

Machine Learning-Accelerated Discovery of Piezoelectric Hybrid Perovskites

Yu Lim Kim

Department of Chemistry
Soongsil University

2026 New Uzbekistan University – Soongsil University, G-LAMP, International Workshop



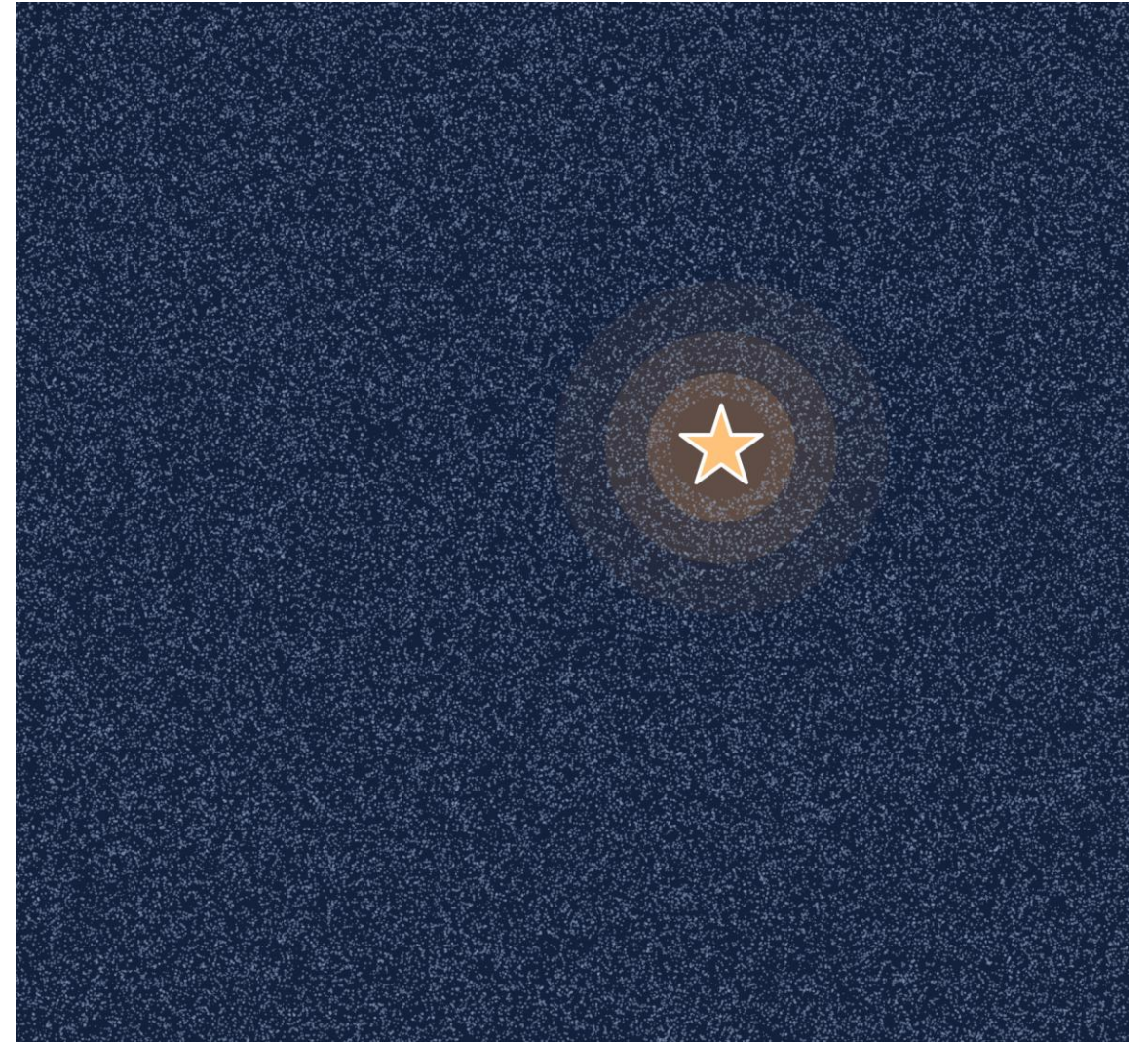
Today's Question

**Can AI find a crystal that
turns a push into
a voltage?**

Millions of candidates.

One is the needle.

Checking a single one takes days of quantum
simulation.



Contents

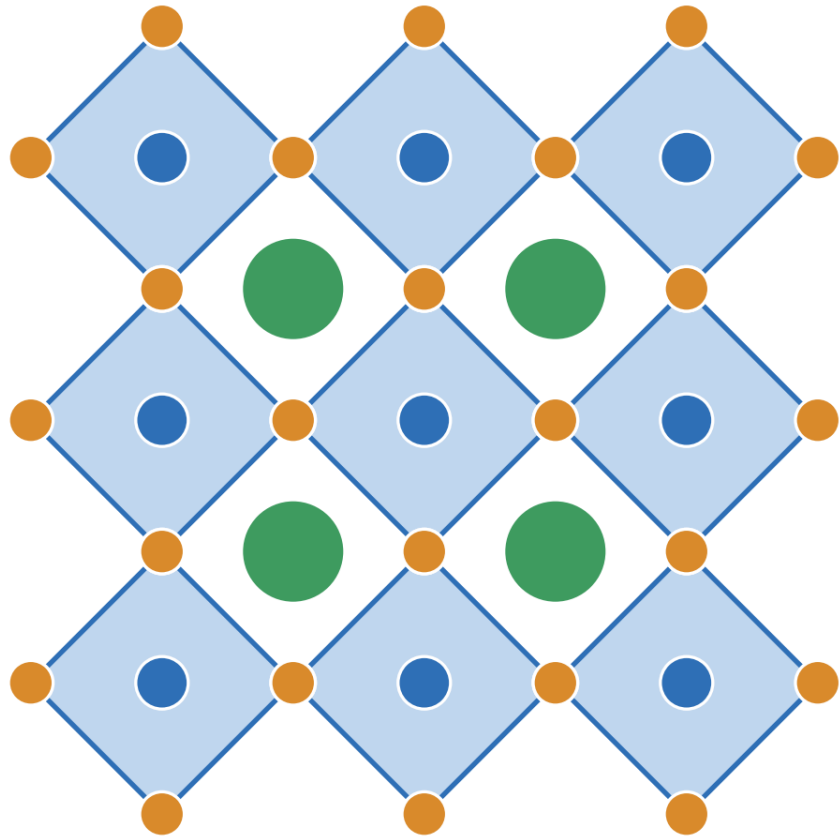
- 1. Introduction:** Perovskites, Piezoelectricity & How We Compute Them
- 2. Topic 1:** Screening Millions of HOIPs with DFT + ML
- 3. Topic 2:** Inventing New Organic Cations
- 4. Topic 3:** Predicting Full Piezoelectric Tensors with Equivariant Networks
- 5. Outlook**

Contents

- 1. Introduction: Perovskites, Piezoelectricity & How We Compute Them**
- 2. Topic 1: Screening Millions of HOIPs with DFT + ML**
- 3. Topic 2: Inventing New Organic Cations**
- 4. Topic 3: Predicting Full Piezoelectric Tensors with Equivariant Networks**
- 5. Outlook**

Perovskite: One Pattern, a Whole Structure

Not one material, but a structural pattern: ABX_3



Top view: corner-sharing BX_6 octahedra; A sits in the holes



Large cation in the cage

Ca^{2+} , Ba^{2+} , Cs^{+} ... or a molecule



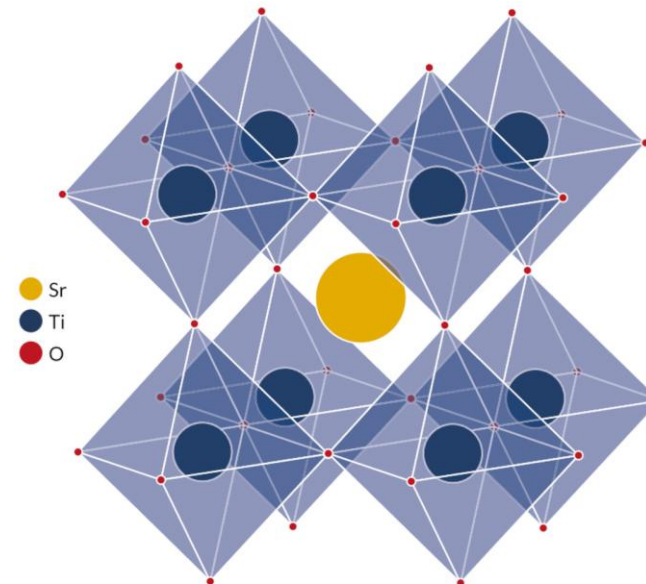
Small metal ion at the octahedron center

Ti^{4+} , Zr^{4+} , Pb^{2+} , Sn^{2+}



Anion at the octahedron corners

O^{2-} , I^{-} , Br^{-} , Cl^{-}

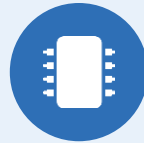


Perovskites Are Already All Around You

Change the letters A, B, X → change the job



Solar cells



Capacitors



Ultrasound imaging



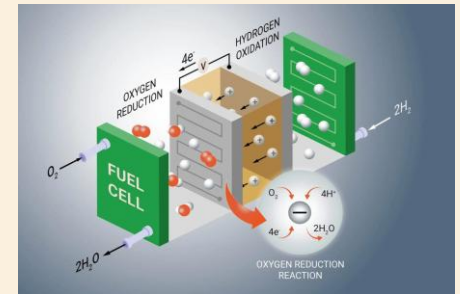
LEDs & displays



X-ray detectors

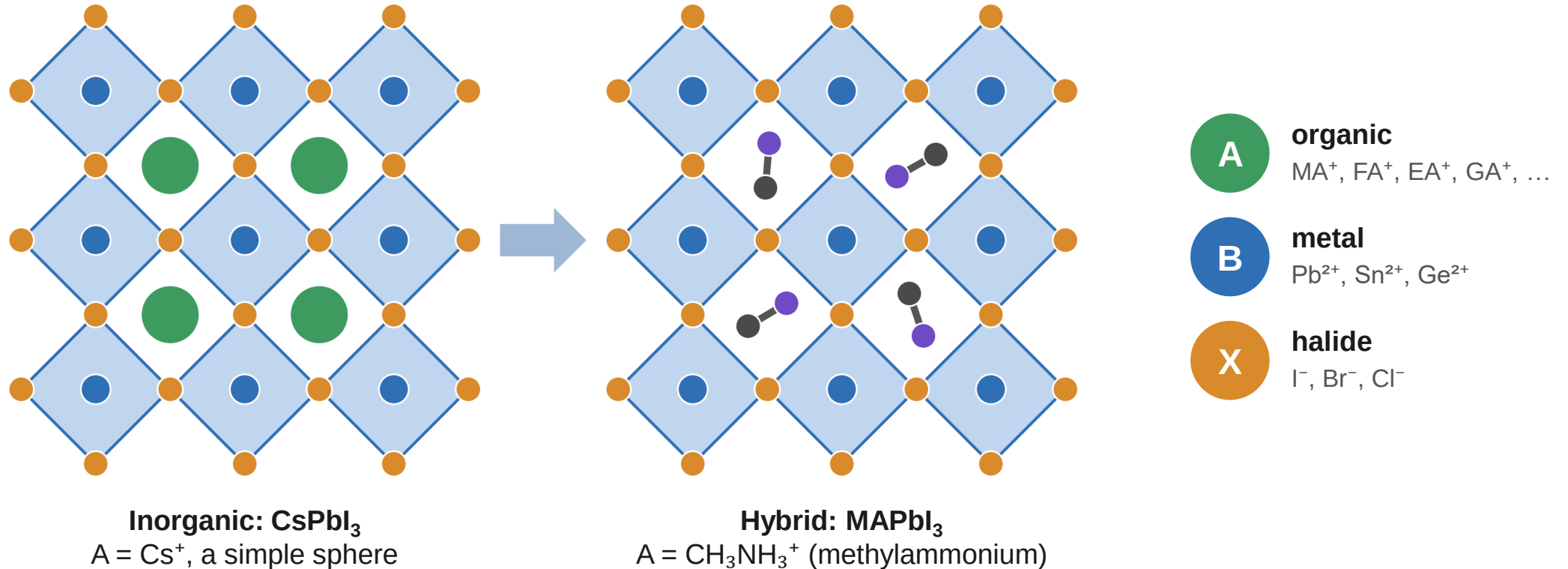


Fuel cells & catalysis



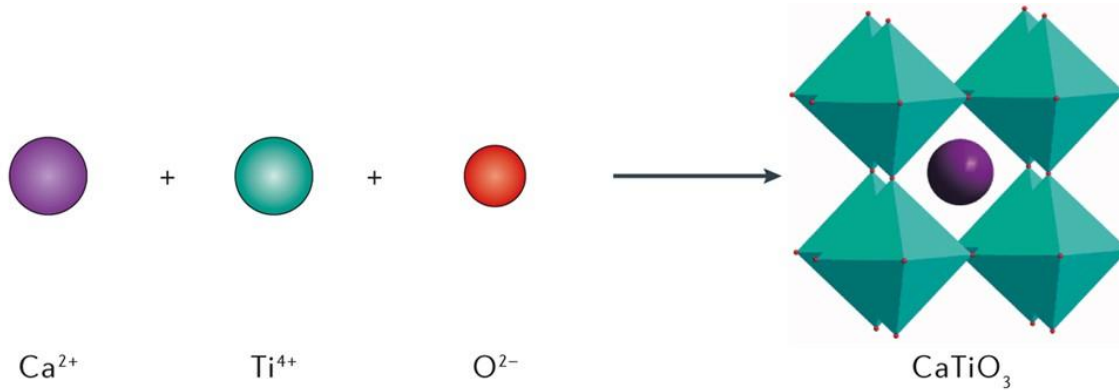
Hybrid Organic–Inorganic Perovskites (HOIPs)

Swap the sphere in the cage for a molecule



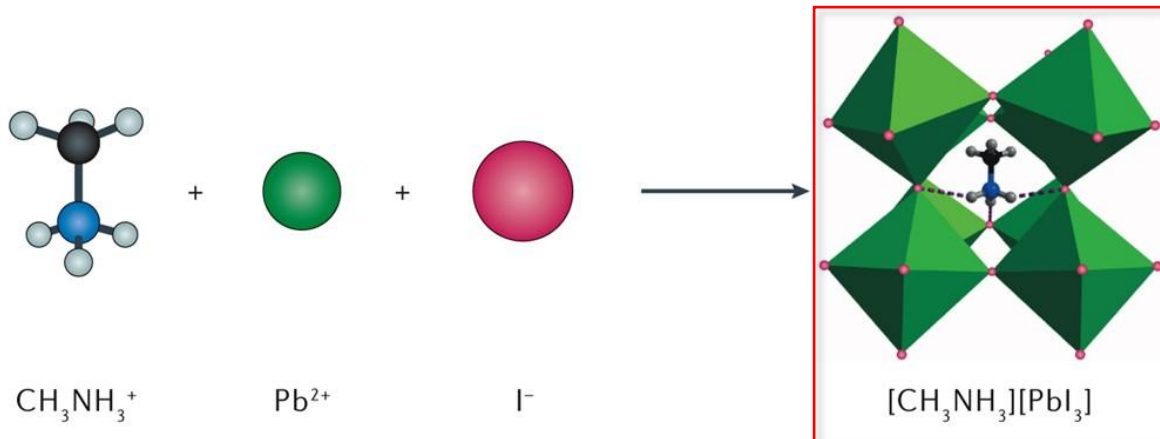
Hybrid Organic–Inorganic Perovskites (HOIPs)

a



(a) Example of an **inorganic perovskite oxide**, CaTiO_3 (orthorhombic, Pbnm).

b



(b) Example of a **HOIP with an organic cation** at the A-site, MAPbI_3 (orthorhombic, Pnma; MA = methylammonium).

Why Go Hybrid? Inorganic vs. Hybrid Perovskites

Same pattern, very different properties

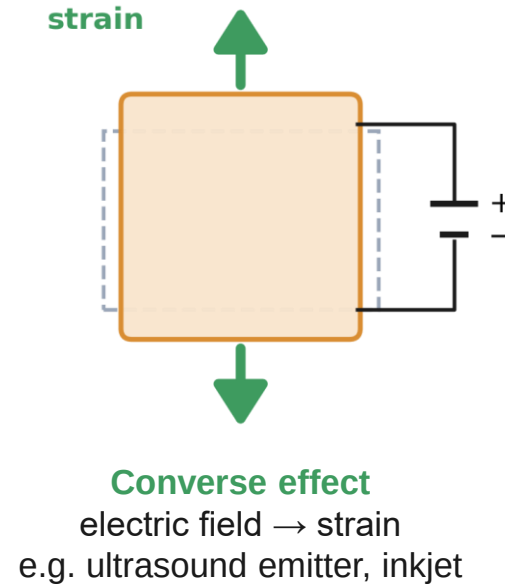
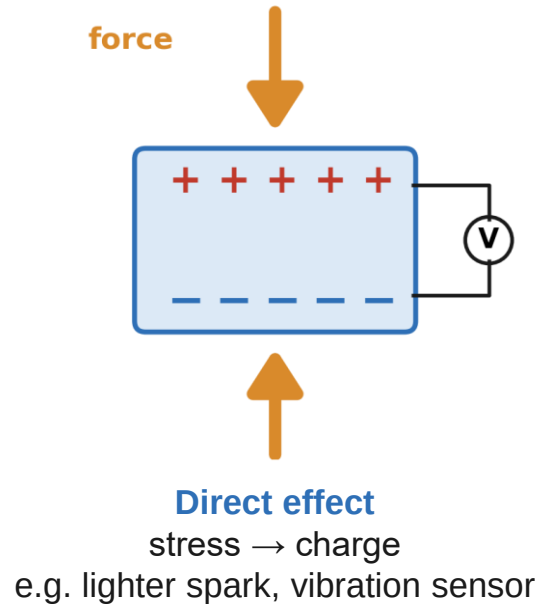
	Inorganic perovskites (PZT, BaTiO ₃)	Hybrid perovskites (HOIPs)
Processing	Sintering above 800 °C	Solution-processed at ~100–150 °C
Mechanics	Hard and brittle	Soft and flexible → wearables
Design space	A handful of inorganic ions	Thousands of organic molecules
Symmetry breaking	Ceramics must be electrically poled	Molecular ordering → built-in polarity
Toxicity	PZT contains lead	Lead-free Sn / Ge options
Functions	Mostly one job	Piezo + ferro + optoelectronic

Needs improvement: stability (moisture, heat) and an enormous design space.

We can't synthesize everything → we need computation to choose.

Piezoelectricity

Squeeze it → voltage. Apply voltage → it deforms.



