Data-based methods to accelerate discovery of novel materials and find new properties in old ones

François-Xavier Coudert

@fxcoudert

www.coudert.name

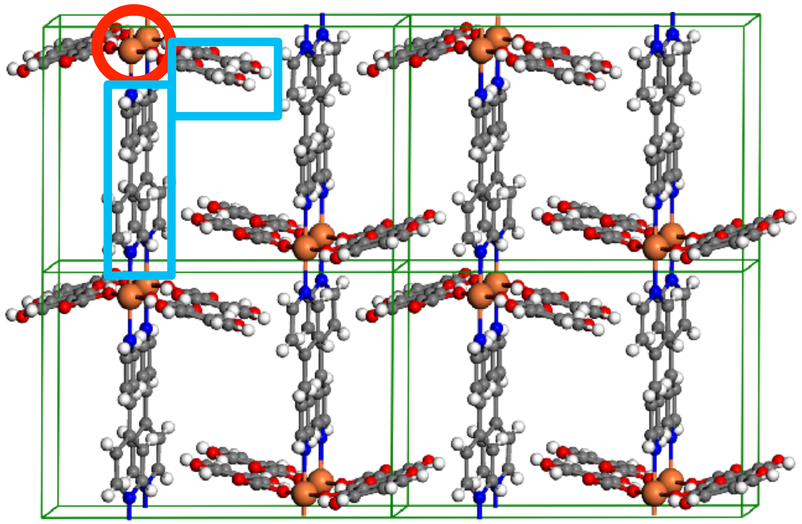
Chimie ParisTech

PSL UNIVERSITÉ PARIS

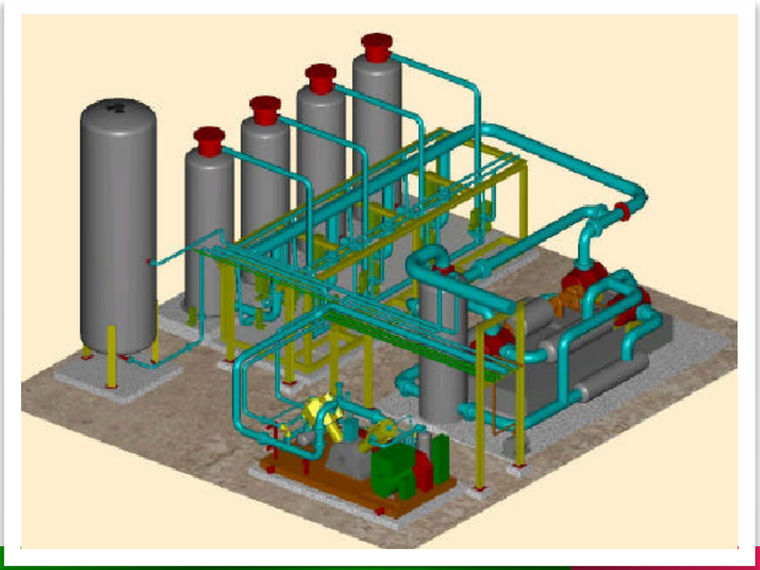
cnrs

Metal–Organic Frameworks

Cristalline, organic–inorganic hybrid nanoporous materials



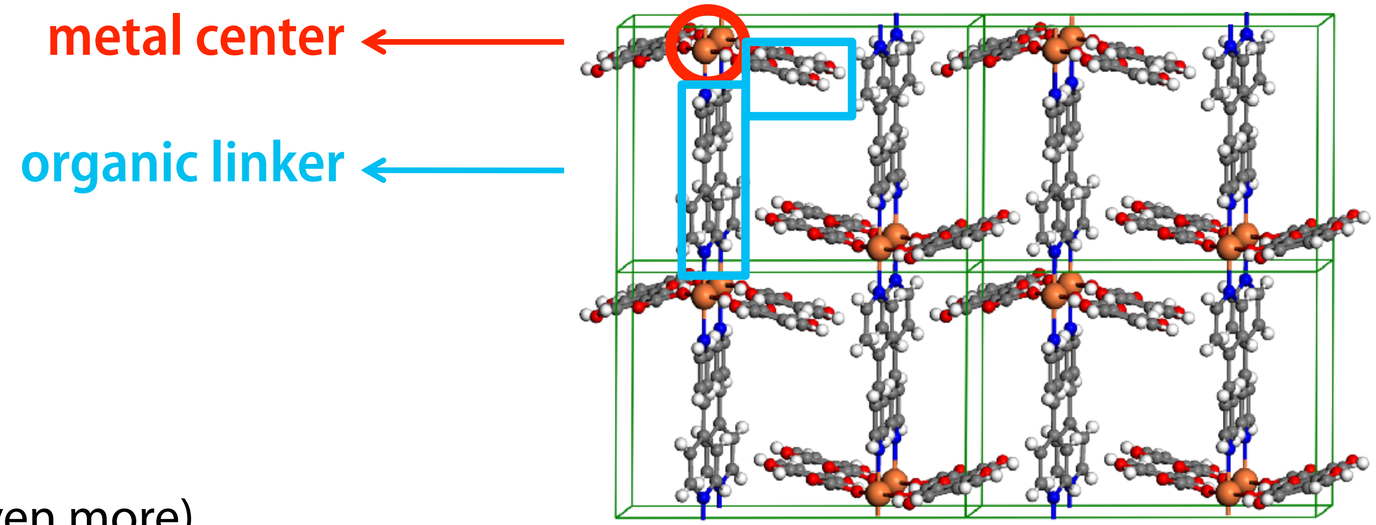
- Flexibility of coordination chemistry: pore geometry and topology
- Versatility of organic chemistry: pore size and internal surface



- Applications: gas adsorption, catalysis, sensing, delivery, …
- **High structural flexibility of their frameworks**
- **Important limitation for applications: hydrothermal & mechanical stability**

Metal–Organic Frameworks

Cristalline, organic–inorganic hybrid nanoporous materials

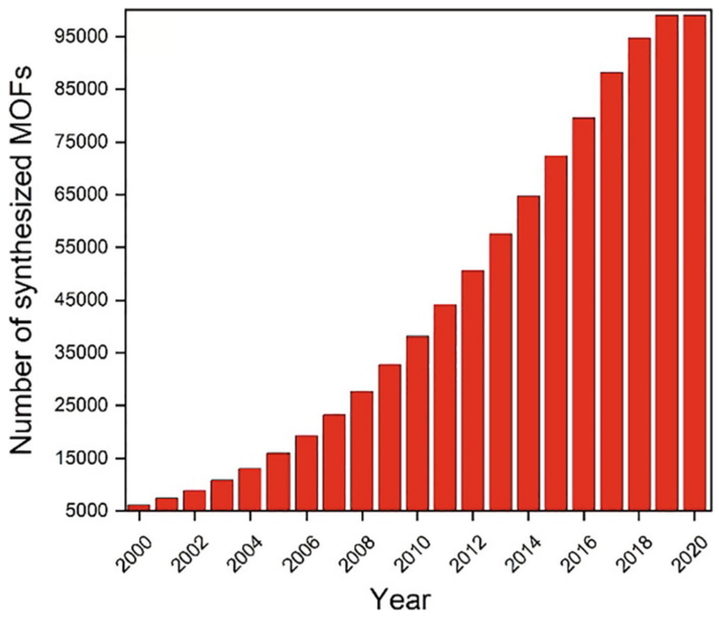


- Metals (many) and inorganic blocks (even more)
- Organic linkers… and functionalization
- Multivariate MOFs are possible
- Topology
- Guest molecule… or guests

- Thermodynamic space: temperature, pressure, composition

Why data-based methods?

Daglar & Keskin, 2020



☆ **Since this meeting started, 61 MOF papers were tweeted**

Figure 4. Most frequent emojis in followers' profiles. National flags were excluded from the analysis.

Coudert, *Chem. Mater.* 2023

ML methods for chemical sciences

- **Property prediction**: from structure or from composition
  (supervised learning, data obtained experimentally or computationally)
- **High-throughput screening**: applying predictor at large scale
- Analysis and exploration of diversity, **clustering of molecules**
- **Generative ML methods**: creating new molecules, new materials
- **Text and data mining**: a lot of information in published literature, in notebooks
- **AI for synthesis prediction**: propose a synthesis method/protocal,
  possibly drive robotic chemistry lab
- ML to **improve computational chemistry**:
  using machine learning to design new force fields, new DFT functionals, etc.

- … and many more…

Can we predict mechanical properties of crystalline materials?

