomitant formation of the so-called $HPbI_3$ in halide perovskite systems is not feasible [18]. Hence, we opted to exclude its possibilities. Others, such as a morphotropic phase boundary in $MA_xFA_{1-x}PbI_3$ [19] were not considered either. Data from ten references were not included in the database because they contain data on materials in which the components of A-site compounds were not necessarily added to replace the A-site cations in the perovskite. Furthermore, the additives may have been used after the initial film formation. Therefore, it is not possible to infer whether the additional phases are due to segregated phases because they are systems above a solubility limit, incomplete dissolution for kinetic reasons, etc. In other cases, additives containing cations or anions that may have been used (e.g. [20]) where it is not possible to guarantee that they have been eliminated from the final materials.

Supplementary Note 2: Models' construction

When choosing the model complexity, the constructed code allows modifying two parameters:

nNonzeroCoefs = d; d = 1, 2, 3 or 4 - try models up to a maximum of 4D (4 terms).

nFeaturesPerSisIter = k; consider k new candidates per iteration.

The code only works if $d * k$ is lower than or equal to the number of descriptors generated. Usually, d = 3 (3D) or 4 (4D) and k =10 or 15 are used. For situations with a small number of descriptors, it is advisable to manipulate the values of the variables d and k so that their product is always less than or equal to the length of the Xout array. To perform SISSO regression in MATLAB using the Statistics and Machine Learning Toolbox, two files must be compiled (generateDescriptors.m and SissoRegressor.m), which are found in two independent scripts available at the following link: https://github.com/NREL/SISSORegressor_MATLAB/tree/main.

The function generateDescriptors(xCell, xVars) has the inputs xCell (a matrix of input feature cells, where each cell is a matrix of features with the same units) and xVars (variable names for the input features), and the class SissoRegressor(nNonzeroCoefs, nFeaturesPerSisIter) has the inputs nNonzeroCoefs (number of descriptors in the equation) and nFeaturesPerSisIter (number of iterations).

First, the database must be expanded with the combination of new descriptors (combined operations) that are called in the generateDescriptors script as operatorA1, operatorA2, and so go, according to the following list:

A1 - Square root

A2 - Square power

A3 - Cube power

A4 - Cube root

A5 - Fourth root

A6 - Fourth power

A7 - Fifth root

A8 - Fifth power

B1 - Inverse

C1 - Product of features from different groups

C2 - Absolute difference between features from the same group

C3 - Sum of features

Variables C2 and C3 were only enabled when there were different descriptors in the same unit of measurement. With the feature matrix expanded with the combination of operators, regression is performed using SISSO. Changes had to be made to the generateDescriptors.m file to include the C3 operator and update the syntax. The code used is given in: https://github.com/fbminussi-web/Phase-transition-data/blob/main/Code.docx. The entire set of models obtained is given in: https://github.com/fbminussi-web/Phase-transition-data/blob/main/Models.docx. For the present study, we opted to use the 2D relation with nFeaturesPerSisIter = 15 with operators A1, A2, A3, A4, A5, A6, B1, and C1, because it is quite simple and with a reasonably good RMSE, the best among the 2D models. Noteworthy, the RMSE values in the above document are for the training data, whereas in the manuscript, the RMSE values are for the test data.

Supplementary Note 3: MLA results from raw data

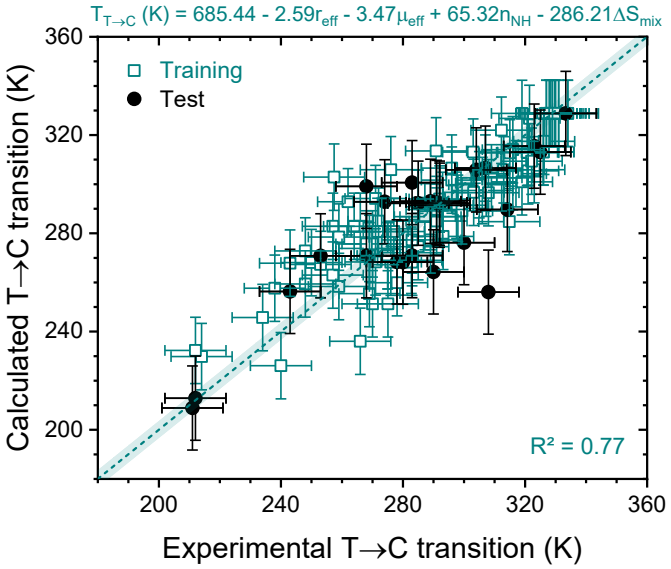


Figure S1 - MLA results before data polishing.

Supplementary Note 4: Visual comparison between MLA and SISSO results

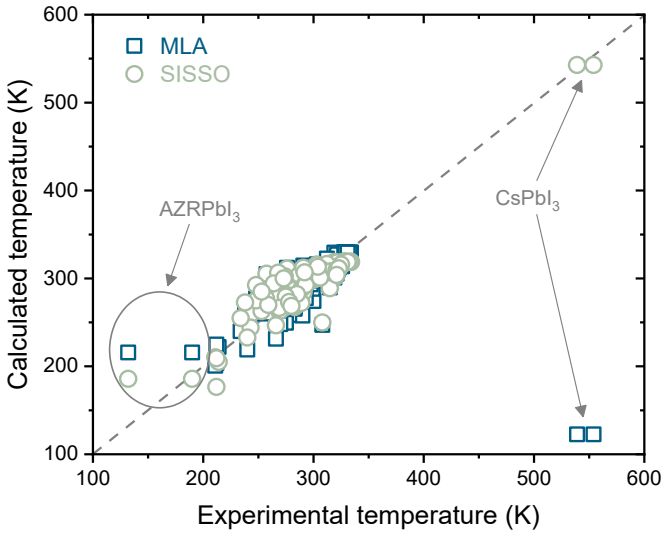


Figure S2 - Calculated temperatures for tetragonal to cubic phase transitions using the two models discussed in the main text. The highlighted data clearly show that MLA performs poorly for compositions with experimental transitions outside the 200-350 K range, but is a bit better for data inside this range.