Piezoelectricity

: the ability of a dielectric material to **convert mechanical energy into electrical energy** (direct effect) and vice versa (converse effect), characterized by the **piezoelectric coefficient** d_{ij} (pC/N)

Why Piezoelectrics Matter

You use dozens of them every day, without noticing



Quartz watch

Quartz rings at $32,768 \text{ Hz} = 2^{15}$
→ halve 15 times = 1 tick/s



Smartphone

Acoustic RF filters pick your
signal out of the air



Car

Knock sensors, fuel injectors,
parking sensors



Hospital

Ultrasound imaging probes



Wearables

Pulse & motion sensors in
smart patches



Energy harvesting

Footsteps and vibrations →
power for tiny sensors

Why we focus here



Get the lead out

Pb is toxic; regulators want lead-free
piezoelectrics



Make it soft

Wearables, implants and flexible sensors need
bendable materials



Sense better

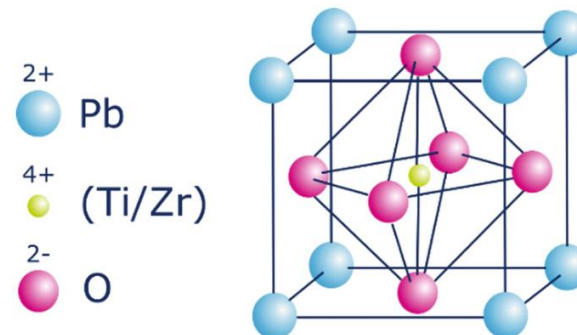
Higher d_{33} → sharper ultrasound, more sensitive
sensors



Search, don't guess

HOIPs: huge design space → compute + ML to
find the needle

Today's champion is PZT
: excellent ($d_{33} \approx 600 \text{ pC/N}$),
but ~60% lead by weight and
brittle as glass.



Piezoelectric Tensor

How piezoelectric is it? The answer is a rank-3 tensor

Linear response, both directions

$$P_i = \sum_{j,k} d_{ijk} \sigma_{jk} \quad \text{charge from stress}$$

$$\eta_{jk} = \sum_i d_{ijk} E_i \quad \text{strain from field}$$

σ = stress, η = strain, $\mathbf{P}$ = polarization, $\mathbf{E}$ = electric field

27

$3 \times 3 \times 3$ components of d_{ijk}

18

σ and η are symmetric 3×3 matrices $\rightarrow d_{ijk} = d_{ikj}$

3×6

Relabel the 6 independent pairs $jk \rightarrow 1 \dots 6$ (Voigt)

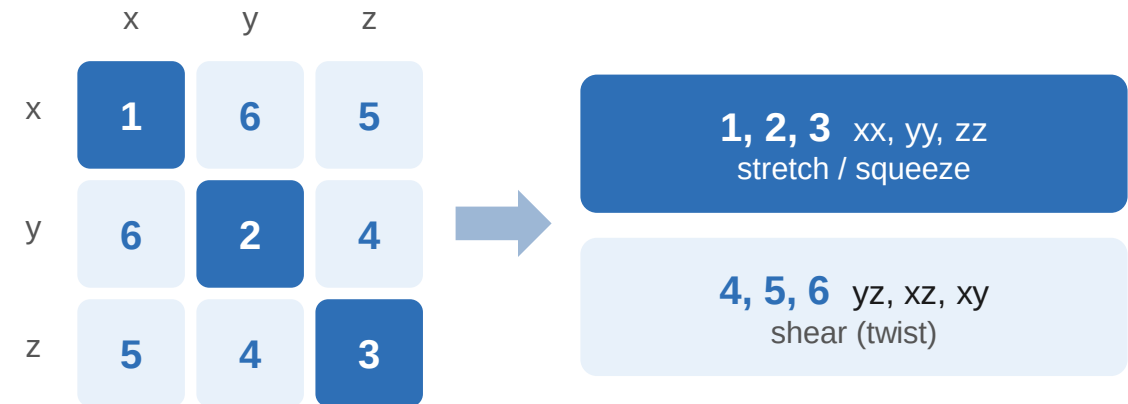
jk	xx	yy	zz	yz	xz	xy
Voigt	1	2	3	4	5	6

Why only 6? Both matrices are symmetric

Stress: $\sigma_{xy} \neq \sigma_{yx}$ would put a net torque on every tiny cube, spinning it infinitely fast. So $\sigma_{xy} = \sigma_{yx}$.

Strain: the antisymmetric part of ∇u is a rigid rotation, and rotating a crystal doesn't deform it.

Voigt notation: a symmetric 3×3 matrix $\rightarrow$ a vector in $\mathbb{R}^6$



Reading the Piezoelectric Matrix

Which way you push → which way the charge goes

Example: point group mm2 (polar HOIPs; axis 3 = polar axis z)

	1	2	3	4	5	6
1	0	0	0	0	d_{15}	0
2	0	0	0	d_{24}	0	0
3	d_{31}	d_{32}	d_{33}	0	0	0

row = direction of charge · column = type of stress

d_{33}

Push along z → charge along z. The headline number.

d_{31}, d_{32}

Push along x or y → charge still collects along the polar axis z.

d_{15}, d_{24}

Shear (twist) → charge along x or y.

Two ways to ask the same question

	e (stress coeff.)	d (strain coeff.)
Input	strain η	stress σ
Question	stretch it → charge?	push it → charge?
Unit	C/m ²	pC/N
Natural for	DFT / DFPT	experiment

$d = e \cdot C^{-1} \rightarrow$ soft crystals win

Small stiffness $C \rightarrow$ large strain per push $\rightarrow$ large d , even with a modest e .

Piezoelectricity (d_{33})

Key d_{33} Benchmarks

Material	d_{33} (pC/N)	Type
PZT	~600	Inorganic ceramic
BaTiO ₃	~190	Inorganic oxide
PVDF (poled)	~33	Piezo-polymer
MASnI ₃	~100.9	Lead-free HOIP
TMCM ₂ SnCl ₆	~137	Double perovskite
MAPbI ₃	~5	Halide HOIP
β -glycine	~200 (d_{16})	Bio-organic
Boc-Dip-Dip peptide	~73	Dipeptide

Limitations of Traditional Materials

Material	Problem
PZT, BaTiO ₃	Toxic Pb, brittle, requires high-T sintering (>800°C)
PVDF	β -phase difficult to obtain; needs mechanical stretching + poling >50 MV/m
ZnO, AlN, GaN	Poor biocompatibility; limited wearable/medical use

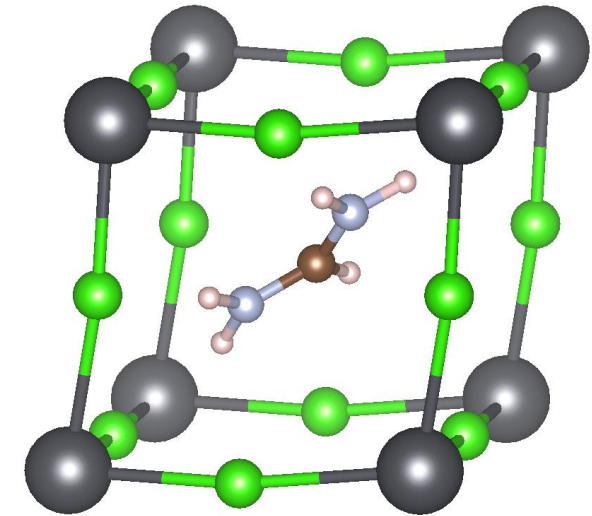
Vijayakanth, Thangavel, et al. "Recent advances in organic and organic–inorganic hybrid materials for piezoelectric mechanical energy harvesting." *Advanced Functional Materials* 32.17 (2022): 2109492.

Motivation

Key Question: Which HOIPs yield high piezoelectric response?

The Challenge

- Conventional piezoelectrics face **lead toxicity**, and high-temperature processing
- **Limited exploration** of organic-inorganic perovskites for piezoelectric applications
- Vast compositional space of organic cations remains largely **unexplored**
- No systematic computational screening framework exists for piezoelectric HOIPs



Computational Chemistry

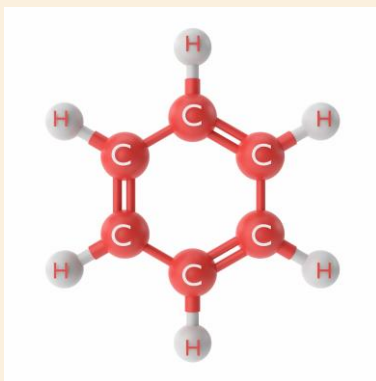
Why not just compute it? Meet the curse of dimensionality

The equation of all chemistry (in principle)

Schrödinger equation

$$\hat{H}\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = E\Psi, \quad \Psi: \mathbb{R}^{3N} \rightarrow \mathbb{C}$$

N electrons $\rightarrow$ one wavefunction on $\mathbb{R}^{3N}$.



10^{126}

Benzene: 42 electrons $\rightarrow \Psi$ on $\mathbb{R}^{126}$.
10 grid points per axis $\rightarrow 10^{126}$ numbers.
(Atoms in the observable universe: $\sim 10^{80}$)

The accuracy–cost ladder

cost scaling

Coupled cluster, CCSD(T)

tens of atoms

$O(N^7)$

Density functional theory (DFT)

hundreds of atoms

$O(N^4)$

Hartree-Fock, HF

hundreds of atoms

$O(N^3)$

Classical force fields

millions+ of atoms

$O(N)$

↑ more accurate, more expensive

Density Functional Theory (DFT)

The trick: trade a 3N-dimensional wavefunction for a 3-dimensional density

Hohenberg–Kohn (1964)

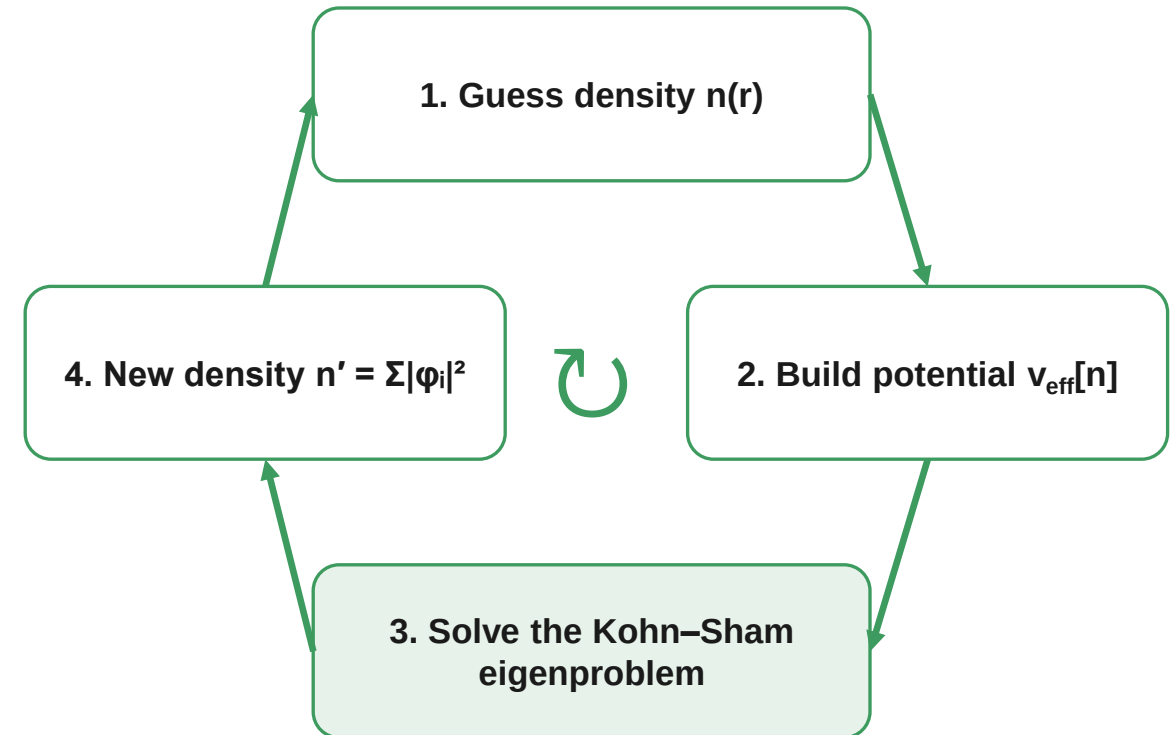
$$E_0 = \min_n E[n], \quad n: \mathbb{R}^3 \rightarrow \mathbb{R}$$

The ground-state energy is a functional of the electron density $n(\mathbf{r})$: 3 dimensions, however many electrons.

Kohn–Sham (1965)

$$E[n] = T_s[n] + \int v n d\mathbf{r} + E_H[n] + E_{xc}[n]$$

Every term is known exactly **except** the exchange–correlation part E_{xc} , which must be approximated (LDA, GGA/PBE, hybrids...).



$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}[n]\right] \phi_i = \epsilon_i \phi_i, \quad n = \sum_i |\phi_i|^2$$

Repeat until $\|n' - n\| < \text{tol}$: a fixed-point iteration $n = F(n)$

Density Functional Perturbation Theory (DFPT)

Material properties are derivatives of the energy, DFPT computes them analytically

$$E(\lambda) = E^{(0)} + \lambda E^{(1)} + \frac{1}{2} \lambda^2 E^{(2)} + \dots$$

Perturbations λ : atomic displacement u , strain η , electric field $\mathcal{E}$. Mixed second derivatives = properties.

	u	η	$\mathcal{E}$
u	K force constants → phonons	Λ internal strain	Z^* Born effective charges
η	Λ internal strain	C elastic constants	$\bar{e}$ piezoelectric (clamped)
$\mathcal{E}$	Z^* Born effective charges	$\bar{e}$ piezoelectric (clamped)	ϵ^∞ dielectric tensor

$\partial^2 E / \partial \lambda_1 \partial \lambda_2$: the Hessian of the energy (symmetric)

Term	Physical meaning	In plain words
$K(u,u)$	Restoring force when an atom is displaced	Springs between atoms → vibrations (phonons)
$C(\eta,\eta)$	Resistance to stretching	How stiff the crystal is
$\epsilon^\infty(\mathcal{E},\mathcal{E})$	How much the electron cloud shifts in an electric field	Dielectric constant
$Z^*(u,\mathcal{E})$	Polarization created by displacing an atom	The "effective" charge an atom actually carries
$\Lambda(u,\eta)$	Force on atoms when the crystal is strained	How much atoms get pushed inside the cell under strain
$\bar{e}(\eta,\mathcal{E})$	Polarization created by stretching	Piezoelectric coefficient (atoms held fixed)

Solve a linear equation for the first-order response
exact derivative

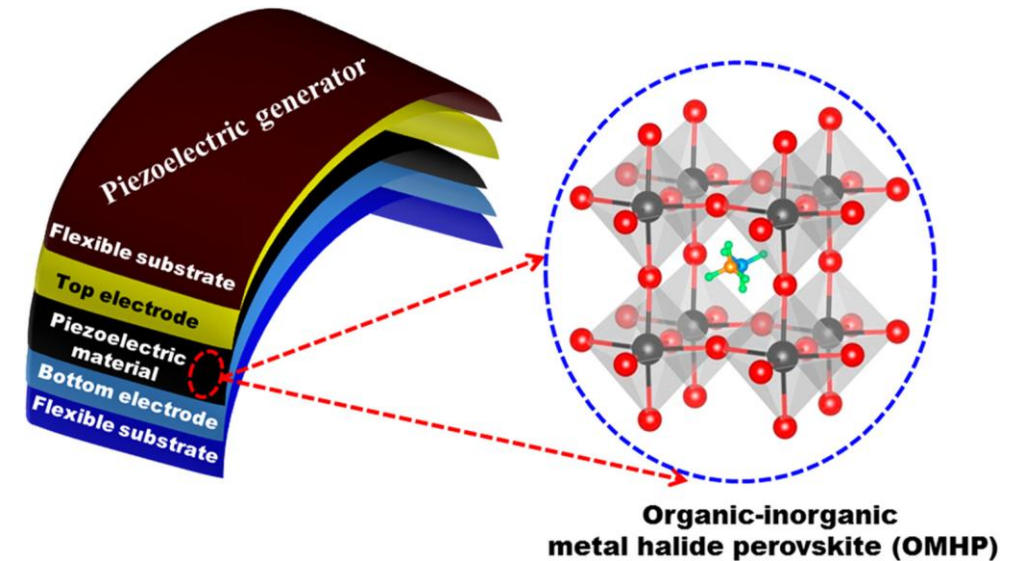
Contents

1. **Introduction:** Perovskites, Piezoelectricity & How We Compute Them
2. **Topic 1: Screening Millions of HOIPs with DFT + ML**
3. **Topic 2:** Inventing New Organic Cations
4. **Topic 3:** Predicting Full Piezoelectric Tensors with Equivariant Networks
5. **Outlook**

HOIPs for Piezoelectrics

Why HOIPs for Piezoelectrics?