PSL
UNIVERSITÉ PARIS

Chimie ParisTech

cnrs

Meta-MOFs

Coudert & Evans, *Coord. Chem. Rev.* **2019**

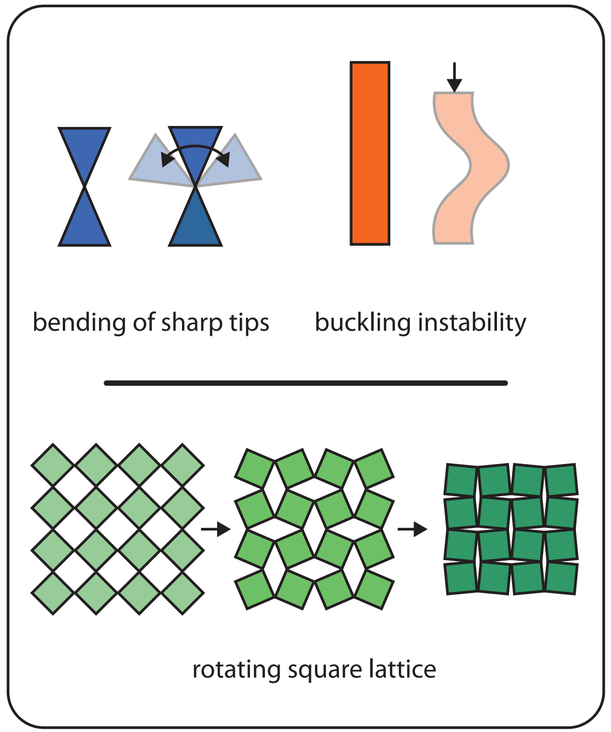
metamaterials

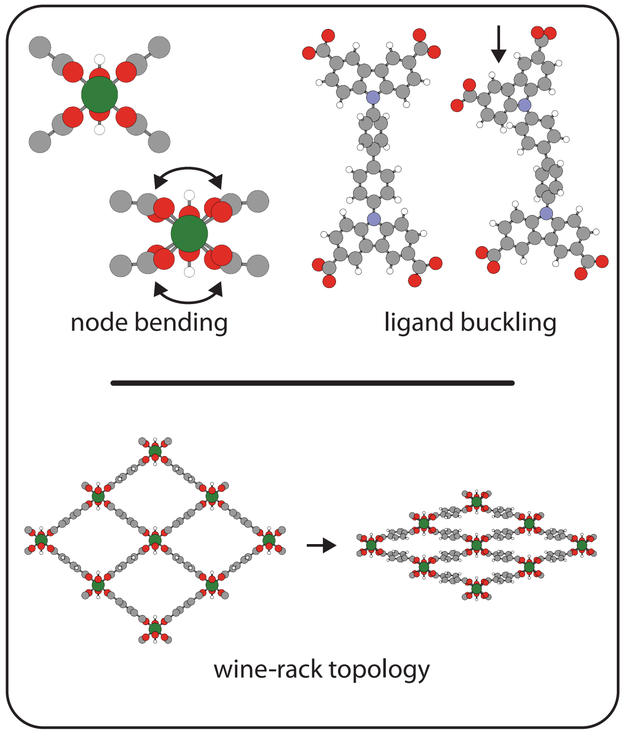
composite material with a structure that exhibits properties <u>not usually found in natural materials</u>

μετά* = *beyond

meta-MOFs

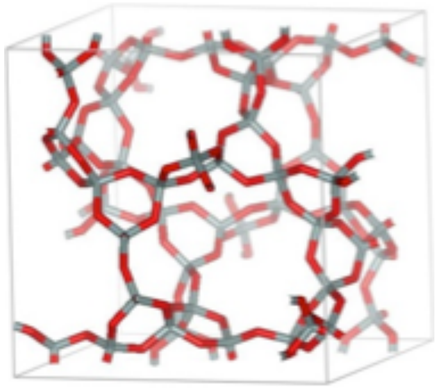
negative thermal expansion
negative compressibility
negative adsorption
breathing
chiral induction
…

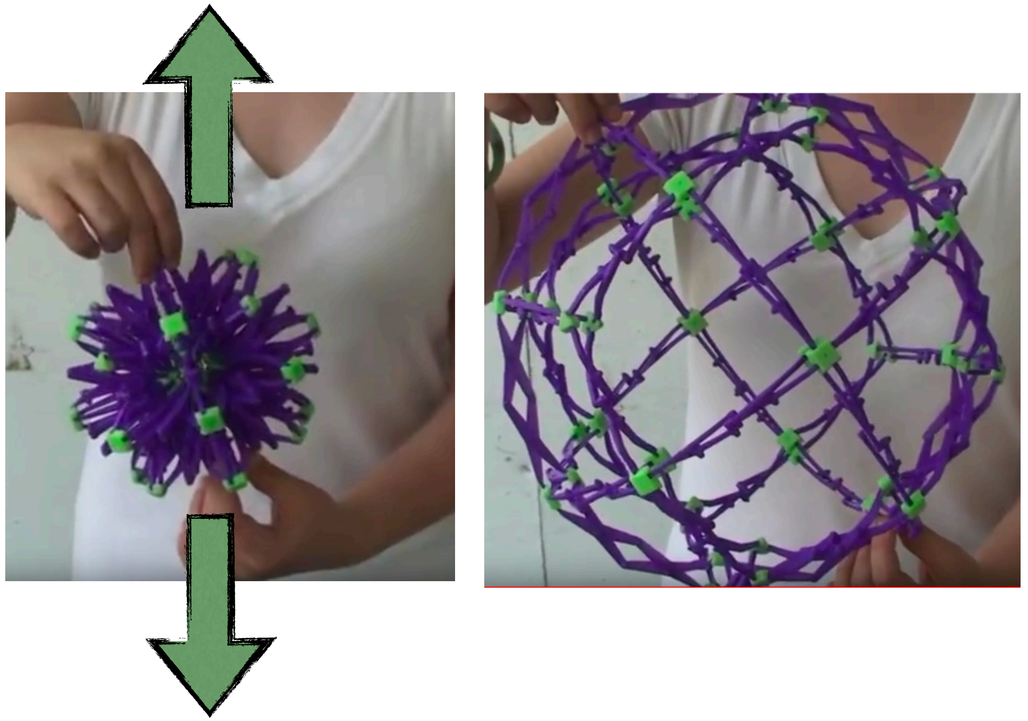




Mechanical properties of crystals

- In 2016, we identified by chance a zeolite with **isotropic auxeticity**





- Only **5 known crystals** with this property!
- Also considered "rare": **negative linear compressibility**

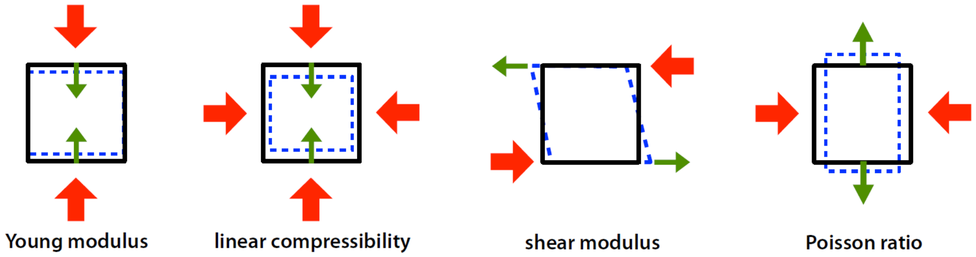
- **How rare are other so-called "rare" mechanical properties?**