References

[1] J.H. Jensen, and J.C. Kromann, "The Molecule Calculator: A Web Application for Fast Quantum Mechanics-Based Estimation of Molecular Properties," J. Chem. Educ. **90**(8), 1093–1095 (2013).

[2] P. Singh, R. Mukherjee, and S. Avasthi, "Acetamidinium-Substituted Methylammonium Lead Iodide Perovskite Solar Cells with Higher Open-Circuit Voltage and Improved Intrinsic Stability," ACS Appl. Mater. Interfaces **12**(12), 13982–13987 (2020).

[3] G. Kieslich, S. Sun, and A.K. Cheetham, "Solid-state principles applied to organic–inorganic perovskites: new tricks for an old dog," Chem. Sci. 5(12), 4712–4715 (2014).

[4] C. Zheng, and O. Rubel, "Aziridinium Lead Iodide: A Stable, Low-Band-Gap Hybrid Halide Perovskite for Photovoltaics," J. Phys. Chem. Lett. **9**(4), 874–880 (2018).

[5] R.D. Shannon, "Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides," Acta Cryst A **32**(5), 751–767 (1976).

[6] M. Simenas, A. Gagor, J. Banys, and M. Maczka, "Phase Transitions and Dynamics in Mixed Three- and Low-Dimensional Lead Halide Perovskites," Chem. Rev. **124**(5), 2281–2326 (2024).

[7] B. Saparov, and D.B. Mitzi, "Organic–Inorganic Perovskites: Structural Versatility for Functional Materials Design," Chem. Rev. **116**(7), 4558–4596 (2016).

[8] G. Kieslich, S. Sun, and A.K. Cheetham, "An extended Tolerance Factor approach for organic–inorganic perovskites," Chem. Sci. **6**(6), 3430–3433 (2015).

[9] A. Kaltzoglou, C.C. Stoumpos, A.G. Kontos, G.K. Manolis, K. Papadopoulos, K.G. Papadokostaki, V. Psycharis, C.C. Tang, Y.-K. Jung, A. Walsh, M.G. Kanatzidis, and P. Falaras, "Trimethylsulfonium Lead Triiodide: An Air-Stable Hybrid Halide Perovskite," Inorg. Chem. **56**(11), 6302–6309 (2017).

[10] M. Parashar, R. Singh, K. Yoo, and J.-J. Lee, "Formation of 1-D/3-D Fused Perovskite for Efficient and Moisture Stable Solar Cells," ACS Appl. Energy Mater. **4**(3), 2751–2760 (2021).

[11] H. Park, A. Ali, R. Mall, H. Bensmail, S. Sanvito, and F. El-Mellouhi, "Data-driven enhancement of cubic phase stability in mixed-cation perovskites," Mach. Learn.: Sci. Technol. **2**(2), 025030 (2021).

[12] S. Hu, A.R. Tapa, X. Zhou, S. Pang, M. Lira-Cantu, and H. Xie, "Formation and stabilization of metastable halide perovskite phases for photovoltaics," Cell Reports Physical Science **5**(2), 101825 (2024).

[13] Y. Zhang, and N.-G. Park, "A thin film (<200 nm) perovskite solar cell with 18% efficiency," J. Mater. Chem. A **8**(34), 17420–17428 (2020).

[14] C.M. Guvenc, S. Toso, Y.P. Ivanov, G. Saleh, S. Balci, G. Divitini, and L. Manna, "Breaking the Boundaries of the Goldschmidt Tolerance Factor with Ethylammonium Lead Iodide Perovskite Nanocrystals," ACS Nano **19**(1), 1557–1565 (2025).

[15] S. Kralj, K. Artuk, A. Wieczorek, N. Orlov, Z. Eftekhari, R. Saive, E. Garnett, S. Siol, C.M. Wolff, and M. Morales-Masis, "Template-Assisted Growth of $Cs_x FA_{1-x} PbI_3$ with Pulsed Laser Deposition for Single Junction Perovskite Solar Cells," Advanced Energy Materials **15**(24), 2406033 (2025).

[16] J.A. Vigil, A. Hazarika, J.M. Luther, and M.F. Toney, "$FA_x Cs_{1-x} PbI_3$ Nanocrystals: Tuning Crystal Symmetry by A-Site Cation Composition," ACS Energy Lett. **5**(8), 2475–2482 (2020).

[17] F. Wang, H. Yu, H. Xu, and N. Zhao, "$HPbI_3$: A New Precursor Compound for Highly Efficient Solution-Processed Perovskite Solar Cells," Adv Funct Materials **25**(7), 1120–1126 (2015).

[18] W. Ke, I. Spanopoulos, C.C. Stoumpos, and M.G. Kanatzidis, "Myths and reality of HPbI3 in halide perovskite solar cells," Nat Commun **9**(1), 4785 (2018).

[19] T. Hainer, E. Fransson, S. Dutta, J. Wiktor, and P. Erhart, "A morphotropic phase boundary in MA1−xFAxPbI3: linking structure, dynamics, and electronic properties," Nat Commun **16**(1), 8775 (2025).

[20] N. Cheng, W. Li, M. Zhang, H. Wu, S. Sun, Z. Zhao, Z. Xiao, Z. Sun, W. Zi, and L. Fang, "Enhance the performance and stability of methylammonium lead iodide perovskite solar cells with guanidinium thiocyanate additive," Current Applied Physics **19**(1), 25–30 (2019).