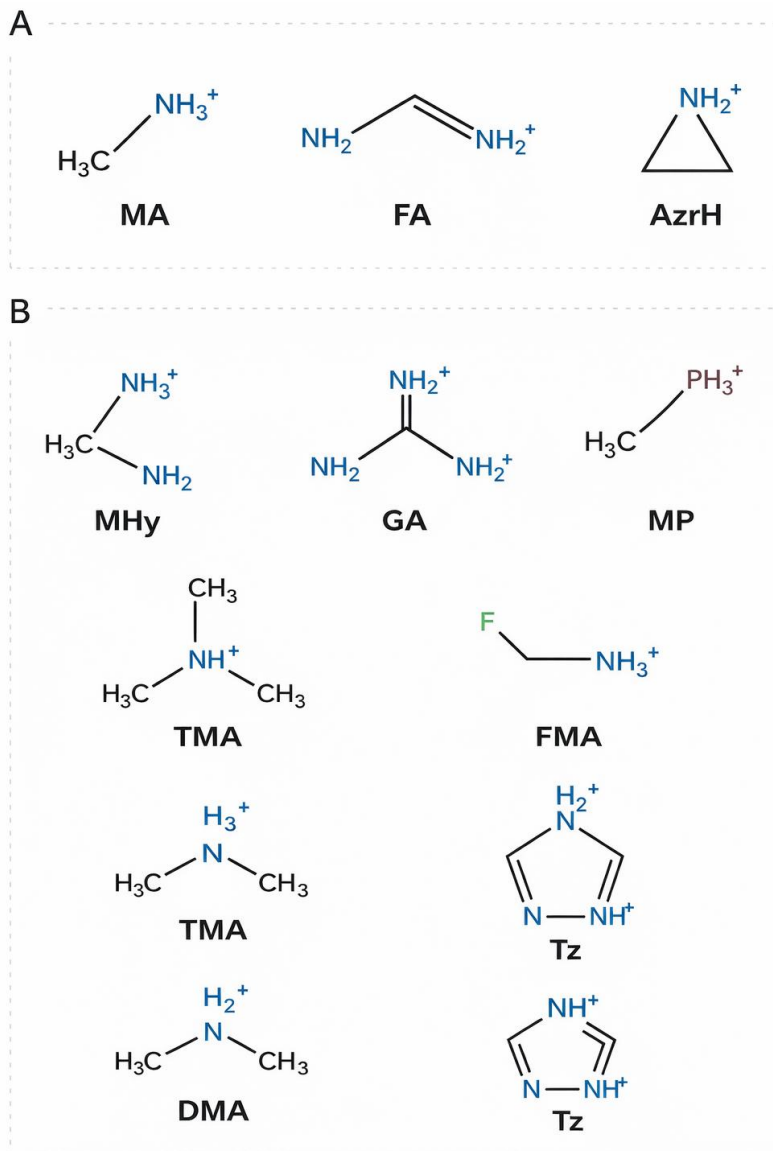
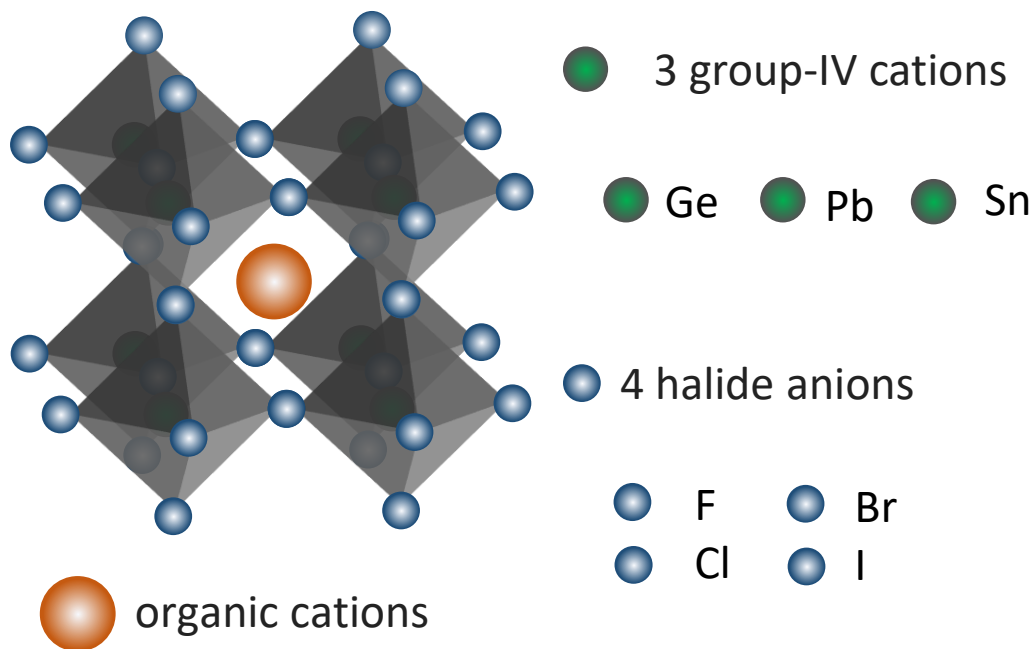
- **Non-centrosymmetric structures** driven by organic cation ordering → intrinsic polarization
- **Tunable d_{33}** via A-site cation and halide substitution
- **Solution-processable** at low temperature, no high-temperature sintering
- **Mechanically flexible** : compatible with wearable/flexible devices
- **Multifunctional**: piezo + pyro + ferro + optoelectronic in one material



Jella, Venkatraju, et al. "A comprehensive review of flexible piezoelectric generators based on organic-inorganic metal halide perovskites." *Nano Energy* 57 (2019): 74-93.

Chemical composition of HOIPs

Limited chemical compositions



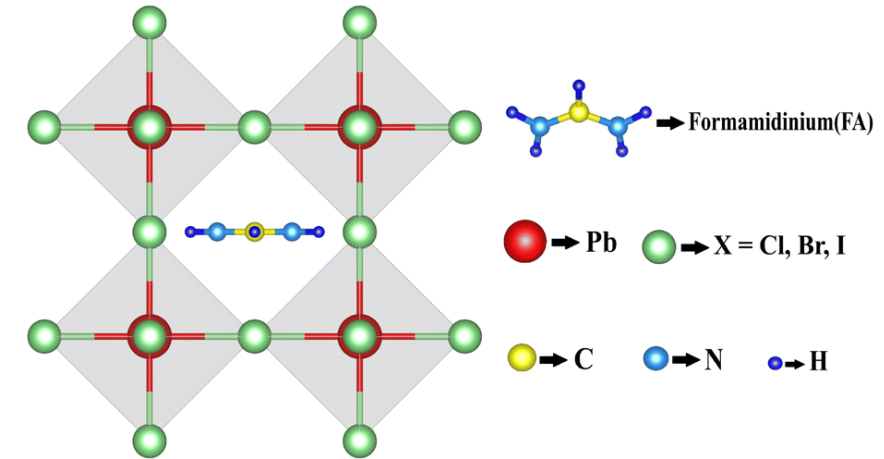
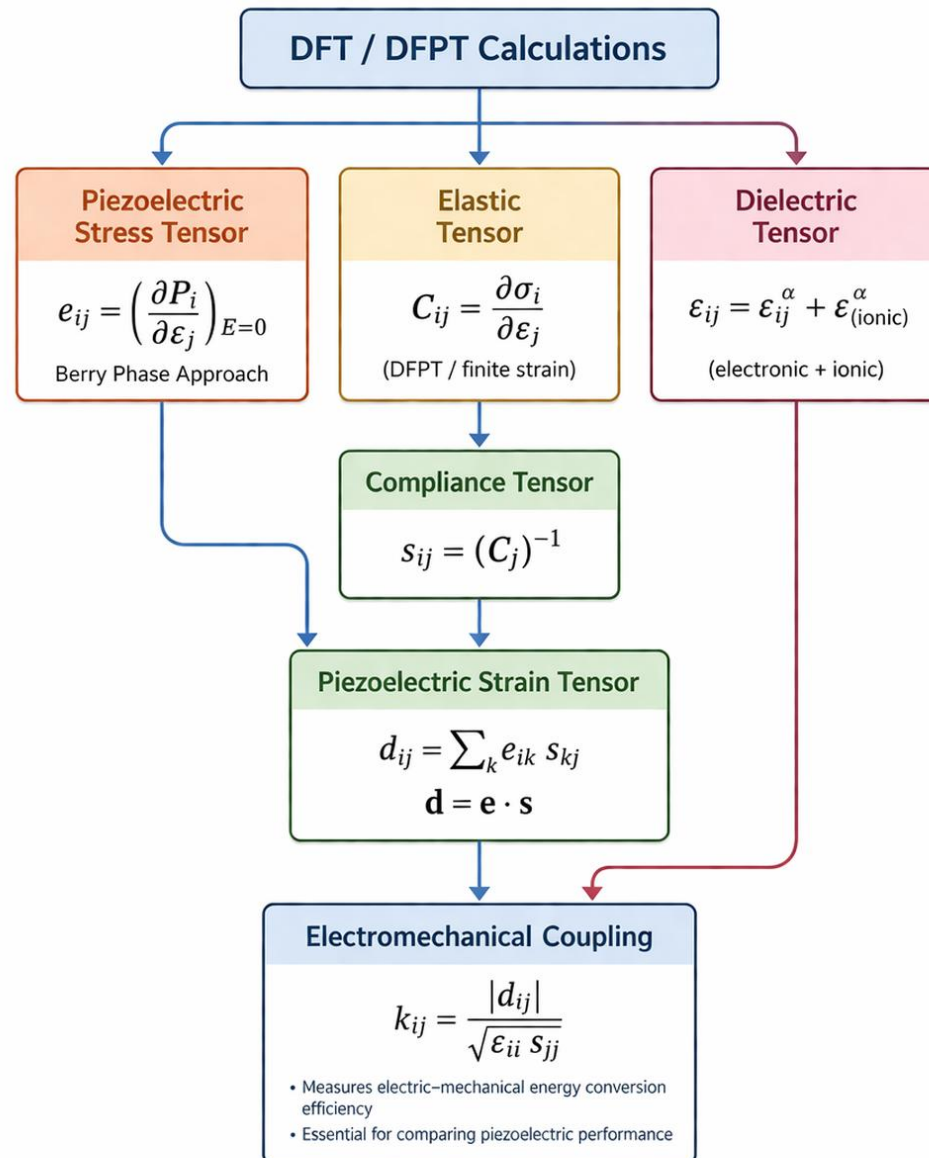
Organic cations used for the formation of 3D hybrid perovskites

A) Cations which were reported to form perovskites with Pb and Sn in combination with Cl, Br, and I.

B) Cations which were reported to form 3D hybrid perovskites only in combination with some selected metals and anions.

DFT simulation of HOIPs

DFPT = Density Functional Perturbation Theory



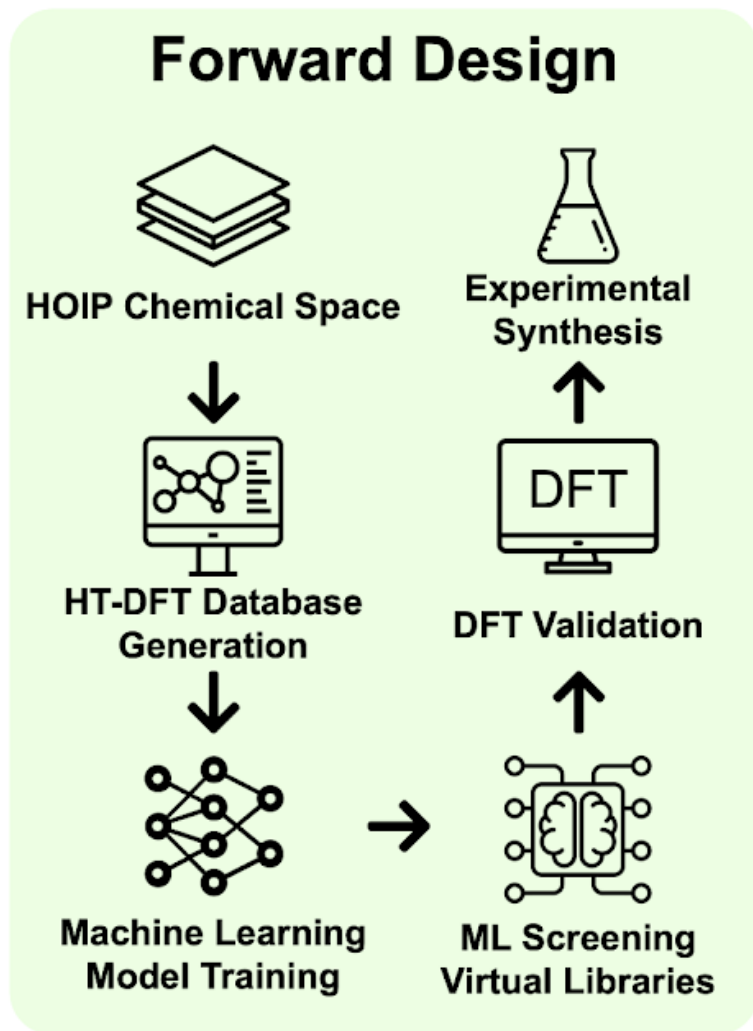
Structural representation of FAPbX₃ (X = Cl, Br, I).

Piezoelectrical response (in C m⁻²) and strain tensor (in pC N⁻¹) for FAPbX₃ (X = Cl, Br and I)

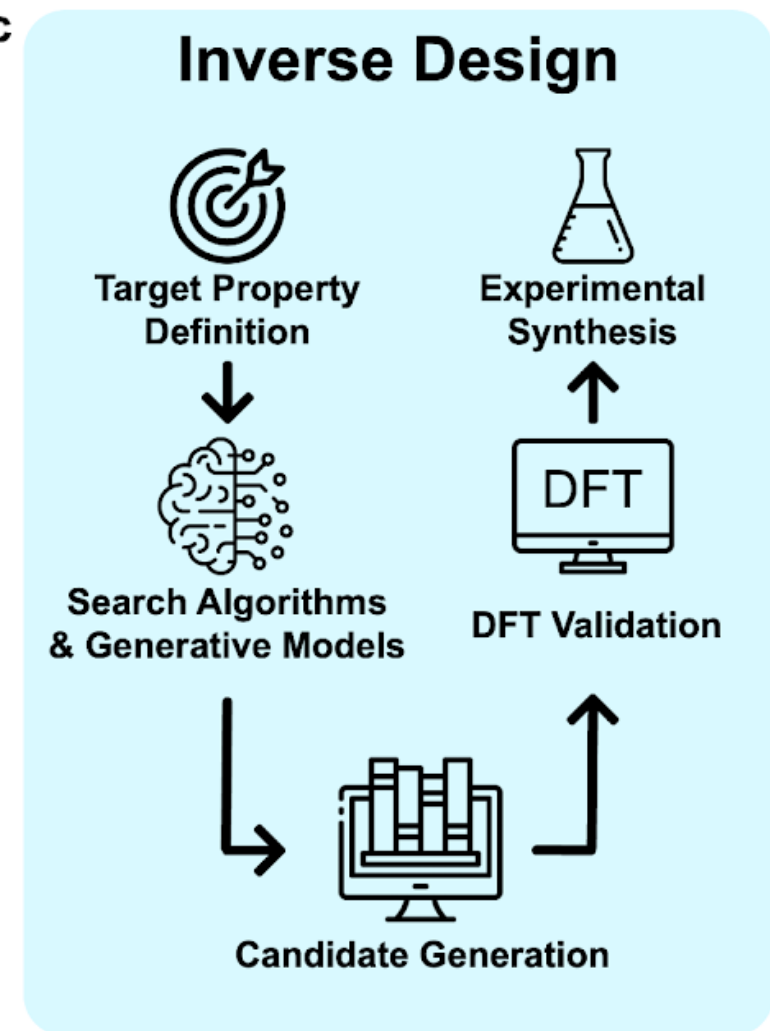
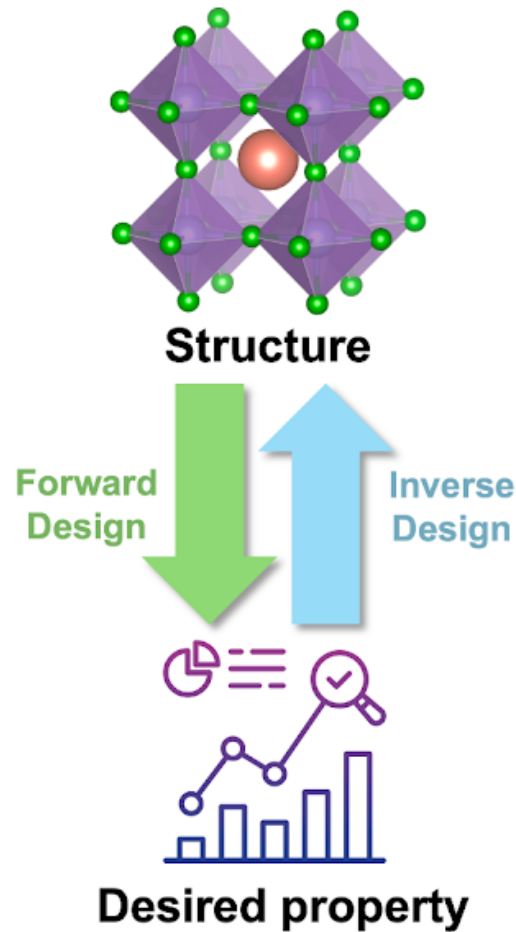
X	e_{16}	e_{25}	e_{31}	e_{32}	e_{33}	d_{33}
Cl	0.29	0.36	0.05	-0.2	-1.55	-28.99
Br	0.17	0.29	0.12	-0.03	-0.31	-6.95
I	0.43	0.37	0.24	-0.17	0.64	16.73

Chenkual, Lalthakimi, et al. "Unveiling the optical and piezoelectric properties of organic-inorganic hybrid perovskites FAPbX₃ (X = Cl, Br, I): a DFT study." *New Journal of Chemistry* 49.35 (2025): 15322-15332.