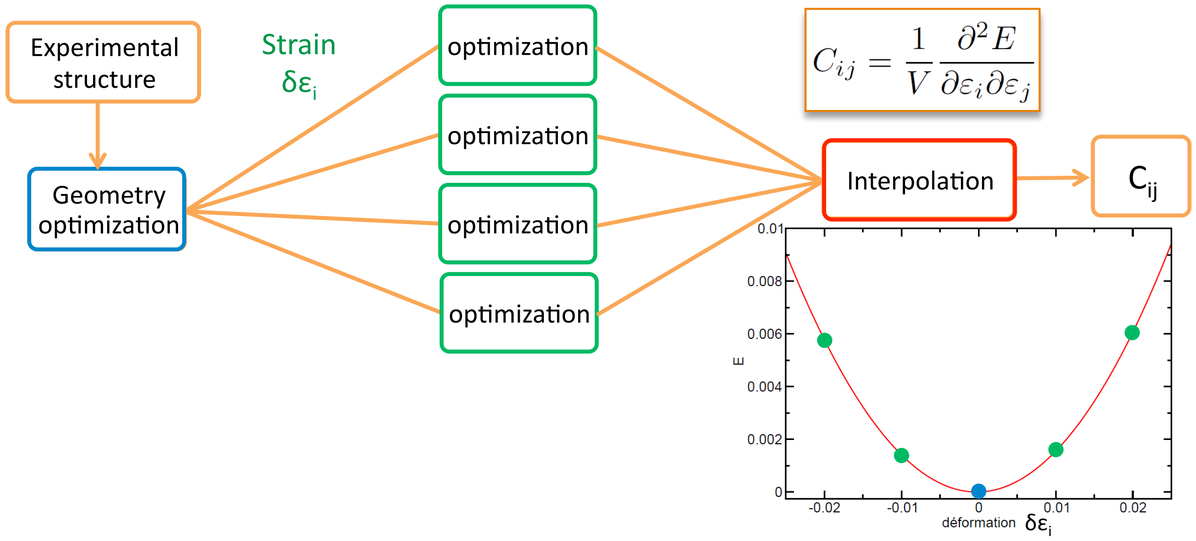
Mechanical properties of crystals

$$\sigma_{ij} = C_{ijkl}\,\epsilon_{kl}$$

$$C_{ijkl} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & & & \\ \cdot & C_{22} & C_{23} & & & \\ \cdot & \cdot & C_{33} & & & \\ & & & C_{44} & & \\ & & & & C_{55} & \\ & & & & & C_{66} \end{pmatrix}$$

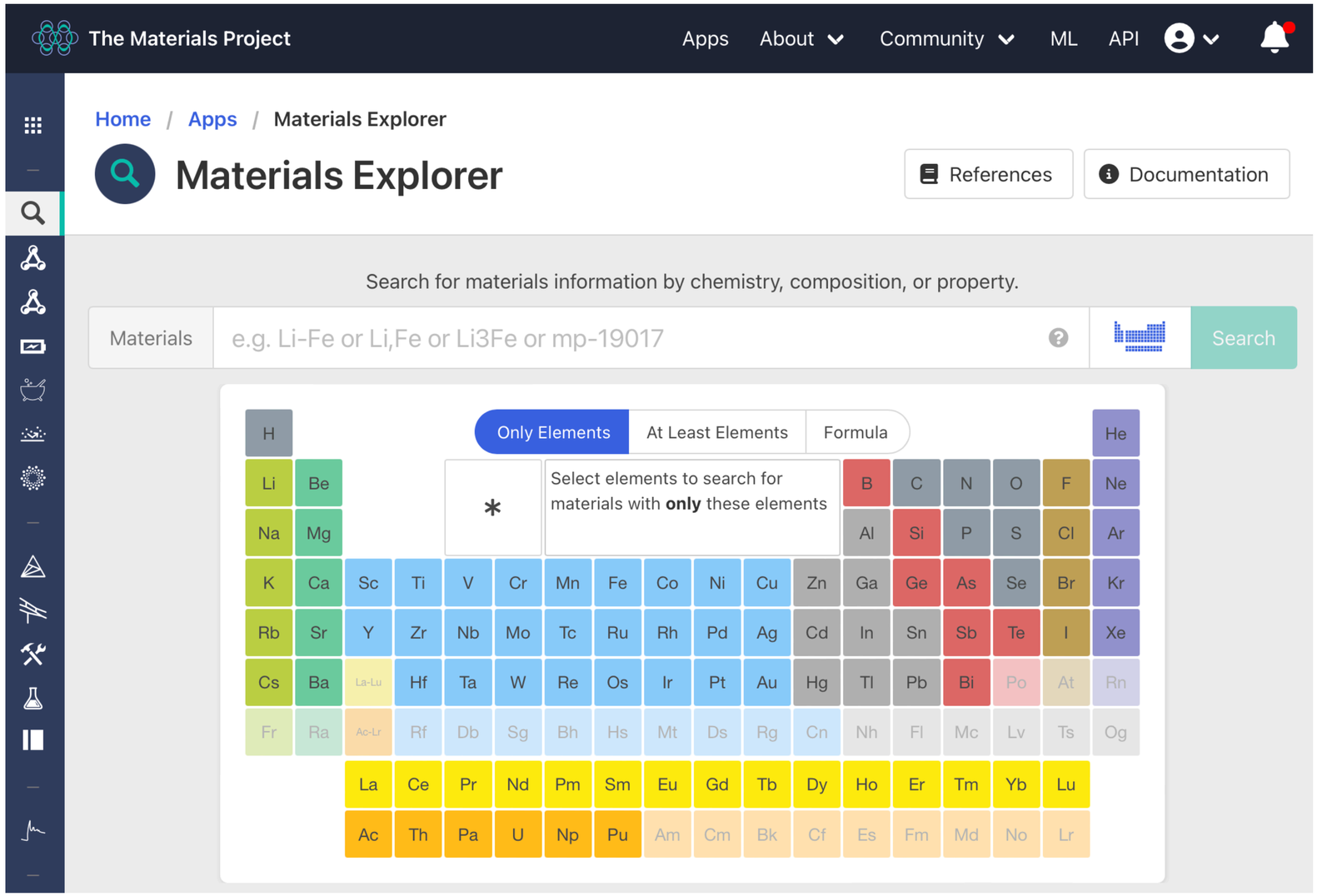
Elastic tensor

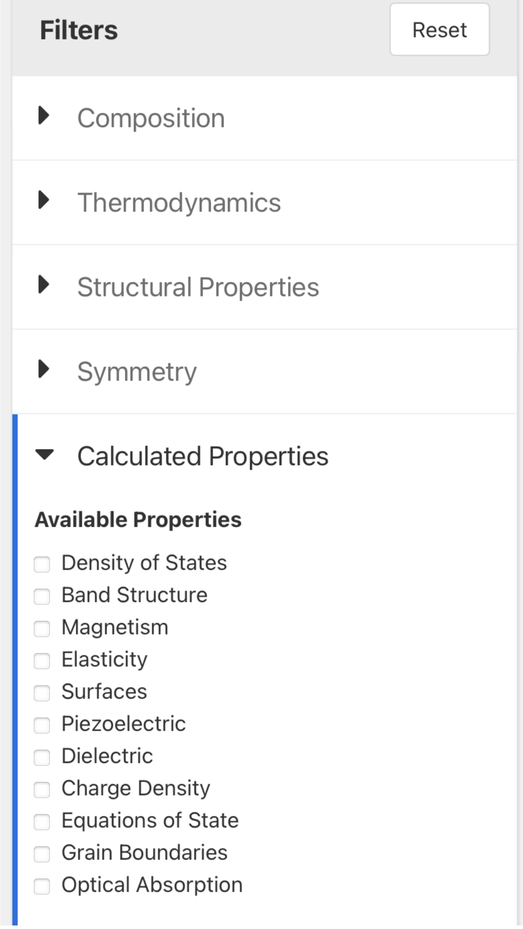




- Elasticity is an anisotropic property
- Experimentally difficult to determine
- "Relatively easy" to compute from DFT

- Most people only care about the bulk modulus, but there is a lot more information!

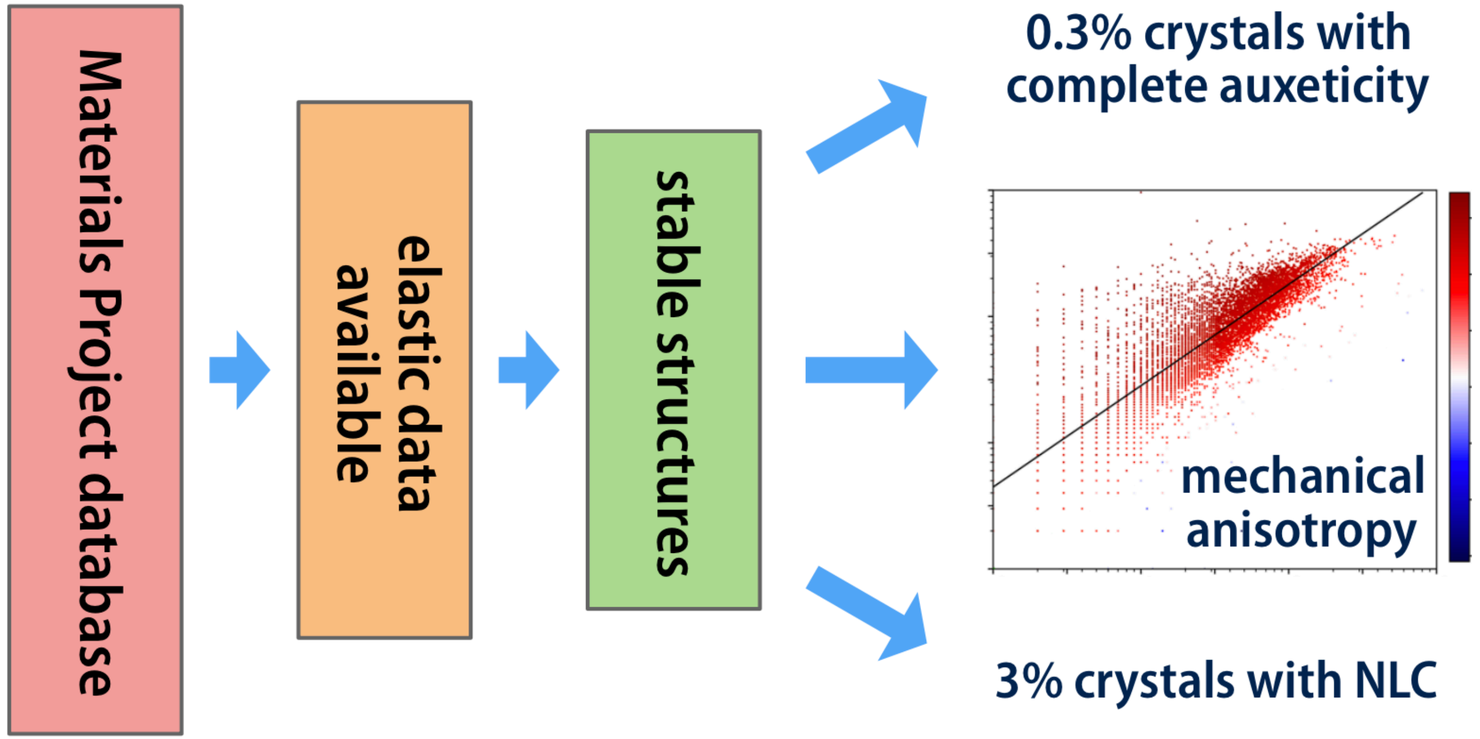

The Materials Project
Apps
About
Community
ML
API
Home / Apps / Materials Explorer
Materials Explorer
References
Documentation
Search for materials information by chemistry, composition, or property.
Materials
e.g. Li-Fe or Li,Fe or Li3Fe or mp-19017
Search
Only Elements
At Least Elements
Formula
Select elements to search for materials with only these elements
*
H
He
Li Be
B C N O F Ne
Na Mg
Al Si P S Cl Ar
K Ca Sc Ti V Cr Mn Fe Co Ni Cu Zn Ga Ge As Se Br Kr
Rb Sr Y Zr Nb Mo Tc Ru Rh Pd Ag Cd In Sn Sb Te I Xe
Cs Ba La-Lu Hf Ta W Re Os Ir Pt Au Hg Tl Pb Bi Po At Rn
Fr Ra Ac-Lr Rf Db Sg Bh Hs Mt Ds Rg Cn Nh Fl Mc Lv Ts Og
La Ce Pr Nd Pm Sm Eu Gd Tb Dy Ho Er Tm Yb Lu
Ac Th Pa U Np Pu Am Cm Bk Cf Es Fm Md No Lr


Filters
Reset
Composition
Thermodynamics
Structural Properties
Symmetry
Calculated Properties
Available Properties
Density of States
Band Structure
Magnetism
Elasticity
Surfaces
Piezoelectric
Dielectric
Charge Density
Equations of State
Grain Boundaries
Optical Absorption

Quantifying anomalous behavior

- Materials Project: 133,691 inorganic compounds
- Elastic data at DFT level for 13,621 structures

- Systematic tensorial analysis to answer this simple question:
  mechanical metamaterials are rare, but how rare exactly?



Chibani & Coudert, *Chem. Sci.* **2019**

Quantifying anomalous behavior

Table 1 List of completely auxetic materials in the Materials Project database, with extremal values of directional Poisson's ratio, and isotropic average

Material ID	Structure	Synthesized	ν_{min}	ν_{max}	$\langle\nu\rangle$
mp-1021516	K_2Sn	No	−0.26	−0.20	−0.21
mp-9580	$TlGaSe_2$	Yes	−0.94	−0.24	−0.59
mp-982773	Na_3Tl	No	−0.50	−0.20	−0.4
mp-862769	$RbGe_3$	No	−1.25	−0.17	−0.18
mp-974789	Rb_3Sn	No	−0.75	−0.73	−0.62
mp-7621	$KTcO_4$	Yes	−0.41	−0.04	−0.2
mp-36508	$SnHgF_6$	No	−1.08	−0.10	−0.45
mp-15639	$HgRhF_6$	Yes	−0.53	−0.14	−0.4
mp-999274	$RbNaH_2$	Yes	−0.77	−0.47	−0.67
mp-697133	Cs_2CaH_4	Yes	−0.56	−0.32	−0.47
mp-27718	$CsHgBr_3$	Yes	−0.15	−0.06	−0.12
mp-865080	$NaCeAu_2$	No	−0.35	−0.29	−0.3
mp-13925	Cs_2NaYF_6	Yes	−0.85	−0.77	−0.82
mp-7961	Sr_3SnO	Yes	−0.08	−0.08	−0.09
mp-989580	Cs_2KNF_6	No	−0.18	−0.07	−0.14
mp-989523	Rb_2NaAsF_6	No	−0.31	−0.20	−0.26
mp-4051	$AlPO_4$	Yes	−0.58	−0.05	−0.28
mp-631316	Li_2GaSb	No	−0.05	−0.05	−0.05
mp-866229	Ca_2SnHg	No	−0.74	−0.65	−0.7
mp-2739	TeO_2	Yes	−0.77	−0.37	−0.54
mp-989536	Cs_2LiNF_6	No	−0.78	−0.75	−0.75
mp-867920	$K_2Rh_2O_5$	No	−0.57	−0.00	−0.27
mp-21200	$PuGa_2$	Yes	−0.45	−0.07	−0.28
mp-989590	Ca_6Sn_2NF	No	−0.58	−0.53	−0.55
mp-20457	InP	Yes	−0.86	−0.77	−0.81
mp-1025524	Zr_2TlC	Yes	−0.20	−0.02	−0.07
mp-1017566	$GePbO_3$	Yes	−0.50	−0.26	−0.38
mp-1008282	Cr_3Fe	Yes	−0.25	−0.04	−0.13

- No clear systematic…
- What do all these materials have in common?
- Can such complex relationships be captured by chemical descriptors? topological descriptors?
- A good case study for deep learning?

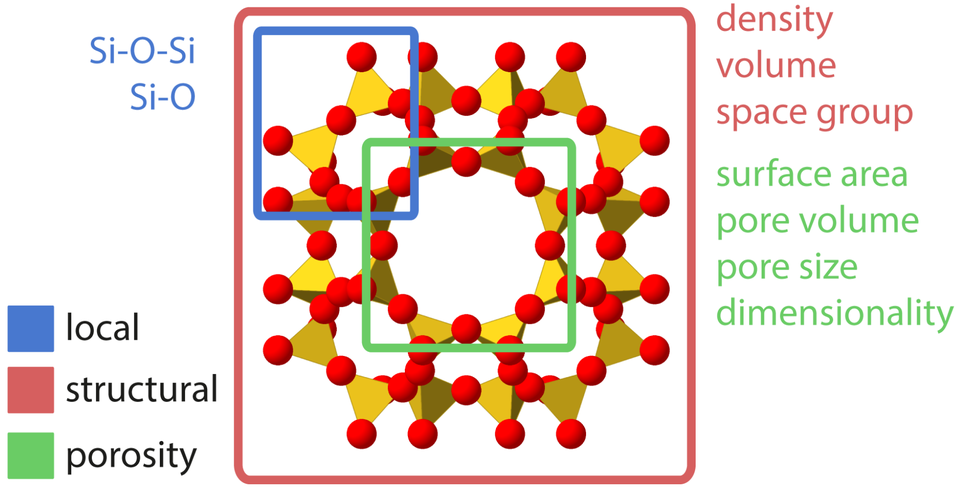
Chibani & Coudert, *Chem. Sci.* **2019**

Predicting mechanical properties

We have a smaller data set (SiO_2 zeolites) that is chemically homogeneous

Different kinds of descriptors are available, with different information:

- Hand-picked geometrical descriptors, relying on our know
- Unbiased/agnostic local geometrical descriptors (e.g. Smooth Overlap of Atomic Positions + PCA)
- Porous network characteristics (Zeo++)
- Topological information?