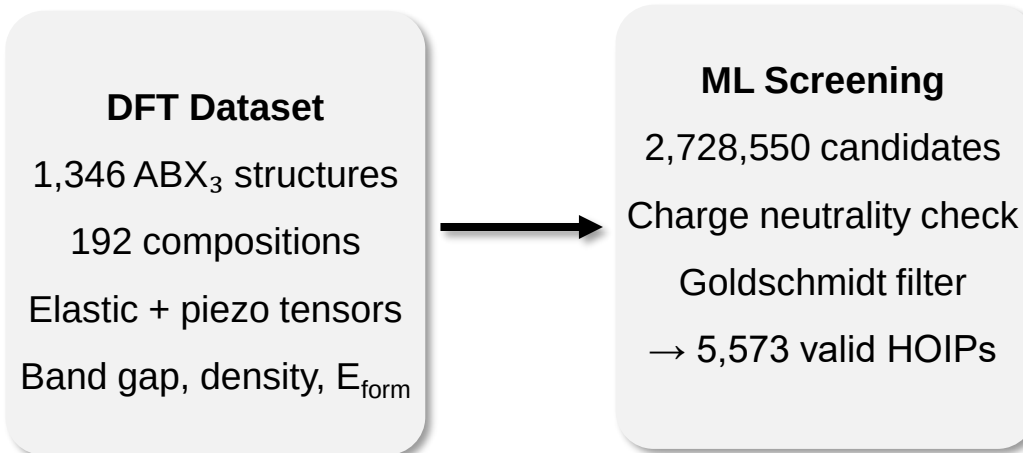
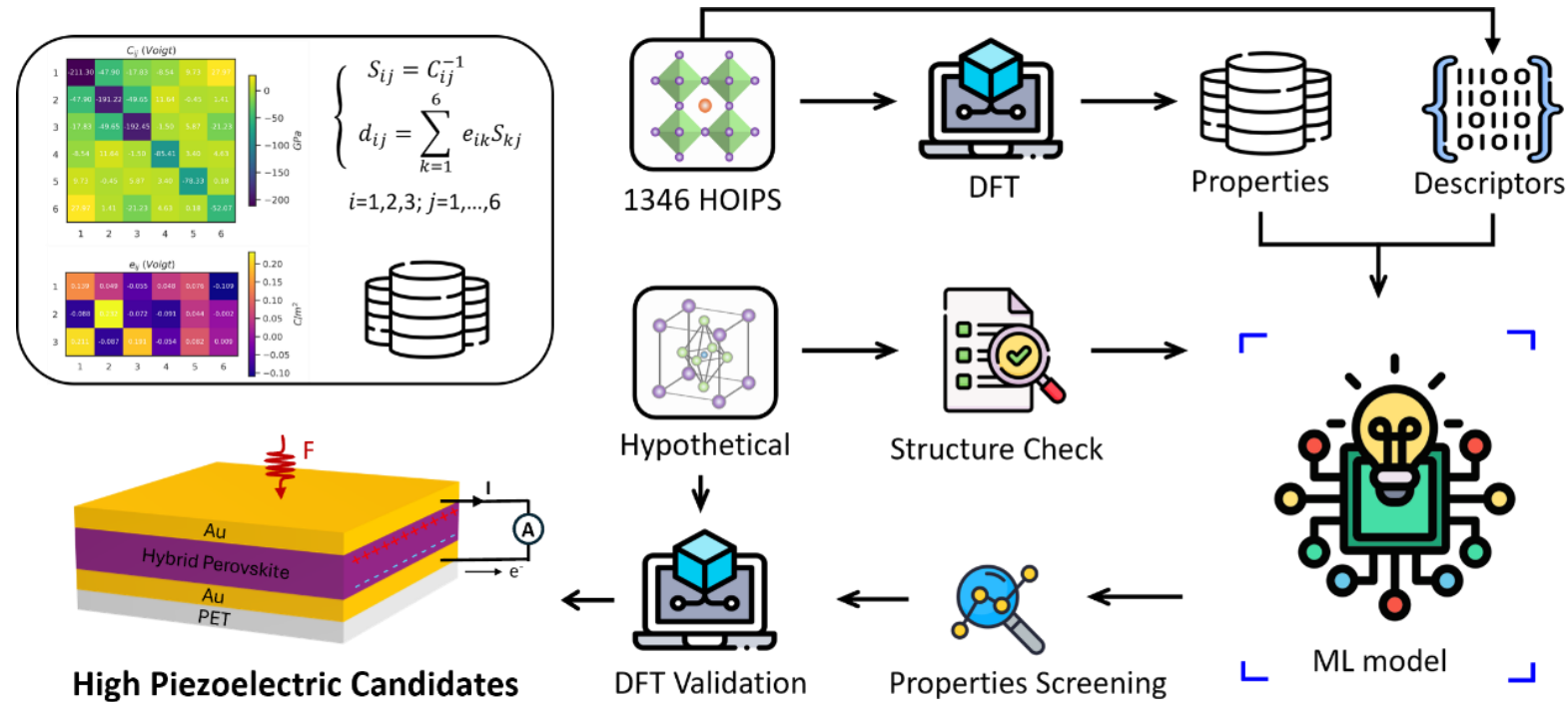
Forward and inverse design of HOIPs



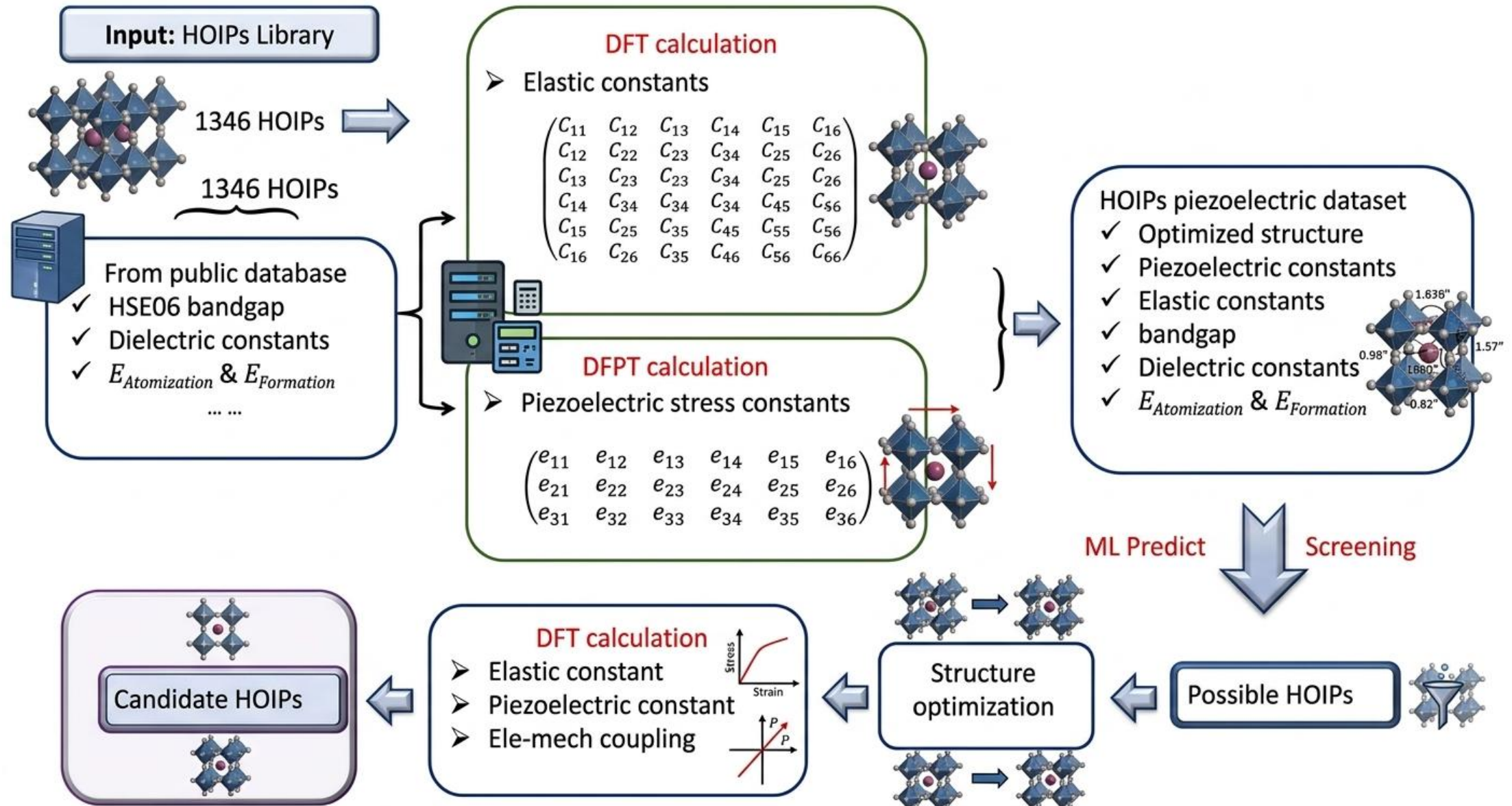
Hybrid Organic-Inorganic Perovskite (HOIP)



DFT-ML Workflow



Workflow of the high-throughput DFT



Rotation-Invariant Piezoelectric Descriptor

Problem with d_{33} alone

DFT relaxation yields HOIPs spanning cubic $\rightarrow$ triclinic symmetries. d_{33} depends on the crystallographic frame $\rightarrow$ biased comparisons across symmetry classes.

Solution: $\sigma_{\max}$

Largest singular value of the Kelvin-scaled 3 $\times$ 6 piezo strain tensor $\tilde{\mathbf{d}}$. Frame-independent (singular values invariant under orthogonal basis change).

Mathematical Definition ($\sigma_{\max}$)

Piezoelectric strain tensor d_{ij} (pC/N)

$$d_{ij} = \sum_{k=1}^6 e_{ik} \cdot S_{kj}$$

$$\begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} = \begin{pmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{pmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{pmatrix}$$

Kelvin-scaled matrix

$$\tilde{\mathbf{d}} = \mathbf{d} \text{diag}(1, 1, 1, 1/\sqrt{2}, 1/\sqrt{2}, 1/\sqrt{2})$$

$$\sigma_{\max} = \|\tilde{\mathbf{d}}\|_2 = \sqrt{\lambda_{\max}(\tilde{\mathbf{d}}^T \tilde{\mathbf{d}})} = \max_{\|\tilde{\boldsymbol{\sigma}}\|_2=1} \|\tilde{\mathbf{d}} \tilde{\boldsymbol{\sigma}}\|_2.$$

Orientation-Free

Comparable across all crystal symmetries

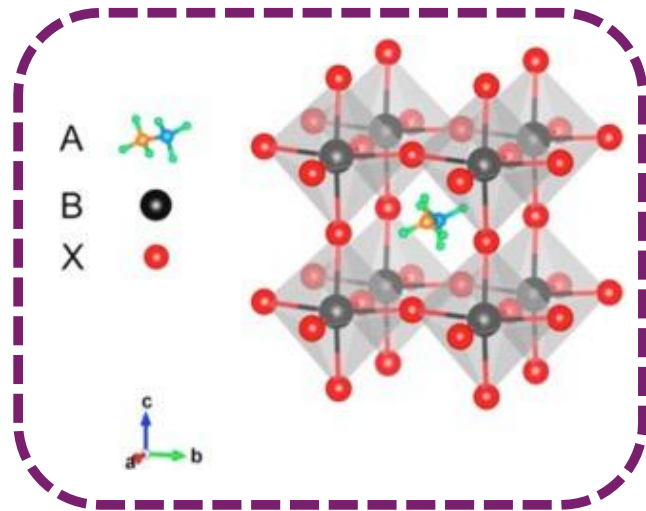
Physically Meaningful

Maximum polarization per unit stress

ML-Friendly

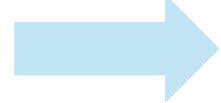
$\log_{10} \sigma_{\max} \rightarrow$ near log-normal distribution

ML models to predict HOIP properties



023333321
33111222244
03334432 331112222333
1411222243 22 022222343
23 1222225 02333332241 122222344
23322211344344443111243333330
01213533334444443245
533334444455556664
92 13444455556665
43 1444455556672
34234455556662
034555566653
01110

D
S
c
r
i
b
e



Coulomb matrix

36.9	23.3	14.3	9.3	14.3	9.3	14.3	9.3	23.3	36.9	23.3	14.3	23.3
------	------	------	-----	------	-----	------	-----	------	------	------	------	------

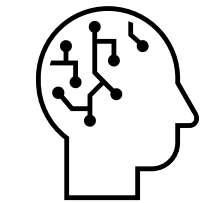
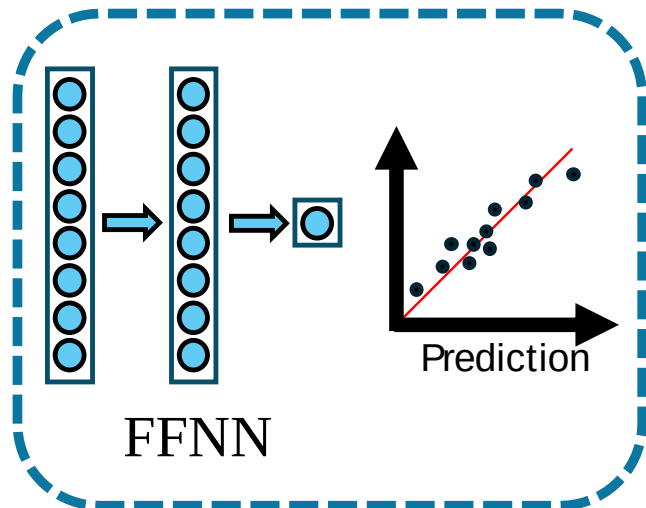
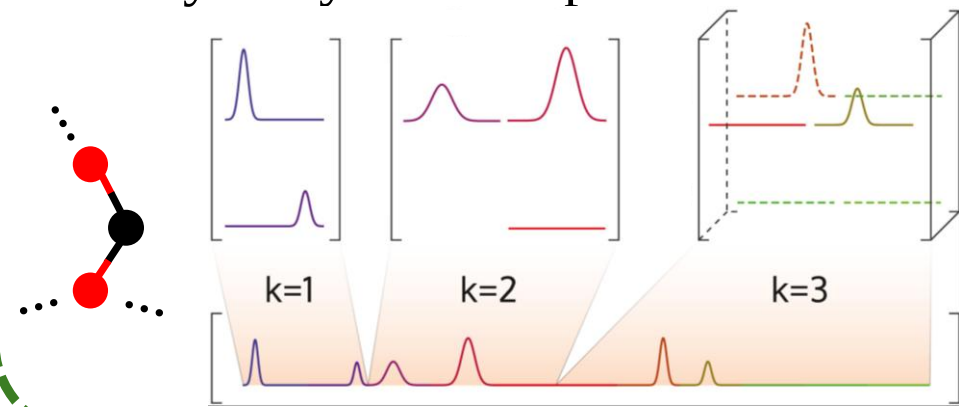
Ewald sum matrix

-14.3	-2.0	-5.9	-2.0	-5.9	-2.0	-5.9	-2.0	-2.0	-14.3	-2.0	-5.9	-2.0
-------	------	------	------	------	------	------	------	------	-------	------	------	------

Sine matrix

36.9	11.7	7.1	11.7	7.1	11.7	7.1	11.7	11.7	36.9	11.7	7.1	11.7
------	------	-----	------	-----	------	-----	------	------	------	------	-----	------

Many-body tensor representation



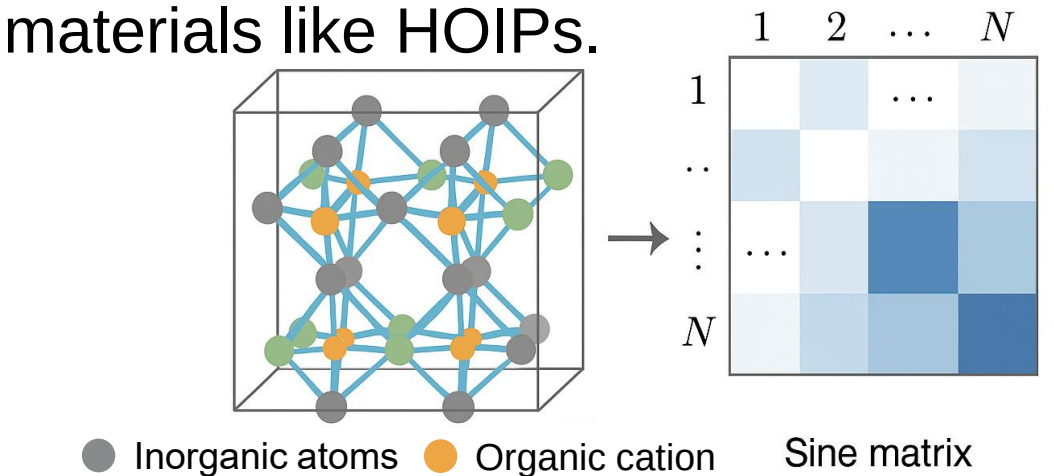
Machine
Learning



Descriptor: sine matrix

The **sine matrix descriptor** is a structural representation used in machine learning models to describe periodic materials like HOIPs.

- The **sine matrix** encodes **interatomic Coulombic interactions** in periodic crystals by considering **sine functions** of lattice positions.
- It is designed to respect the **periodicity** of crystal structures, making it more suitable than simple Coulomb matrices for solid-state systems.
- The descriptor is represented as a **square matrix**, where each element reflects interactions between atom pairs modulated by the lattice periodicity.
- It is **computationally efficient**, which makes it attractive for high-throughput screening.
- It captures **global periodic structure** but has **limited capability** in capturing **complex local chemical environments** or long-range interactions.



$$M_{ij}^{\text{sine}} =$$

$$\begin{cases} \text{for } i = j: & 0.5Z_i^{2.4}, \\ \text{for } i \neq j: & \frac{Z_i Z_j}{\left| \mathbf{B} \cdot \sum_{k=\{x,y,z\}} \hat{\mathbf{e}}_k \sin^2 \left(\pi \hat{\mathbf{e}}_k \mathbf{B}^{-1} \cdot (\mathbf{R}_i - \mathbf{R}_j) \right) \right|} \end{cases}$$

Z_i - nuclear charge
 $\mathbf{R}_i$ - position vector

$\mathbf{B}$ - lattice vectors matrix
 $\hat{\mathbf{e}}_k$ - cartesian unit vectors

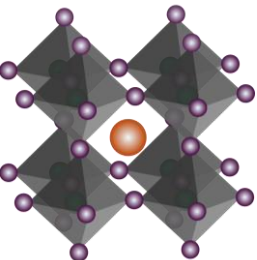
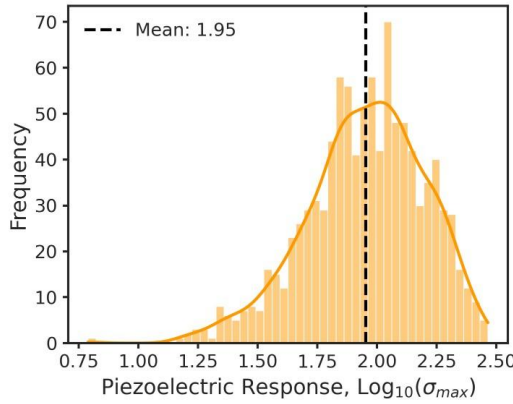
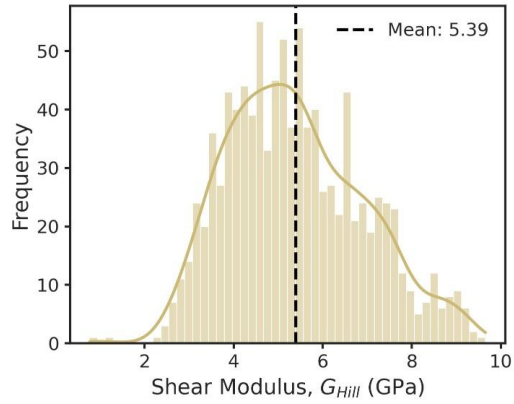
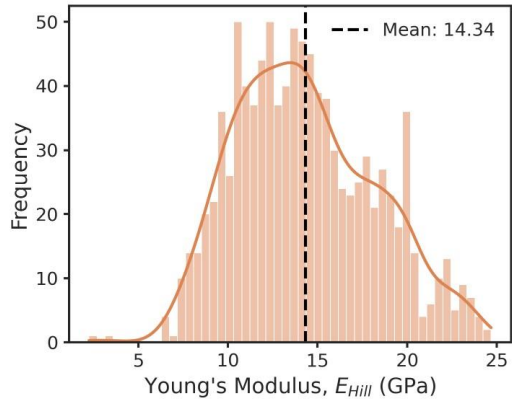
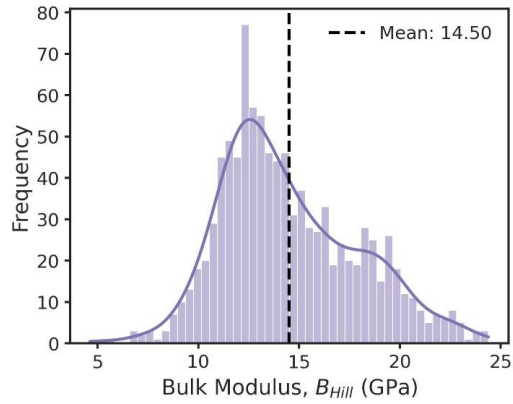
Dataset