- Geometrical descriptors are best
- SOAP + PCA performs generally as well as "smart" descriptors

Evans & Coudert, *Chem. Mater.* **2017**

Hunting for anisotropic zeolites

- Anisotropic mechanical properties are much harder to predict
- Force fields generally perform badly
- What we are looking for is a very rare property

- Let's try a multi-step approach

Gaillac, Chibani & Coudert, *Chem. Mater.* **2020**

Hunting for anisotropic zeolites

590,811 hypothetical structures

from Pophale *et al.*

BKS
force field

462,248 mechanically stable

predicted auxetic: 578

random subset: 742

DFT

DFT

392 stable structures

599 stable structures

machine
learning

trained predictor

Gaillac, Chibani & Coudert, *Chem. Mater.* **2020**

Hunting for anisotropic zeolites

- Force field predicts structures adequately, average mechanical properties "okay", but anisotropic properties are terrible
- GBR model based on geometric descriptors only, trained on DFT data, achieves much better accuracy

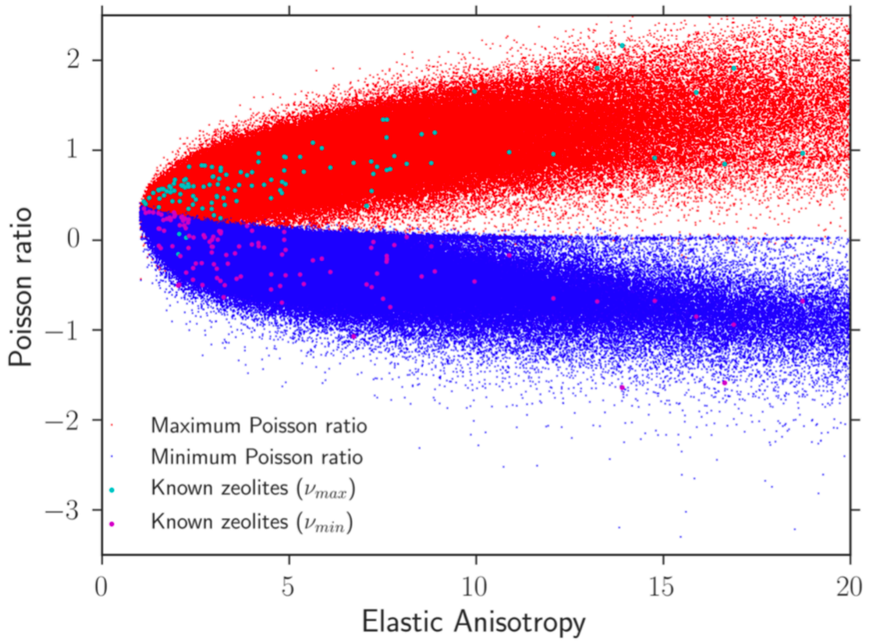


Table 4. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) for the Three Subsets and Their Assembly for the Prediction of the Poisson's Ratio

subset (method)	RMSE (ν_{min})	MAE (ν_{min})	RMSE (ν_{max})	MAE (ν_{max})
all (GBR)	0.39	0.26	0.46	0.32
all (BKS)	1.4	0.51	9.8	2.1

- Future work: extend to zeolitic frameworks with different chemical composition ($AlPO_4$, gallogermanates, etc.) and extra-framework cations

Predicting the full tensor?

Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

PAPER

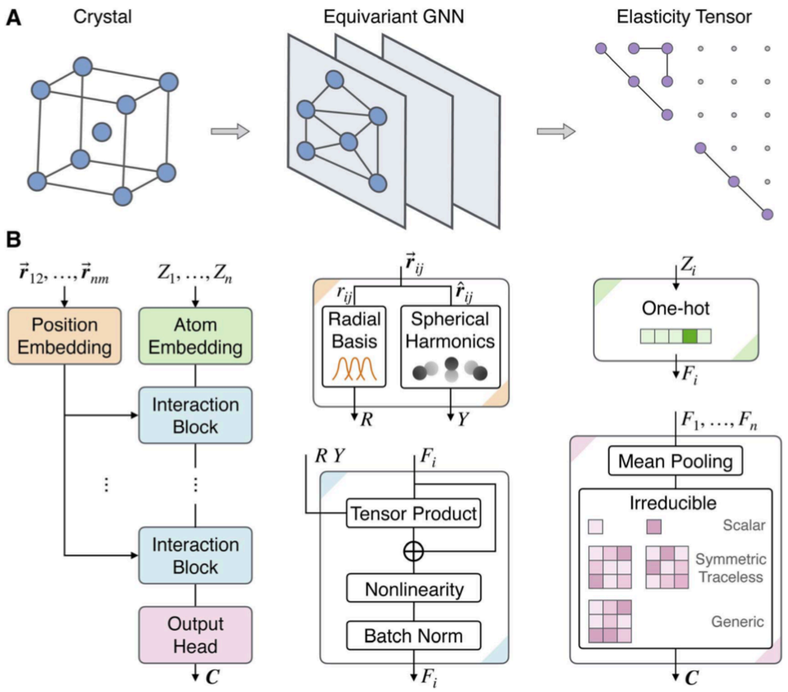
View Article Online
View Journal | View Issue

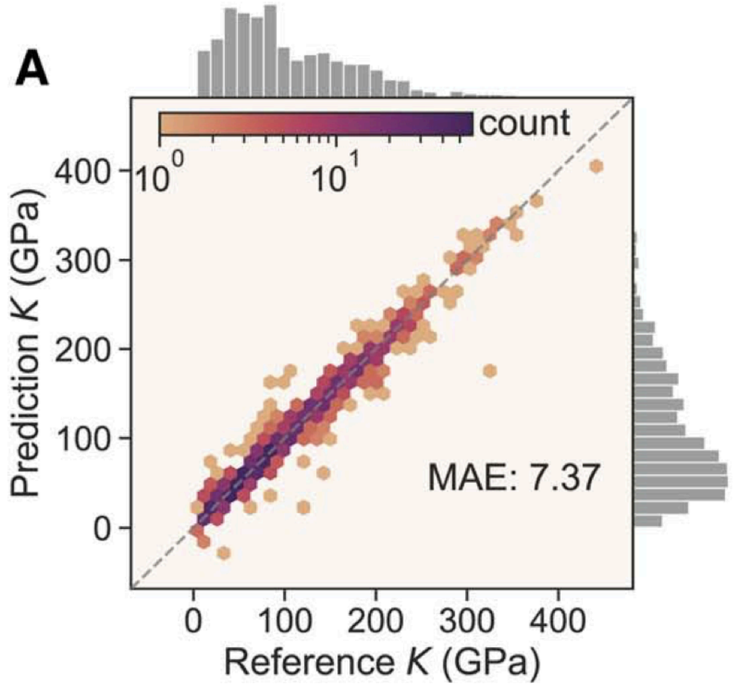
Check for updates

Cite this: *Digital Discovery*, 2024, 3, 869

An equivariant graph neural network for the elasticity tensors of all seven crystal systems†

Mingjian Wen, *[a] Matthew K. Horton, [bc] Jason M. Munro,[b] Patrick Huck [d] and Kristin A. Persson [ef]





Predicting the full tensor?

Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

PAPER

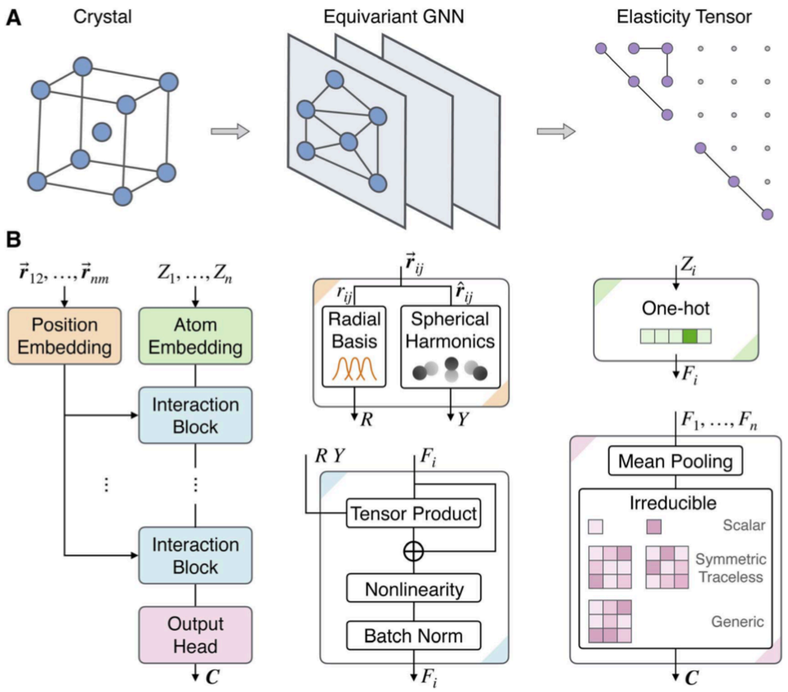
View Article Online
View Journal | View Issue

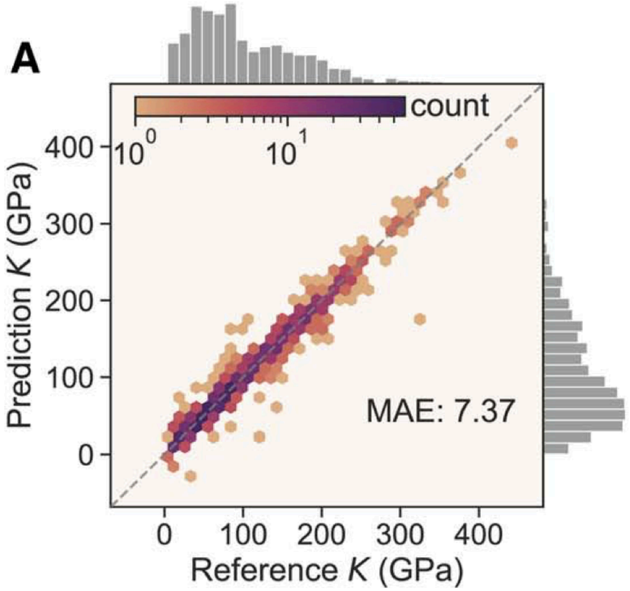
Check for updates

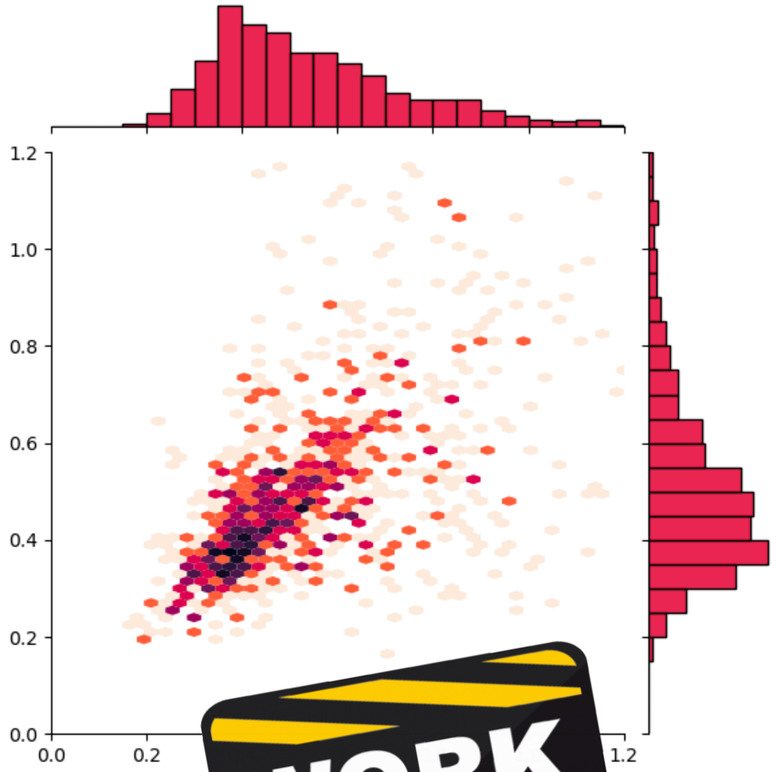
Cite this: *Digital Discovery*, 2024, 3, 869

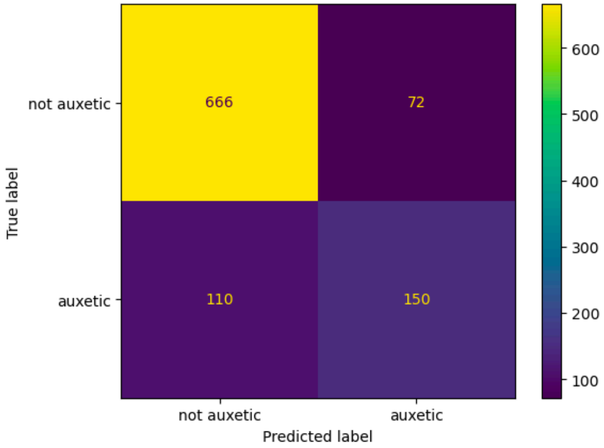
An equivariant graph neural network for the elasticity tensors of all seven crystal systems†

Mingjian Wen, *[a] Matthew K. Horton, [bc] Jason M. Munro,[b] Patrick Huck [d] and Kristin A. Persson [ef]









Can we model amorphous MOFs?