No.	Property	Sources	Detail
#1	Atomization energy, Bandgap, Density, Dielectric constant (electronic), Dielectric constant (ionic), Dielectric constant (total), Refractive index, Relative energy1, Relative energy2, Volume of the unit cell	Scientific data [1]	1,346 structures including 192 ABX ₃ compositions
#2	Bulk modulus, Young's Modulus, Shear modulus, Piezoelectric descriptor	DFT Calculation	1,346 structures including 192 ABX ₃ compositions
#3	Bulk modulus, Young's Modulus, Shear modulus, Piezoelectric descriptor	DFT Calculation	59 structures selected from 5,573 structures (~3M ABX ₃ composition)

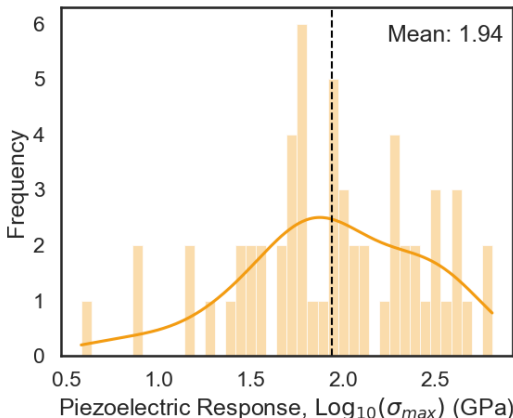
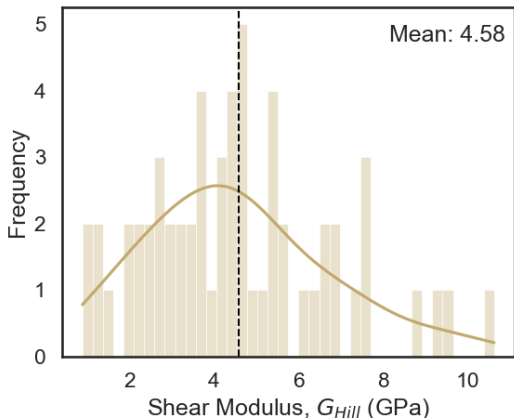
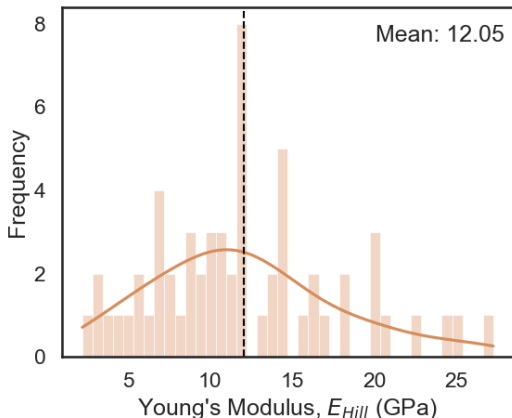
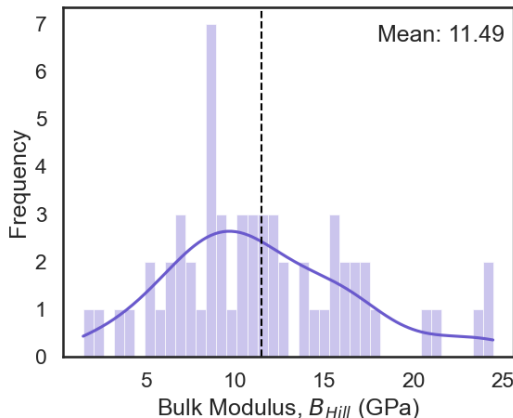
[1] Kim, Chiho, Tran Doan Huan, Sridevi Krishnan, and Rampi Ramprasad. "A hybrid organic-inorganic perovskite dataset." *Scientific data* 4, no. 1 (2017): 1-11.



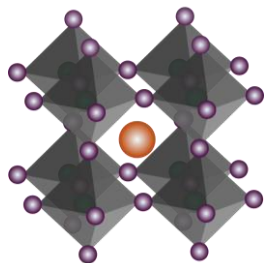
1,346 HOIPs
Scientific data [1]



59 HOIPs from
Candidate library



ML FNN Models

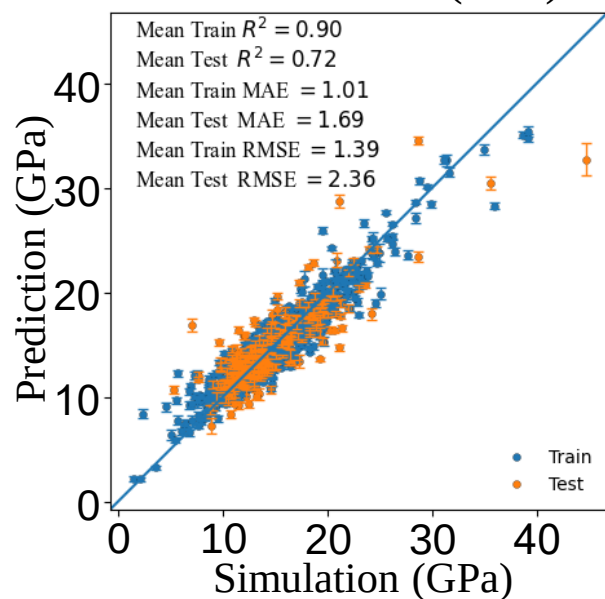


59 HOIPs
Candidate library
(5,573 in total)

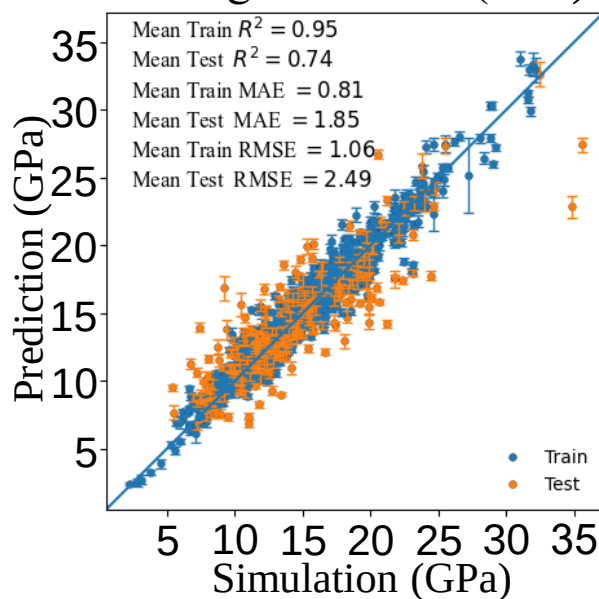
52 samples are introduced through data augmentation to compensate for the structural limitations of the Scientific Data dataset

7 samples for validation

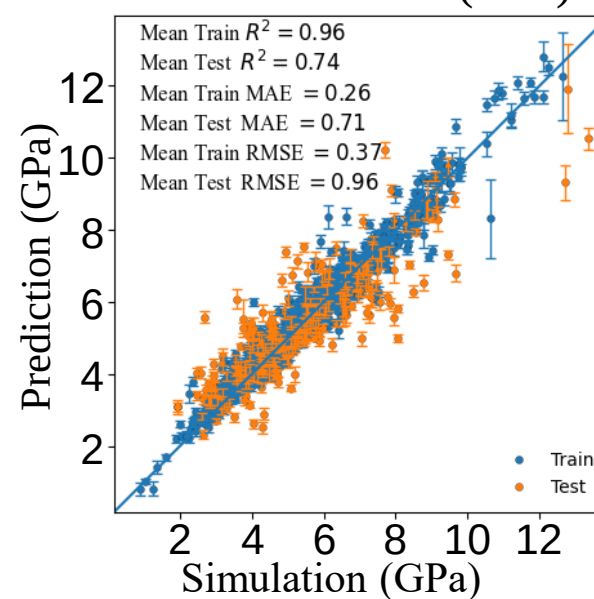
Bulk modulus (GPa)



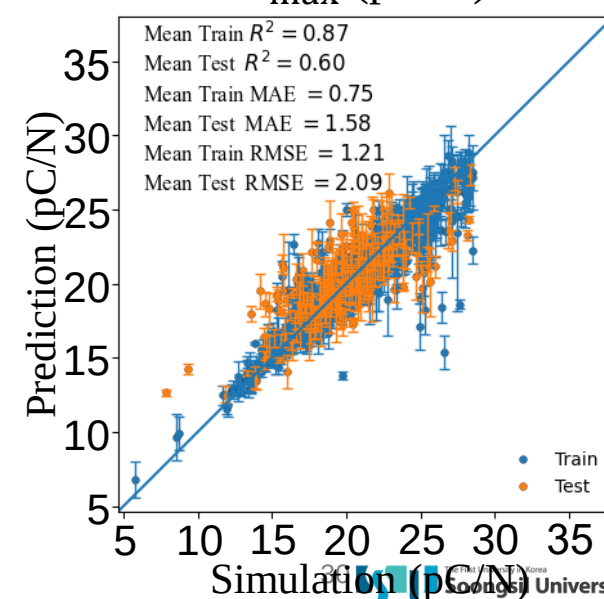
Young's modulus (GPa)



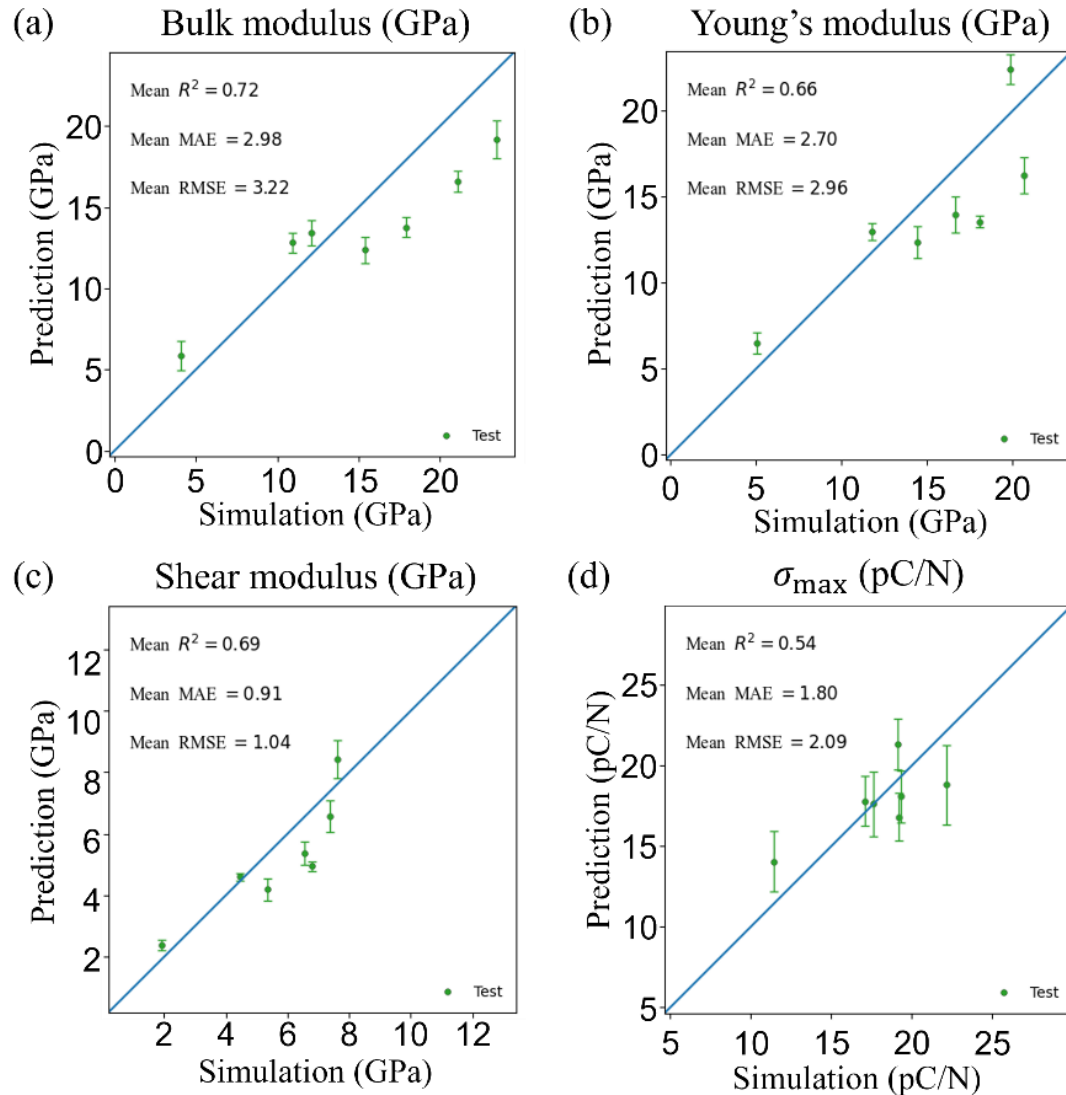
Shear modulus (GPa)



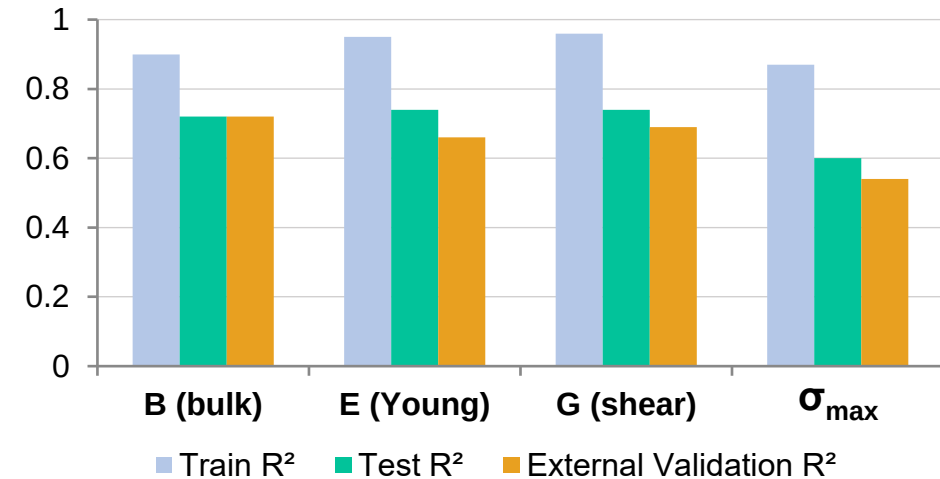
$\sigma_{\max}$ (pC/N)



ML Model performance



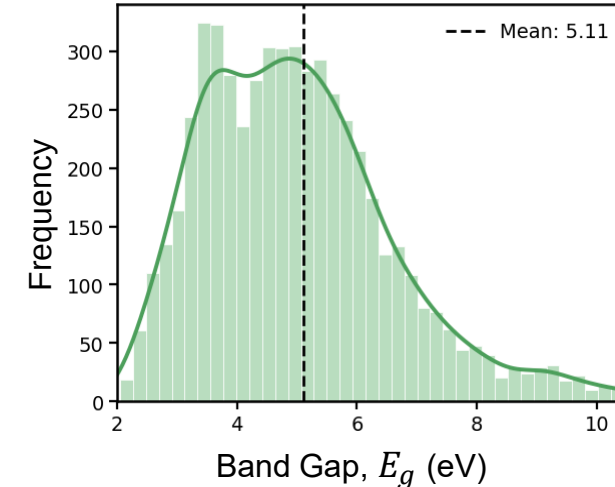
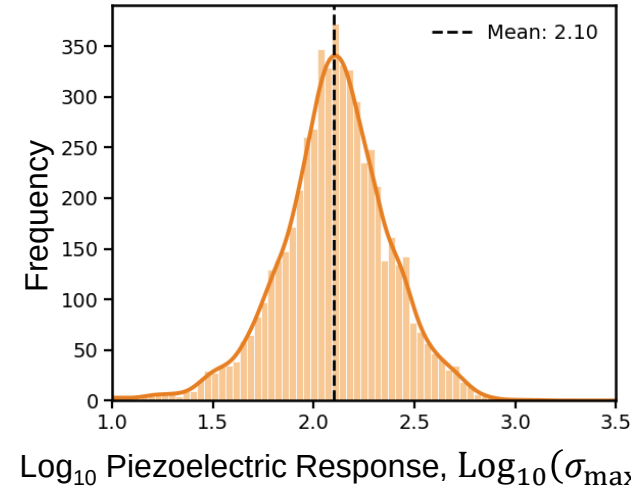
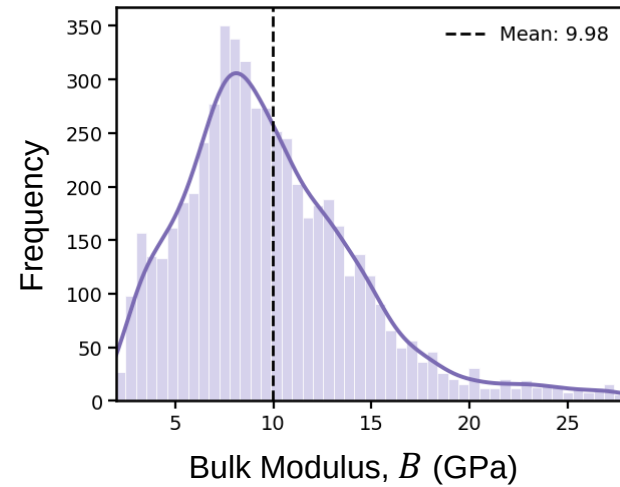
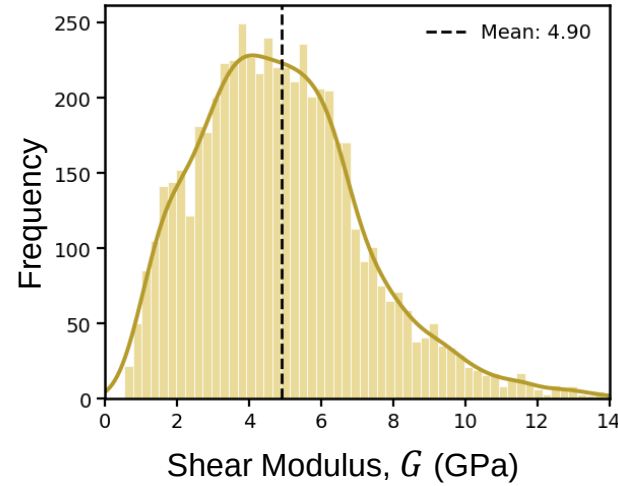
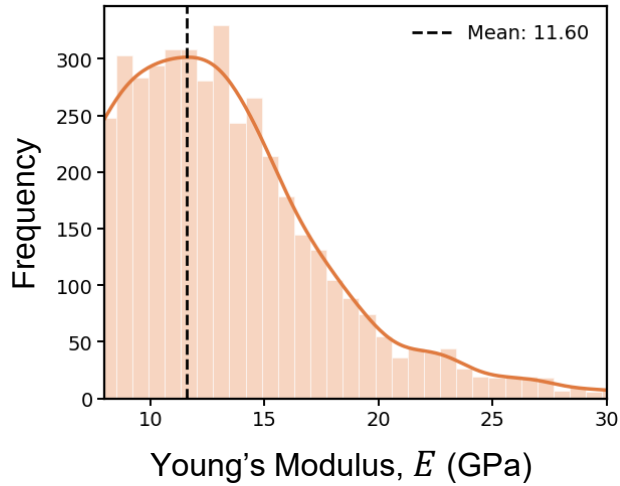
Mechanical Moduli & $\sigma_{\max}$



- ✓ $\sigma_{\max}$: moderate overfitting (train 0.87 vs test 0.60) due to sensitivity to local bonding & octahedral distortions
- ✓ External validation (7 unseen compositions): $R^2=0.54$ – $0.72 \rightarrow$ meaningful transferability confirmed
- ✓ Sine matrix + FNN ensemble outperforms ACSF, Ewald, MBTR for mechanical properties

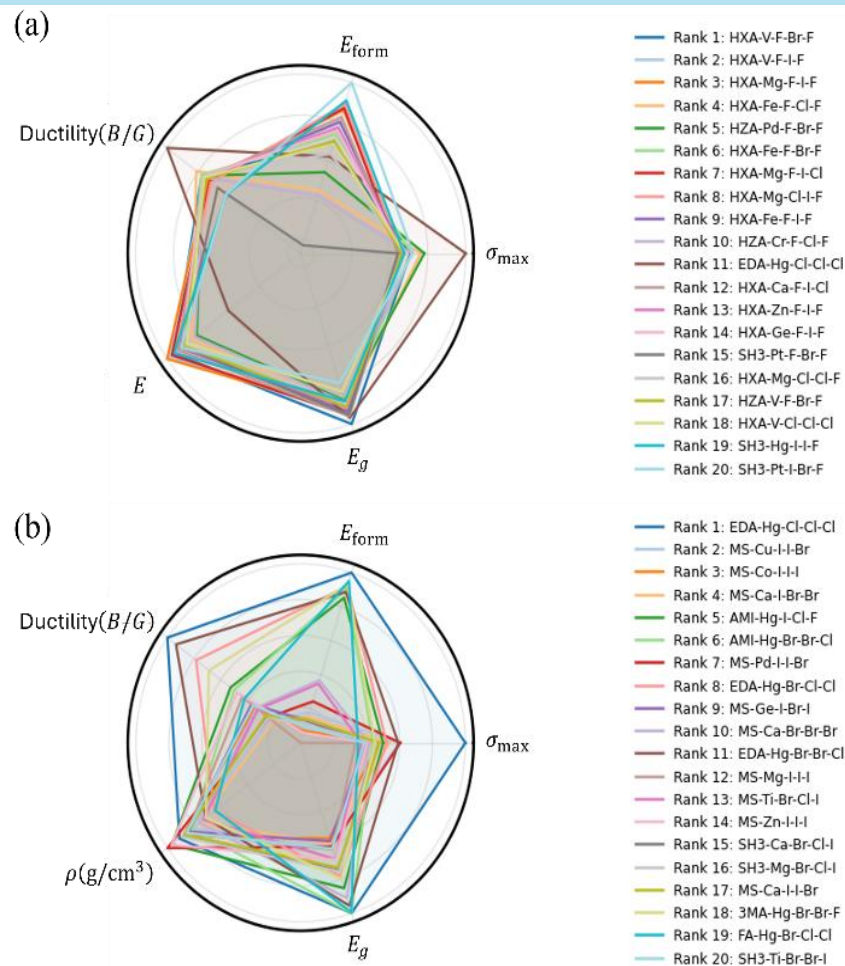
Fig. 7. External validation parity plots for (a) B , (b) E , (c) G , and (d) $\sigma_{\max}$ on seven representative candidate structures selected from the hypothetical ABX_3 library.

ML Screening



Property	Mean Value	Significance
E (Young) / G (Shear) / B (Bulk)	11.6 / 4.90 / 9.98 GPa	Mechanically soft compared to inorganic oxides and conventional ceramics
Pugh ratio B/G	2.32	Indicates ductile-like mechanical behavior
Density ρ	2.87 g/cm ³	Relatively low density compared to many inorganic oxide materials
$\sigma_{\max}$ (mean / top tier)	152.5 pC/N / >400 pC/N	High piezoelectric response for HOIPs
Band gap E_g	5.11 eV	Wide-gap semiconductor

Promising candidates



High-Performance Piezoelectrics

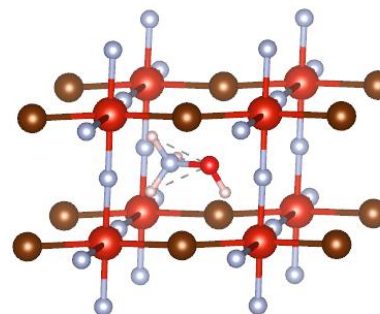
Objective:	High $\sigma_{\max}$ + stiffness (E) + stability
Top A-site:	HXA (hydroxylammonium)
Top B-site:	V^{2+} , Mg^{2+} , Fe^{2+} , Pd^{2+} , Ge^{2+}
Halides:	Mixed F / Cl / Br
#1 Candidate:	HXA-V-F-Br-F
B/G ratio:	~2.2 (moderate ductility)
Application:	Sensors, actuators, energy harvesters
Rigid polar framework $\rightarrow$ high electromechanical efficiency	

Novel Flexible Piezoelectrics

Objective:	High $\sigma_{\max}$ + ductility + low density
Top A-site:	EDA (ethylenediammonium), MS
Top B-site:	Hg^{2+} , Co^{2+} , Ca^{2+}
Halides:	Mixed Cl / Br / I
#1 Candidate:	EDA-Hg-Cl-Cl-Cl
B/G ratio:	~2.5 (high ductility)
Application:	Wearables, deformable electronics
Extended organic cations $\rightarrow$ mechanical adaptability	

Both categories satisfy: $E_{\text{form}} < 0$ (thermodynamic stability) $\cdot$ $E_g > 1.5$ eV (insulating character)

HXA-V-F-Br-F



EDA-Hg-Cl-Cl-Cl

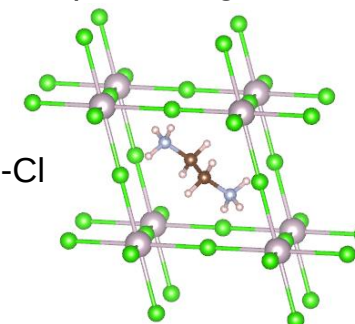


Fig. 9. Radar plots illustrating multi-objective screening of the top 20 HOIP candidates based on ML-predicted properties. (a) High-performance piezoelectric screening, optimized for large $\sigma_{\max}$, mechanical stiffness (E), and thermodynamic stability ($-E_{\text{form}}$). (b) Novel flexible piezoelectric screening, emphasizing $\sigma_{\max}$, ductility (B/G), low density (ρ), and stability for deformable device applications (left), as well as the atomic structures of the top-ranked HOIPs from each screening strategy (right).

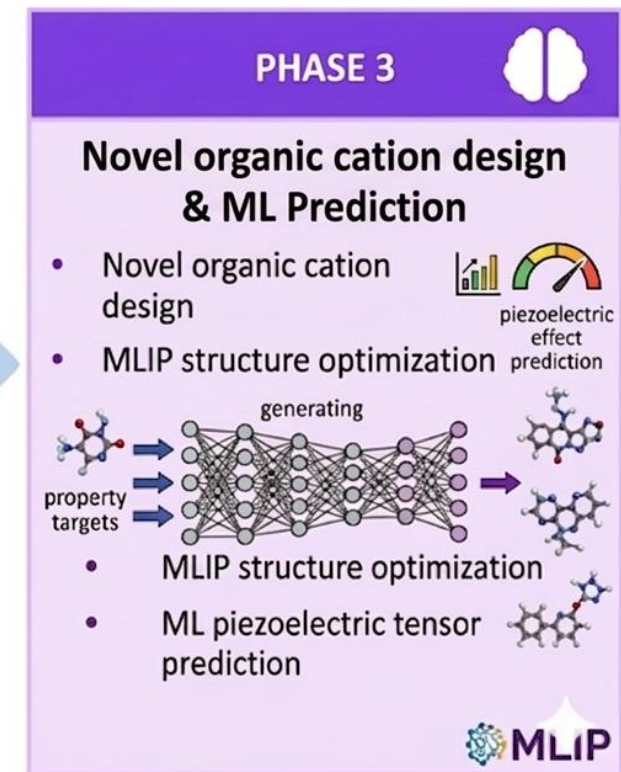
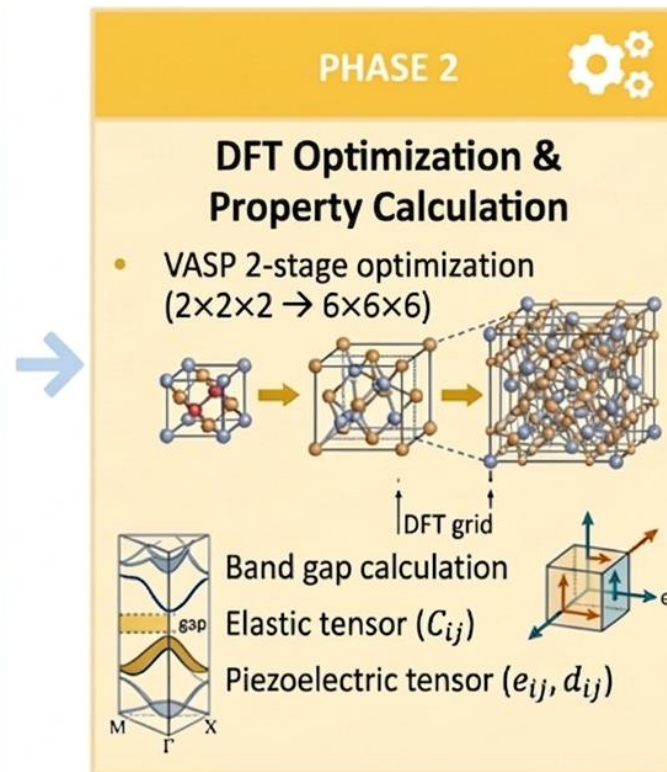
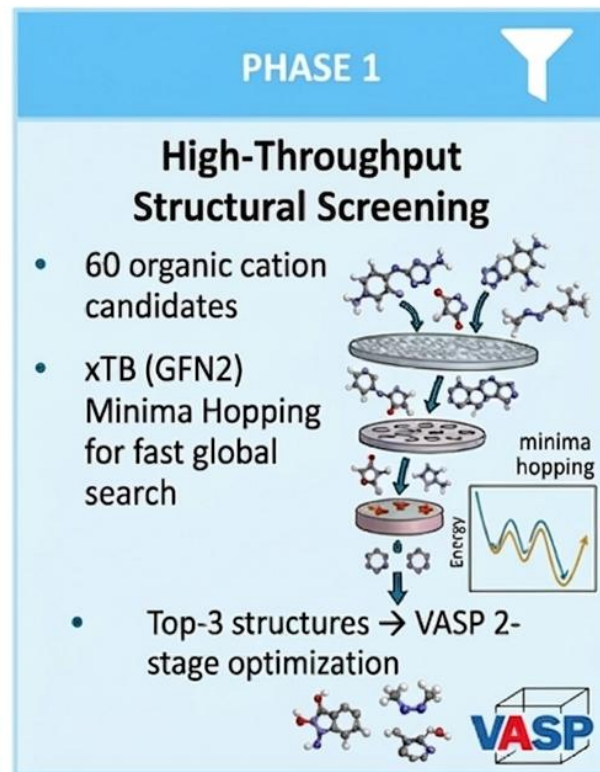
Outlook

- **Expand organic cation space**
→ Current study is limited to a small set of organic cations : broader chemical diversity may unlock new structure–property relationships
- **Direct tensor prediction instead of derived moduli**
→ Move from scalar/mechanical descriptors (e.g., B , G , $\sigma_{\max}$) to **direct prediction of piezoelectric and elastic tensors** for higher physical fidelity

Contents

1. **Introduction:** Perovskites, Piezoelectricity & How We Compute Them
2. **Topic 1:** Screening Millions of HOIPs with DFT + ML
3. **Topic 2: Inventing New Organic Cations**
4. **Topic 3:** Predicting Full Piezoelectric Tensors with Equivariant Networks
5. **Outlook**

Workflow

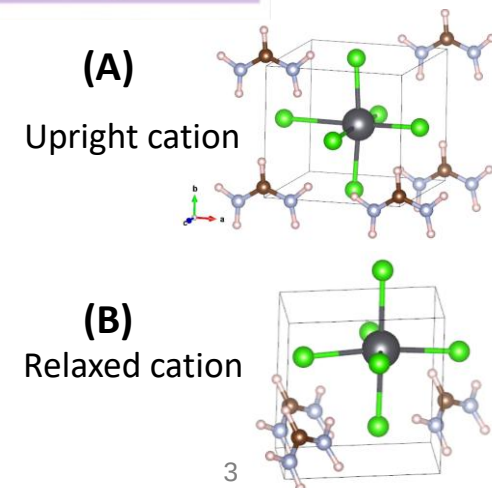


Assisted by AI

Organic Cation Orientation Critically Affects Elastic Property Calculations

FAPbCl₃

Elastic Tensor (C_{ij})	C_{11}	C_{12}	C_{13}	C_{22}	C_{23}	C_{33}	C_{44}	C_{55}	C_{66}	d_{33}
(A)	45.4869	12.297	10.2978	34.4206	9.2324	50.0444	-24.8757	-36.6142	5.4023	-0.578
(B)	21.8954	1.1158	5.9125	9.1841	1.1776	21.9387	3.5180	8.0370	3.4982	23.02



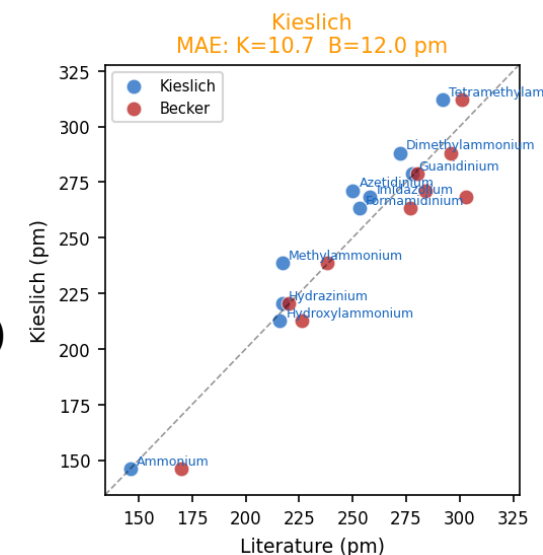
Organic Cation Library: 60 Candidates

Selection Criteria

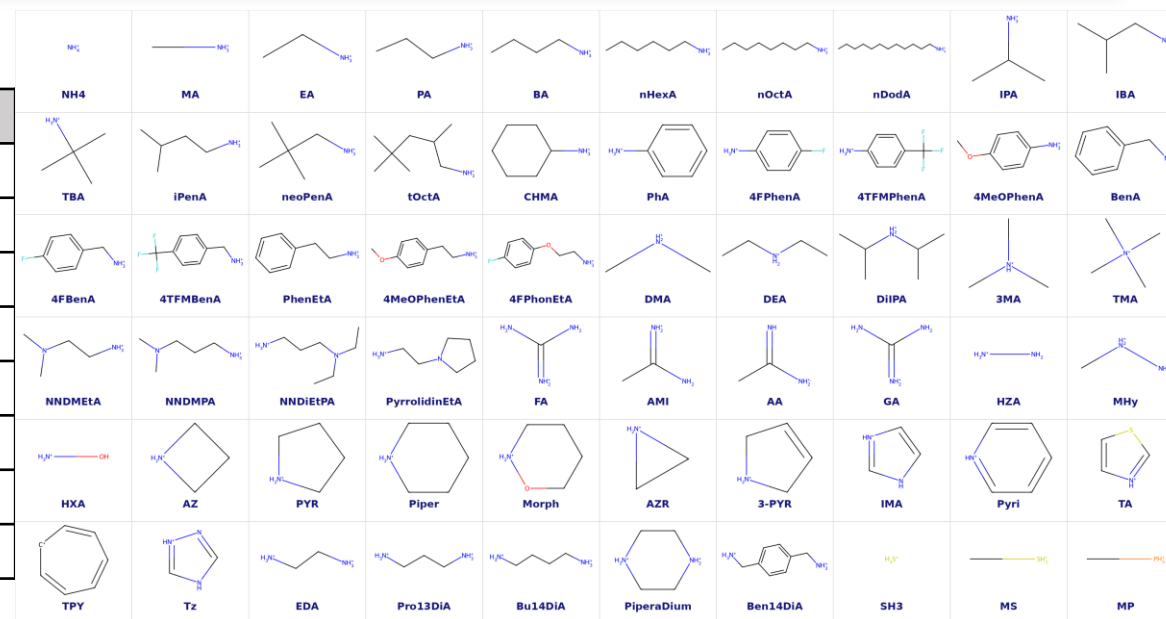
- Goldschmidt tolerance factor: 0.70 – 1.22
- Octahedral factor (μ): 0.4 – 0.9
- Charge: +1 (monovalent) and +2 (divalent) cations
- Charge neutrality: $qA + qB + 3qX = 0$
- Synthetic accessibility considered