PSL UNIVERSITÉ PARIS

Chimie ParisTech

cnrs

Modelling amorphous MOFs

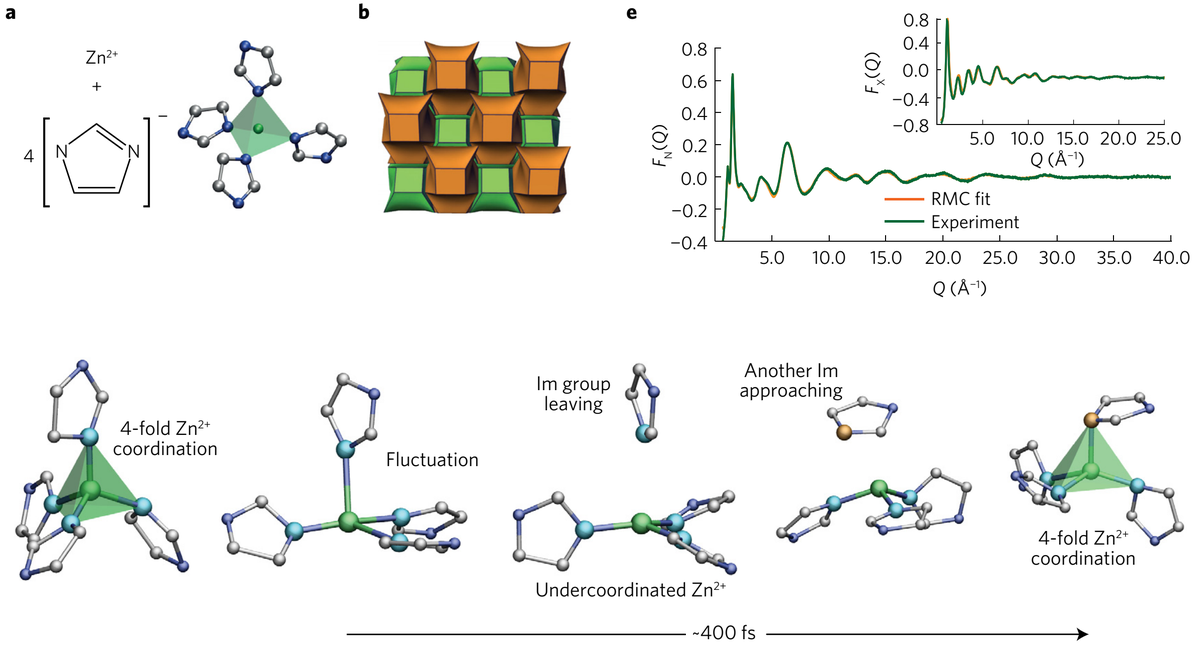
nature materials

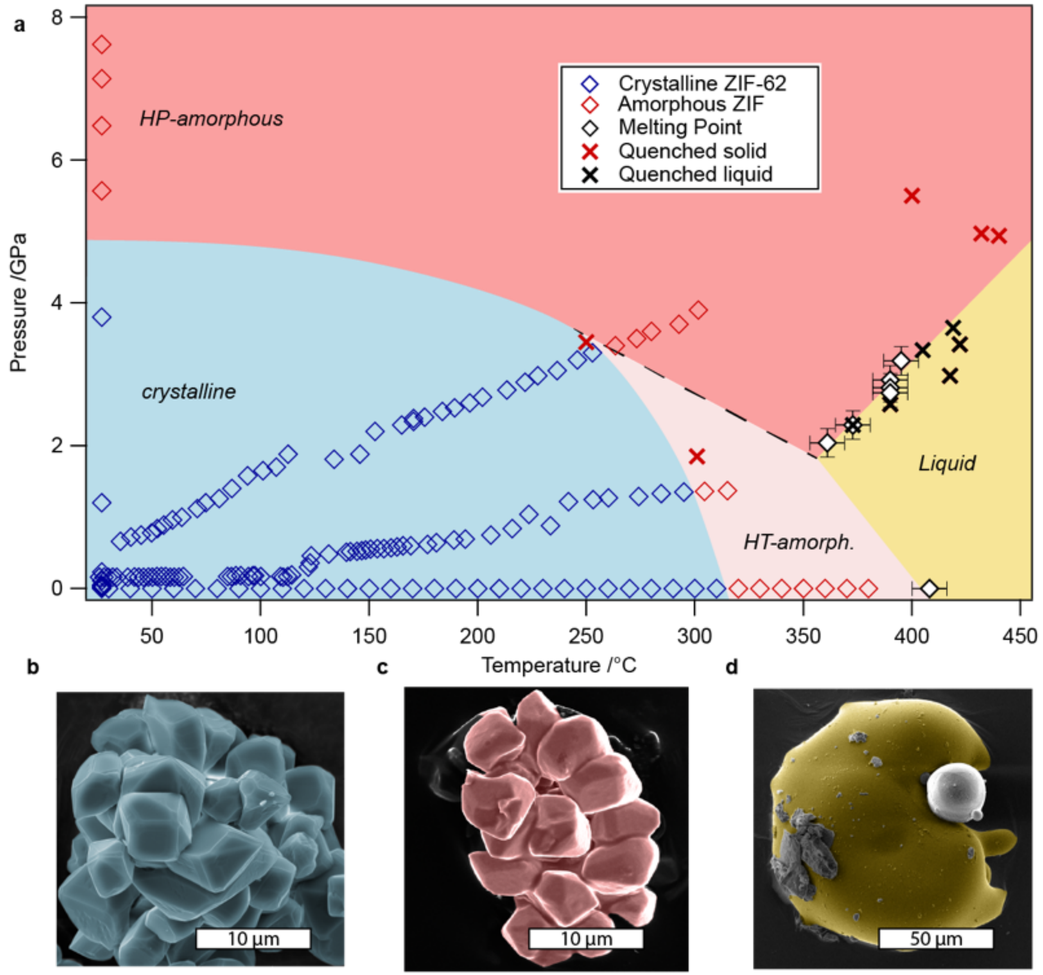
ARTICLES

PUBLISHED ONLINE: 9 OCTOBER 2017 | DOI: 10.1038/NMAT4998

Liquid metal–organic frameworks

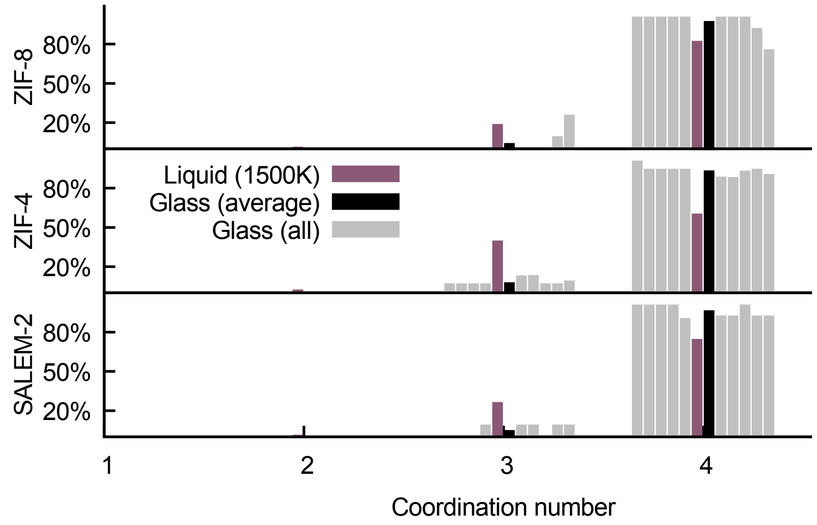
Romain Gaillac[1,2], Pluton Pullumbi[2], Kevin A. Beyer[3], Karena W. Chapman[3], David A. Keen[4], Thomas D. Bennett[5]* and François-Xavier Coudert[1]*

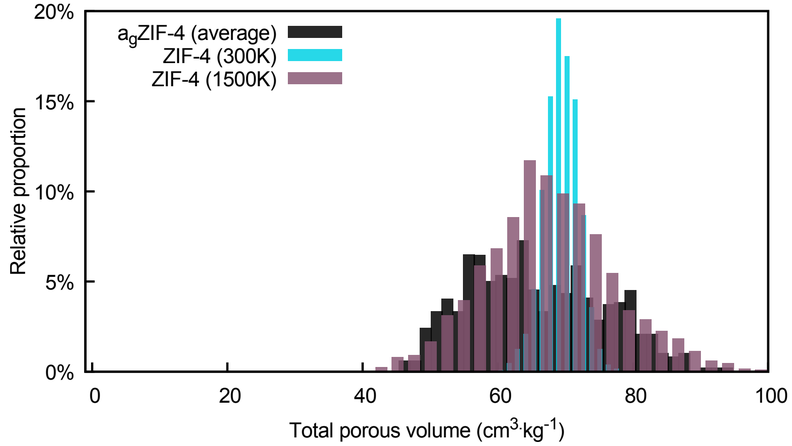


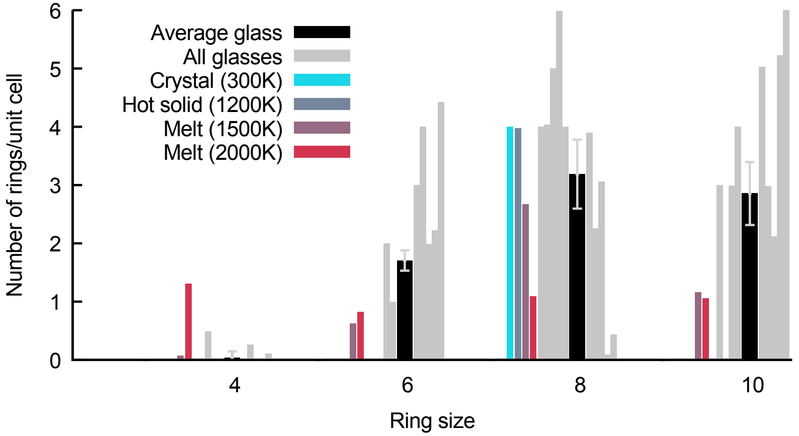


Modelling amorphous MOFs

- Bond breaking and formation: ab initio molecular dynamics
- Disorder, no periodicity: large simulation box sizes
- Slow dynamics: long simulation times