Effective Radius: Kieslich Method

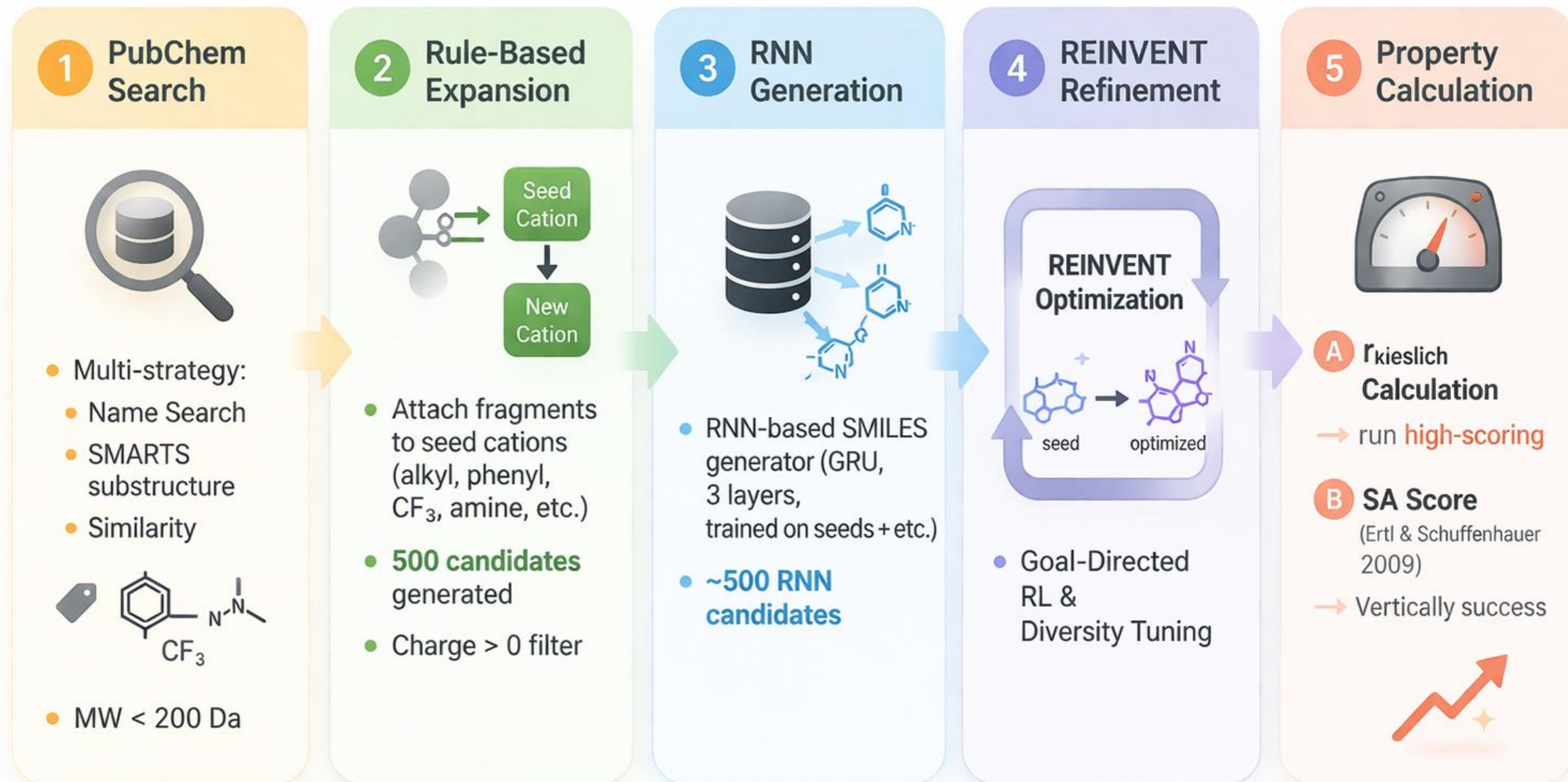
- $r_{\text{eff}} = r_{\text{mass}} + r_{\text{ion}}$ (Kieslich et al. 2014)
- RDKit multi-conformer generation (10 confs)
- Top-5 lowest energy conformer average
- MMFF force field optimization (UFF fallback)
- B/X radii: pymatgen $\rightarrow$ reference value fallback



Category	Count	Examples
Primary amines (linear/branched)	13	MA, EA, PA, BA, IPA, TBA, neoPenA ...
Aromatic / Benzylic / Phenethyl	9	PhA, BenA, PhenEtA, 4FPhenA, 4FBenA ...
Secondary / Tertiary / Quaternary	5	DMA, DEA, DiIPA, 3MA, TMA
Formamidinium / Hydrazinium	7	FA, AMI, AA, GA, HZA, MHy, HXA
Cyclic (saturated / aromatic)	10	AZ, PYR, Piper, IMA, Pyri, TA, Morph ...
Tert-amine primary NH_3^+	4	NNDMEtA, NNDMPA, NNDiEtPA, PyrrolidinEtA
Diammonium (+2)	5	EDA, Pro13DiA, Bu14DiA, PiperaDium, ...
Sulfonium / Phosphonium / Others	7	SH3, MS, MP, NH4, CHMA, nHexA, nOctA ...



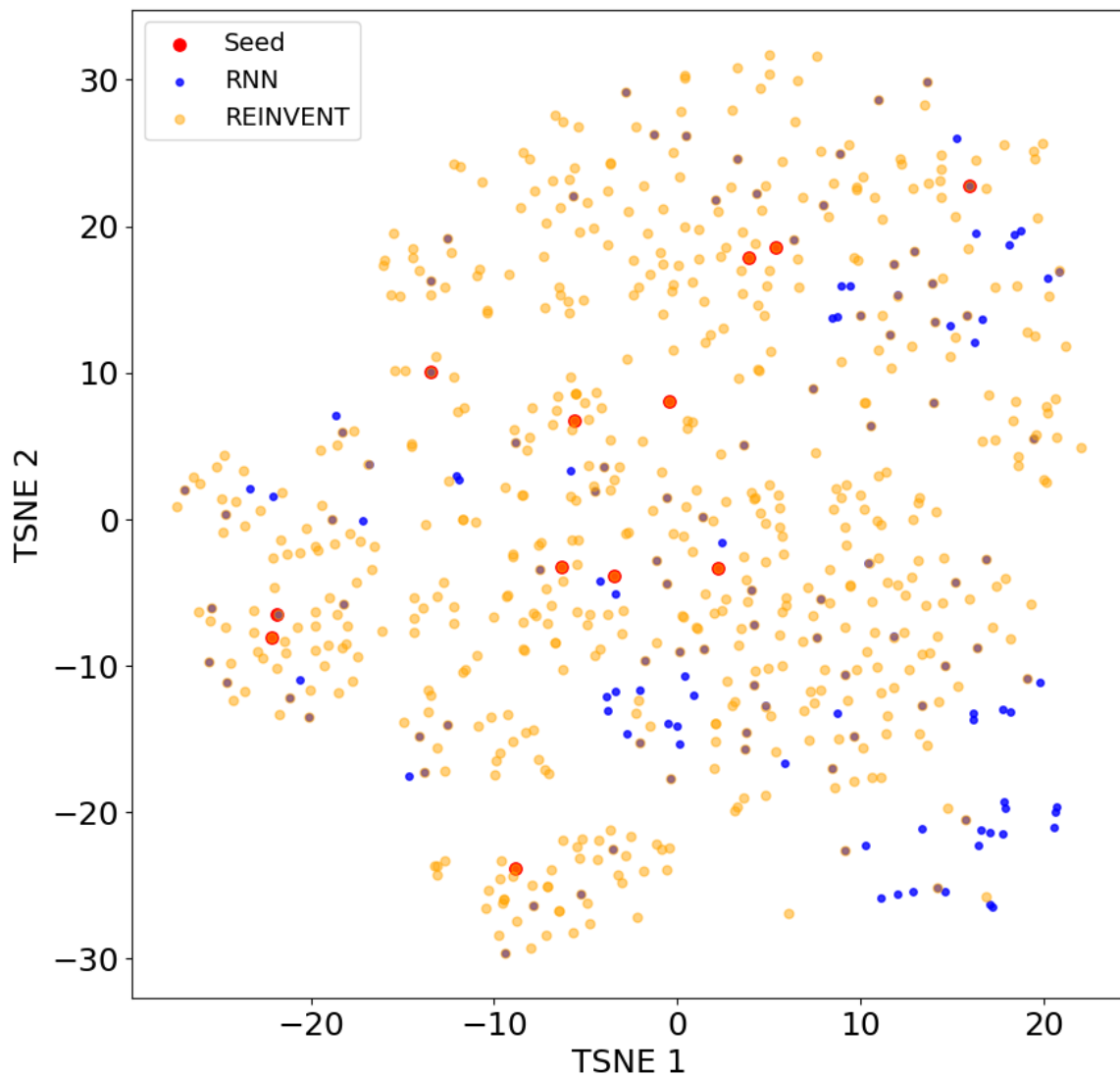
Novel Organic Cation Workflow



Assisted by AI

Novel Organic Cation

Generated cations fill new regions of chemical space

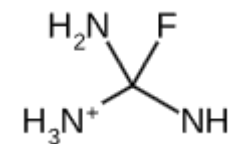
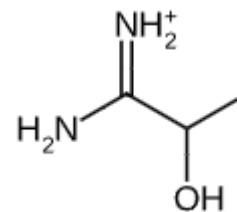
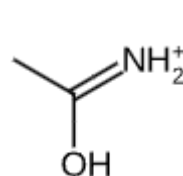
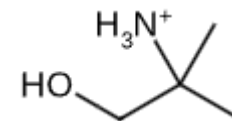
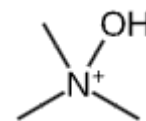
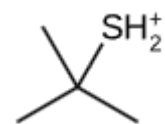


Filtering criteria

- ✓ Ionic radius < 3.4 (decided from seed cations in Goldschmit tolerance factor)
- ✓ SA (Synthetic Score) < 6

- 1) Seed 12 filtered among 60 cations
- 2) After RNN 143
- 3) After REINVENT 598

TSNE plot contains only structures after filtering



Expand Organic Cations

Seed A-site

60 organic cations

~890 ABX_3

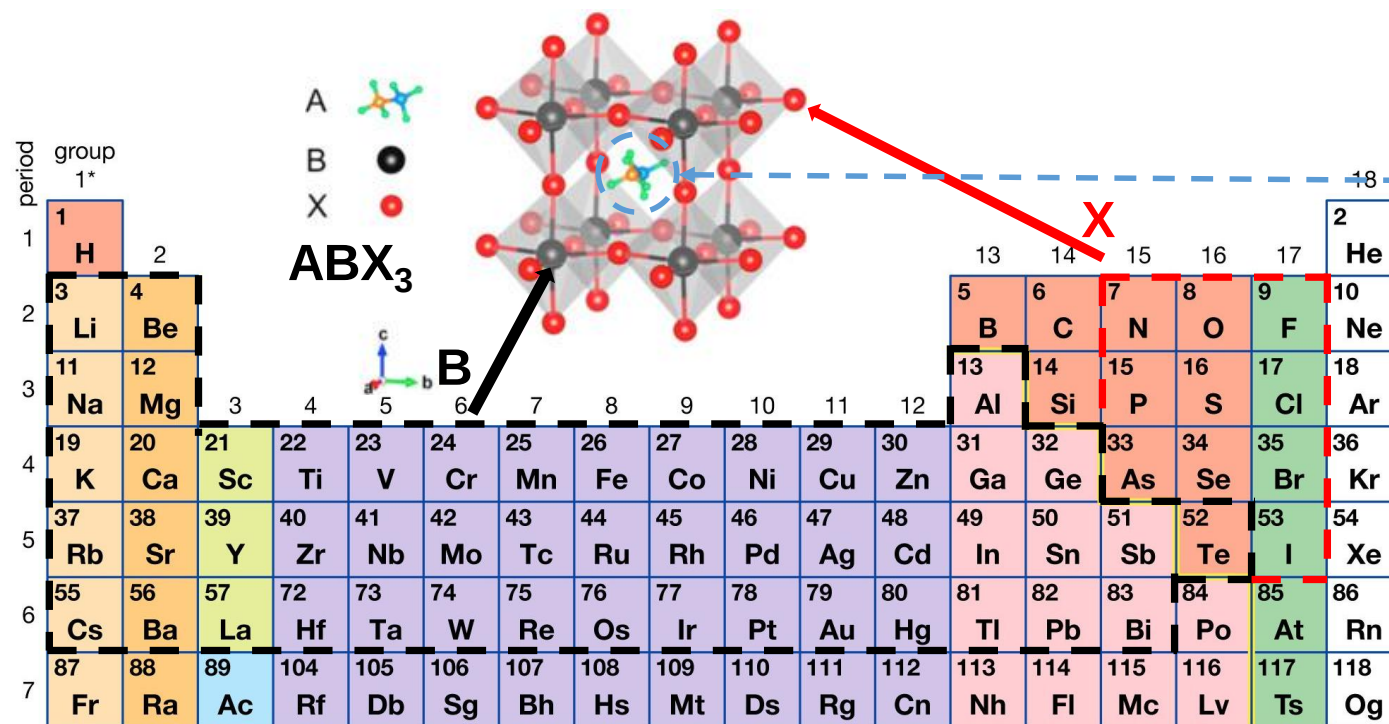


Same B, X
Chemical space

New A-site

60 organic cations

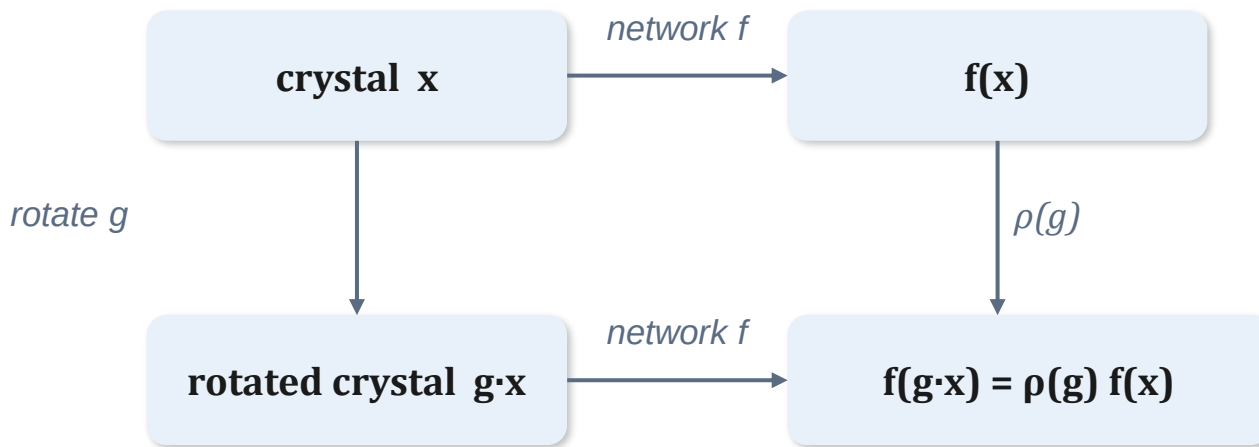
~20,000 ABX_3



Contents

1. **Introduction:** Perovskites, Piezoelectricity & How We Compute Them
2. **Topic 1:** Screening Millions of HOIPs with DFT + ML
3. **Topic 2:** Inventing New Organic Cations
4. **Topic 3: Predicting Full Piezoelectric Tensors with Equivariant Networks**
5. **Outlook**

Symmetry Is Free Data: Equivariant Networks



$$f(g \cdot x) = \rho(g) f(x) \text{ for every } g \in E(3)$$

Why it matters

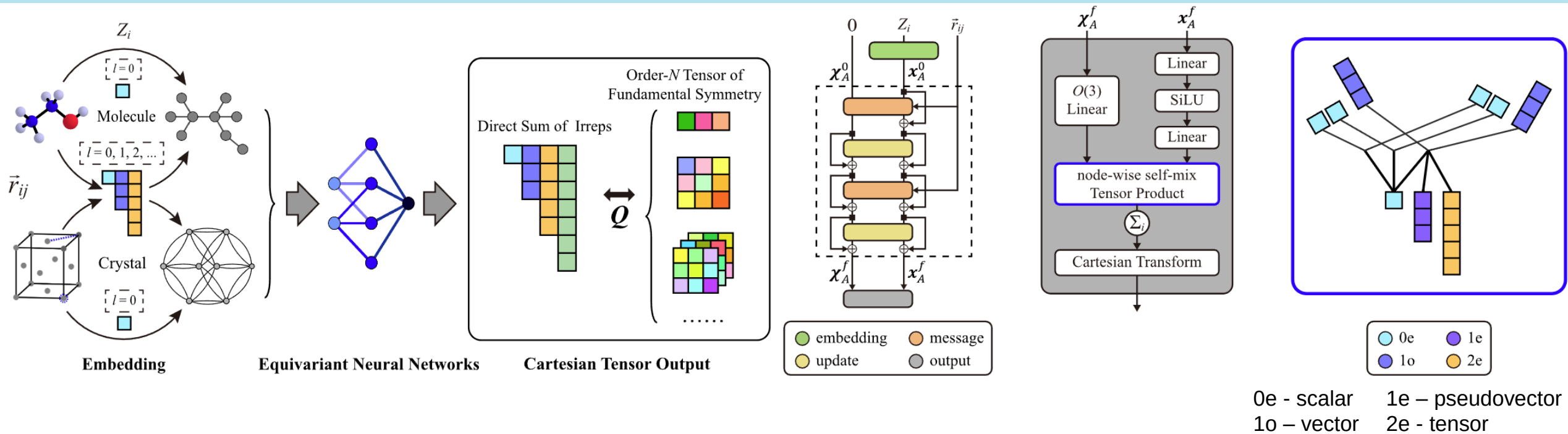
- A rotated crystal is the same crystal — the network never has to learn this from data.
- Data augmentation samples a few rotations; equivariance gives all of them for free.

Physical quantities = representations of $SO(3)$

Quantity	Irreps (l)	Dimension
Energy, band gap, $\sigma_{\max}$	0	1
Force, polarization	1	3
General rank-2 tensor	$0 \oplus 1 \oplus 2$	$1+3+5 = 9$
Piezoelectric tensor (rank 3)	$1 \oplus 1 \oplus 2 \oplus 3$	$3+3+5+7 = 18$
Elastic tensor (rank 4)	$0 \oplus 0 \oplus 2 \oplus 2 \oplus 4$	$1+1+5+5+9 = 21$

Tensor-output heads assemble the answer from these blocks (e3nn notation: 0e, 1o, 1e, 2e, ...) — see XPaiNN on the next slide.

ML model methods

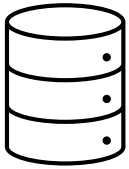


- Schematic of learning and predicting the Cartesian tensor with the present **XPaiNN model**.
- General workflow of **predicting the Cartesian tensor with an equivariant GNN**.
- Model architecture of XPaiNN.
- Cartesian tensor output module designed in this work, which takes both invariant and equivariant node-hidden representations from the message passing block as input and passes them into the node-wise self-mix layer (marked within the blue frame) with suitable transformations applied.
- Illustration of node-wise self-mix tensor products, the input $2 \times 0e + 1 \times 1o$ Irreps feature χ forms the output of $1 \times 0e + 1 \times 1e + 1 \times 2e$, representations for an order-2 tensor of arbitrary symmetry, each line in bold denotes trainable weights transformed from χ for each legal tensor product path. The Irreps are given in e3nn notation

Elastic tensor & Piezoelectric tensor

Mechanical Elasticity tensor

10,276 samples



Materials Tensor
(MatTen)

Wen, Mingjian, et al. *Digital Discovery* 3.5 (2024): 869-882.

Elasticity tensor $\mathcal{C}$ [GPa]

36 components:

$$C_{ijkl}$$

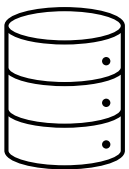
Symmetries:

$$C_{ijkl} = C_{jikl},$$

$$C_{ijkl} = C_{ijlk}$$

Piezoelectricity Piezo stress tensor

4,884 samples



JARVIS-DFT

Choudhary, Kamal, et al. *npj computational materials* 6.1 (2020): 173.

Piezo stress tensor e [C/m^2]

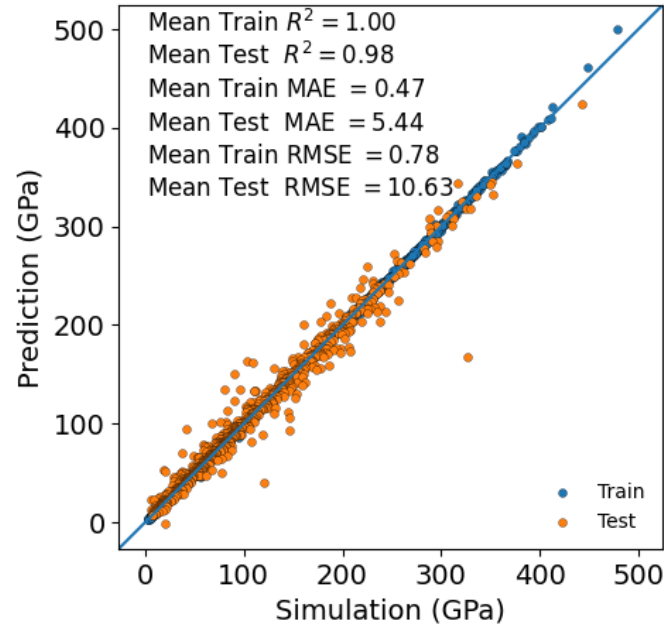
27 components:

$$e_{ijk}$$

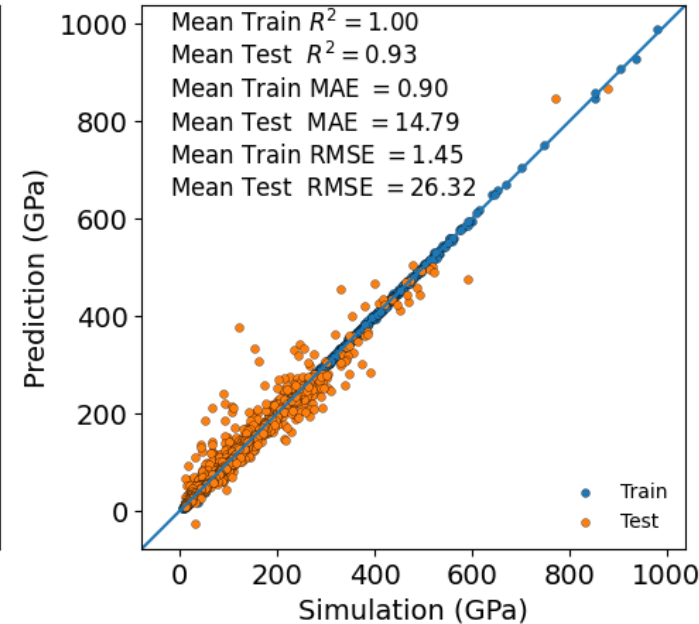
Symmetries:

$$e_{ijk} = e_{ikj}$$

Bulk Modulus



Young's Modulus



Frobenius norm distance

$$F_{\text{norm}} (\downarrow) = 0.24$$

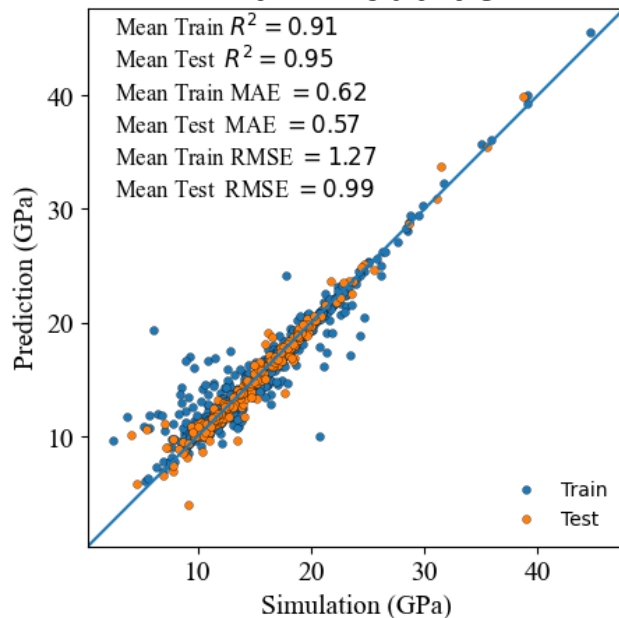
$$d_k = \left\| \mathbf{Y}_{\text{pred}}^{(k)} - \mathbf{Y}_{\text{true}}^{(k)} \right\|_F$$

$$= \sqrt{\sum_{i,j} \left(y_{ij}^{\text{pred}} - y_{ij}^{\text{true}} \right)^2}$$

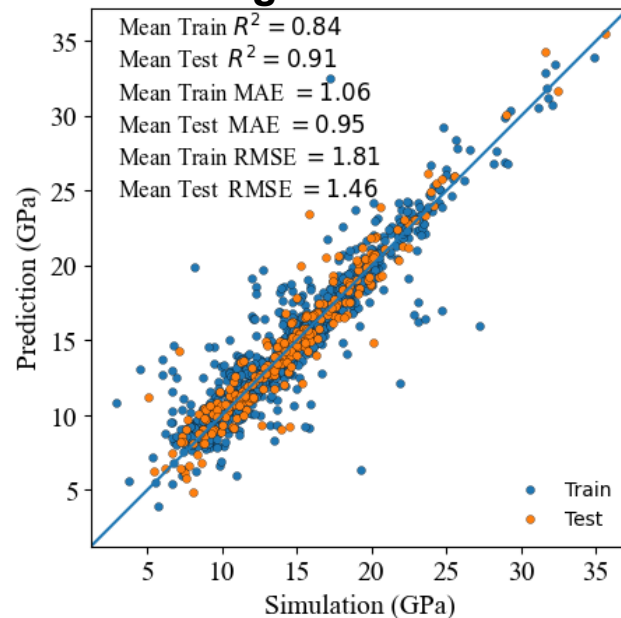
$$F_{\text{norm}} = \frac{1}{N} \sum_{k=1}^N d_k$$

XequiNet Tensor

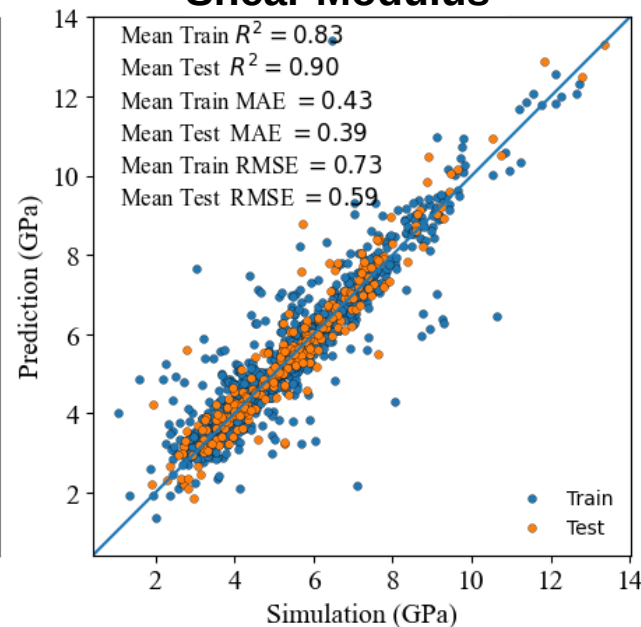
Bulk Modulus



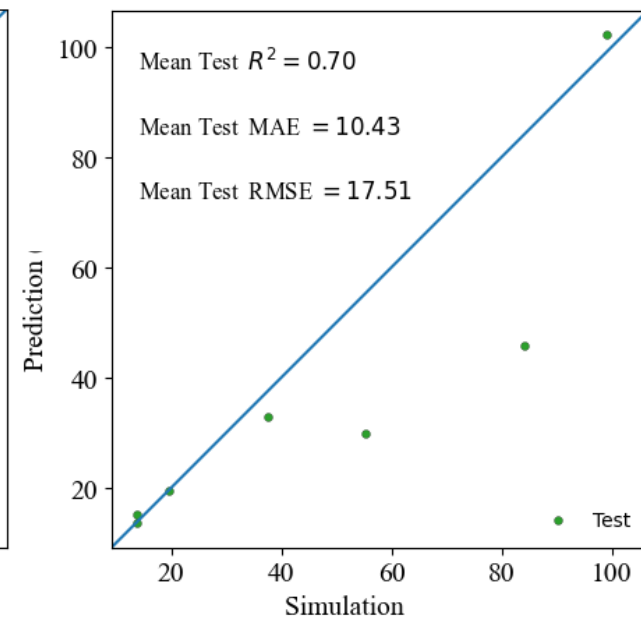
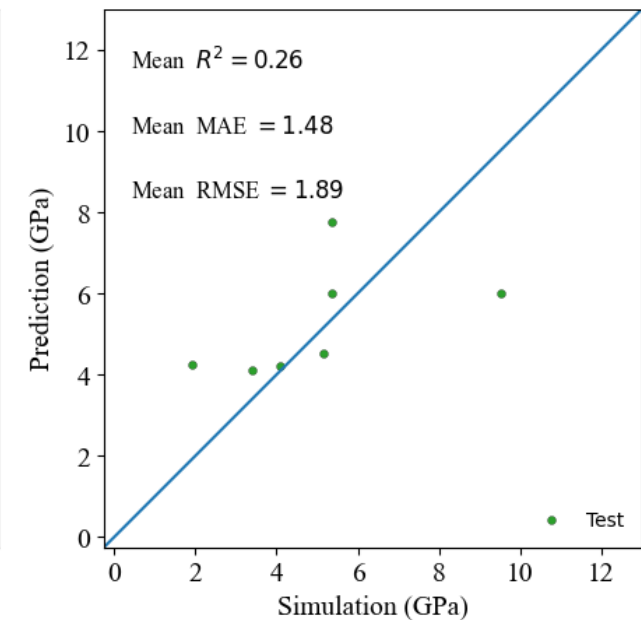
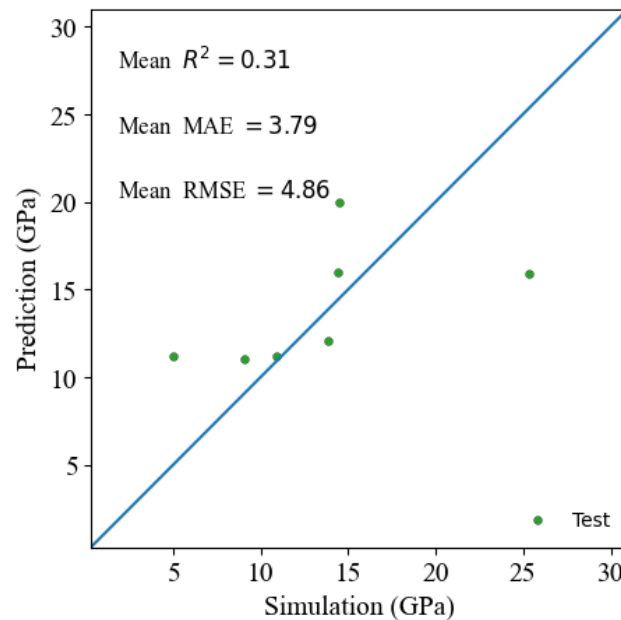
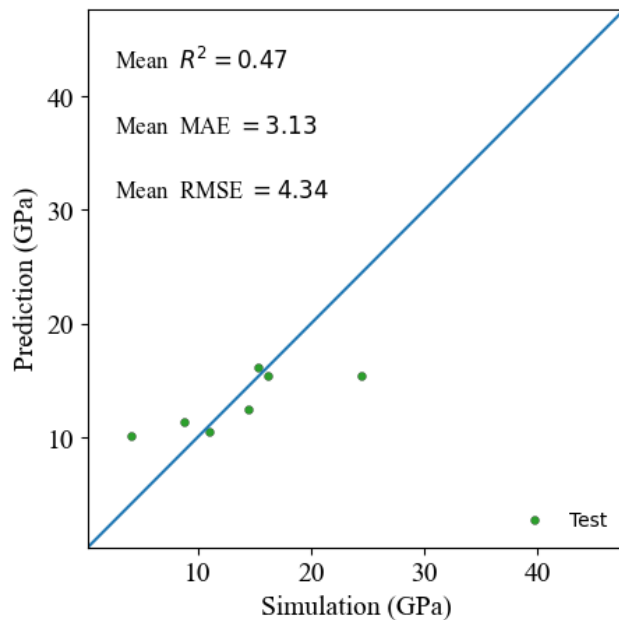
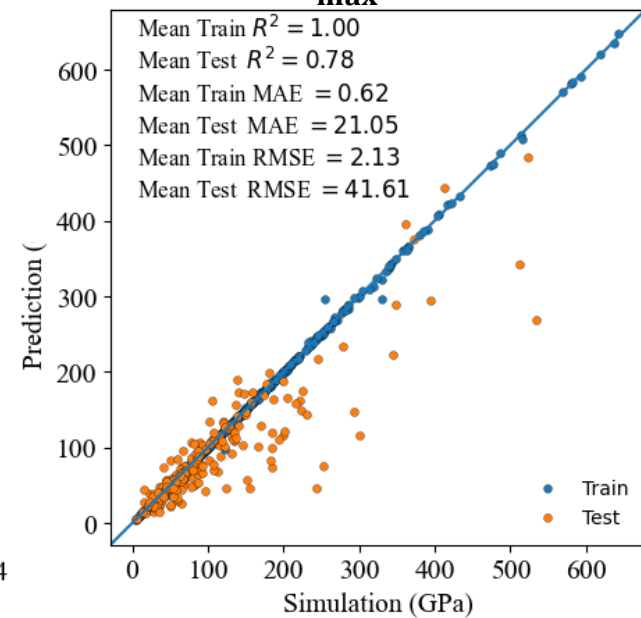
Young's Modulus



Shear Modulus



$\sigma_{\max}$



Contents

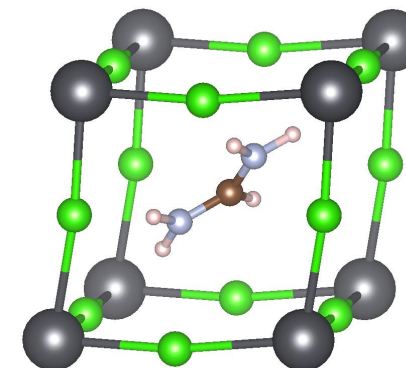
- 1. Introduction:** Perovskites, Piezoelectricity & How We Compute Them
- 2. Topic 1:** Screening Millions of HOIPs with DFT + ML
- 3. Topic 2:** Inventing New Organic Cations
- 4. Topic 3:** Predicting Full Piezoelectric Tensors with Equivariant Networks
- 5. Outlook**

Future Work

Current limitation

Cation	Fast motion (wobbling-in-cone)	Slow motion (jump-like reorientation)
MA ⁺ (MAPbI ₃)	~300 fs	~3 ps
FA ⁺ (FAPbI ₃)	~300 fs	~3-5 ps

- Organic cations exhibit **dynamic orientational disorder**
- Polarization must be evaluated as a **time-averaged quantity**
- Static DFT → insufficient to capture finite-temperature behavior

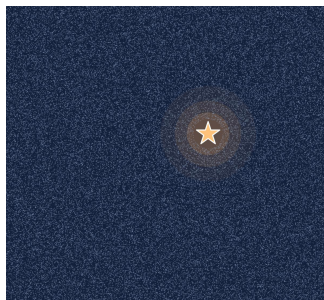


Active Learning: Compute Only Where It Matters

- Uncertainty-aware ML + DFT oracle loop
- DFT only where model confidence is lowest → **max information / cost**
- Examples:
 - MACE committee → MD configurations with high force uncertainty → selective DFT labeling
 - Bayesian optimization → Targeted DFT evaluation in promising + uncertain regions
 - Progressive ML → Iterative retraining with DFT-validated candidates

→ Extend to piezoelectric screening → fully autonomous workflow

Take-Home Message



We asked: *Can AI find a crystal that turns a push into a voltage?*

Yes, when chemical insight guides AI

Topic 1 · Screen

2.7 M $\rightarrow$ 2 families

DFT-trained ML ranked millions of HOIPs and found high-performance and flexible piezoelectric candidates.

regression

Topic 2 · Generate

$\sim 890 \rightarrow \sim 20,000$

New AI-designed organic cations grow the searchable ABX_3 space more than 20-fold.

generative models · RL

Topic 3 · Respect symmetry

Full 3×6 tensor

Equivariant networks predict the whole piezoelectric tensor; the scalars follow.

group equivariance

Next: from frozen snapshots to real temperatures: MLIP molecular dynamics + active learning.
Chemistry decides what to look for and AI make the search possible. Let's solve the rest together.

Acknowledgments

Soongsil University

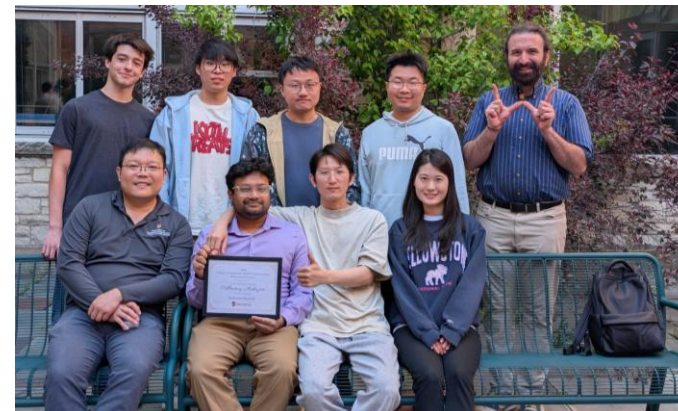


University of Wisconsin-Madison

Prof. Ying Li

Jianxin He

Tianle Yue



Thank you for
listening!