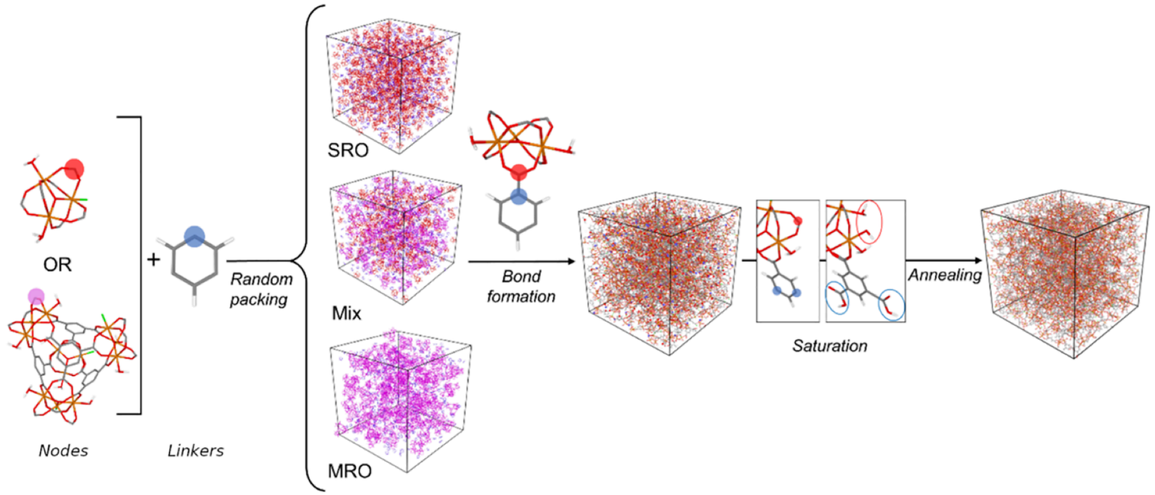




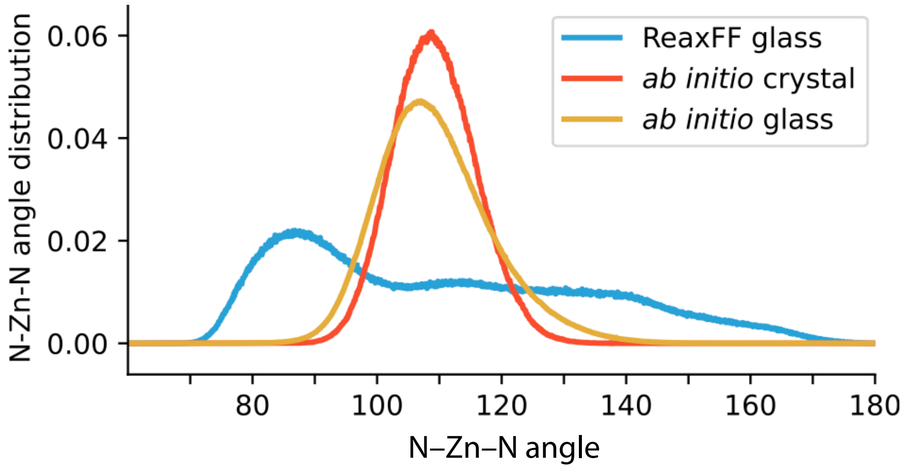
- **Ab initio MD works**, sort of… but is painfully slow

Comparing to other methods

- **Reverse Monte Carlo models, Continuous Random Network**: underconstrained, you get what you put in
- **Polymatic** (Jelfs group): great approach, but needs microscopic description



- **ReaxFF**: *friends don't let friends use ReaxFF*



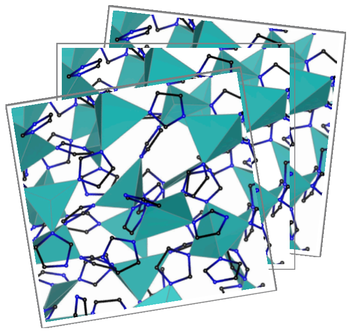
Castel et al, *J Phys Chem C* 2022 (review)

Castel et al, *J Phys Chem C* 2022 (ReaxFF)

New compchem for amorphous MOFs

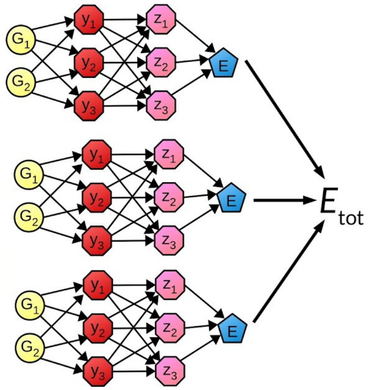
- Ab initio MD describes best the local environment
- Very high computational cost:
  small unit cell (~1000 atoms), short simulations (~200 ps) => millions CPU hours
- Influence of initial configuration (our glasses look too much like the crystal)

- But these *ab initio MD runs* represent a large amount of data



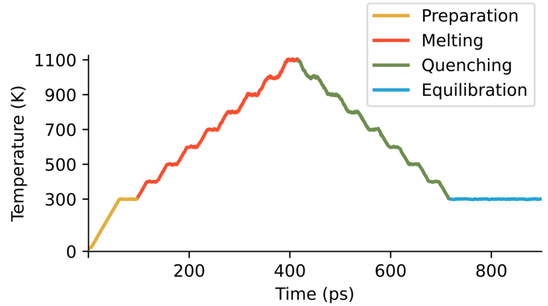
MD configurations
$\boldsymbol{r}$, E_{pot}, $\boldsymbol{F}$, $\boldsymbol{\sigma}$

train



E(3)-equivariant
neural networks

deploy



MD with machine
learned potential

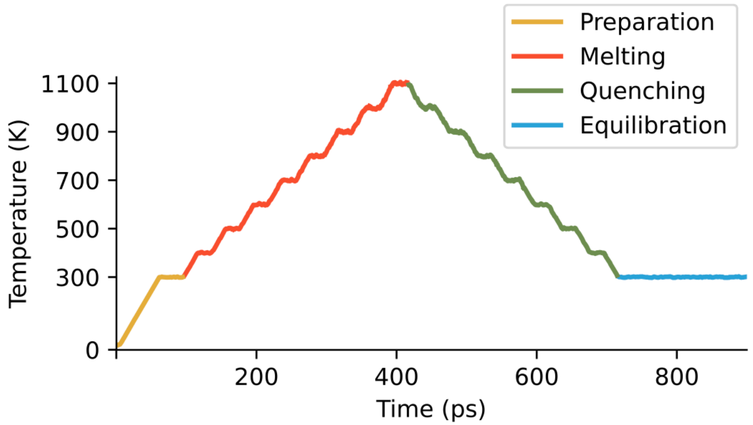
LAMMPS

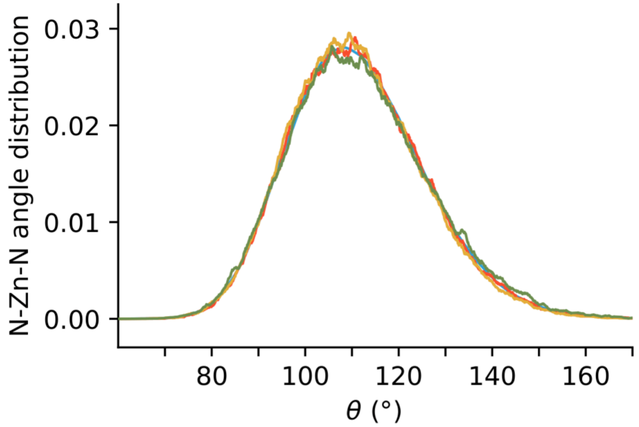
NequIP

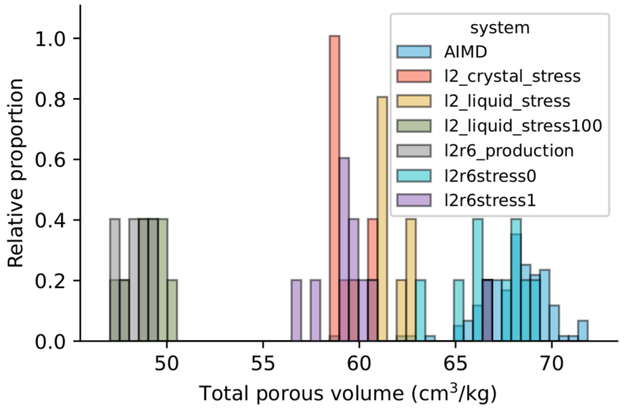
Castel et al, *Digital Discovery*, 2023

New compchem for amorphous MOFs

- Approach #1: Melt-quenching the crystals
  Is already working quite well…







- Approach #2: Using Polymatic + ML potential
  (collaboration with Imperial College London, Kim Jelfs group)

Castel et al, *Digital Discovery*, 2023

What next?

- ML potentials for ZIFs with multiple topologies, multiple linkers
- ML potentials for other amorphous MOFs
- Create a database of amorphous models and associated properties:
  - porosity and adsorption
  - topological analysis
  - dynamics of the framework
  - mechanical and thermal properties *(hard to measure experimentally)*
- **Amorphous phases have a lot of promise, but so little is known about them**

- Pressure/volume curves are very difficult to reproduce with the current methodology
- Learning stress appears to be much more difficult than energies and forces

On this topic...

- Reproducibility of models

Nature Chemistry, 2021

Check for updates

comment

Best practices in machine learning for chemistry

Statistical tools based on machine learning are becoming integrated into chemistry research workflows. We discuss the elements necessary to train reliable, repeatable and reproducible models, and recommend a set of guidelines for machine learning reports.

Nongnuch Artrith, Keith T. Butler, François-Xavier Coudert, Seungwu Han, Olexandr Isayev, Anubhav Jain and Aron Walsh

- Introductory review

APL Materials RESEARCH UPDATE scitation.org/journal/apm

Machine learning approaches for the prediction of materials properties

Cite as: APL Mater. 8, 080701 (2020); doi: 10.1063/5.0018384
Submitted: 15 June 2020 • Accepted: 16 July 2020 •
Published Online: 4 August 2020

View Online Export Citation CrossMark

Siwar Chibani and François-Xavier Coudert[a)]