

innovating nanoscience



AI and machine learning for magnetism (and beyond)

Stefano Sanvito

School of Physics and CRANN, Trinity College Dublin, IRELAND



COMPUTATIONAL
SPINTRONICS

SANVITO RESEARCH GROUP
TRINITY COLLEGE, DUBLIN



CRANN

K. Rossi et al. (a lot of “al.”),

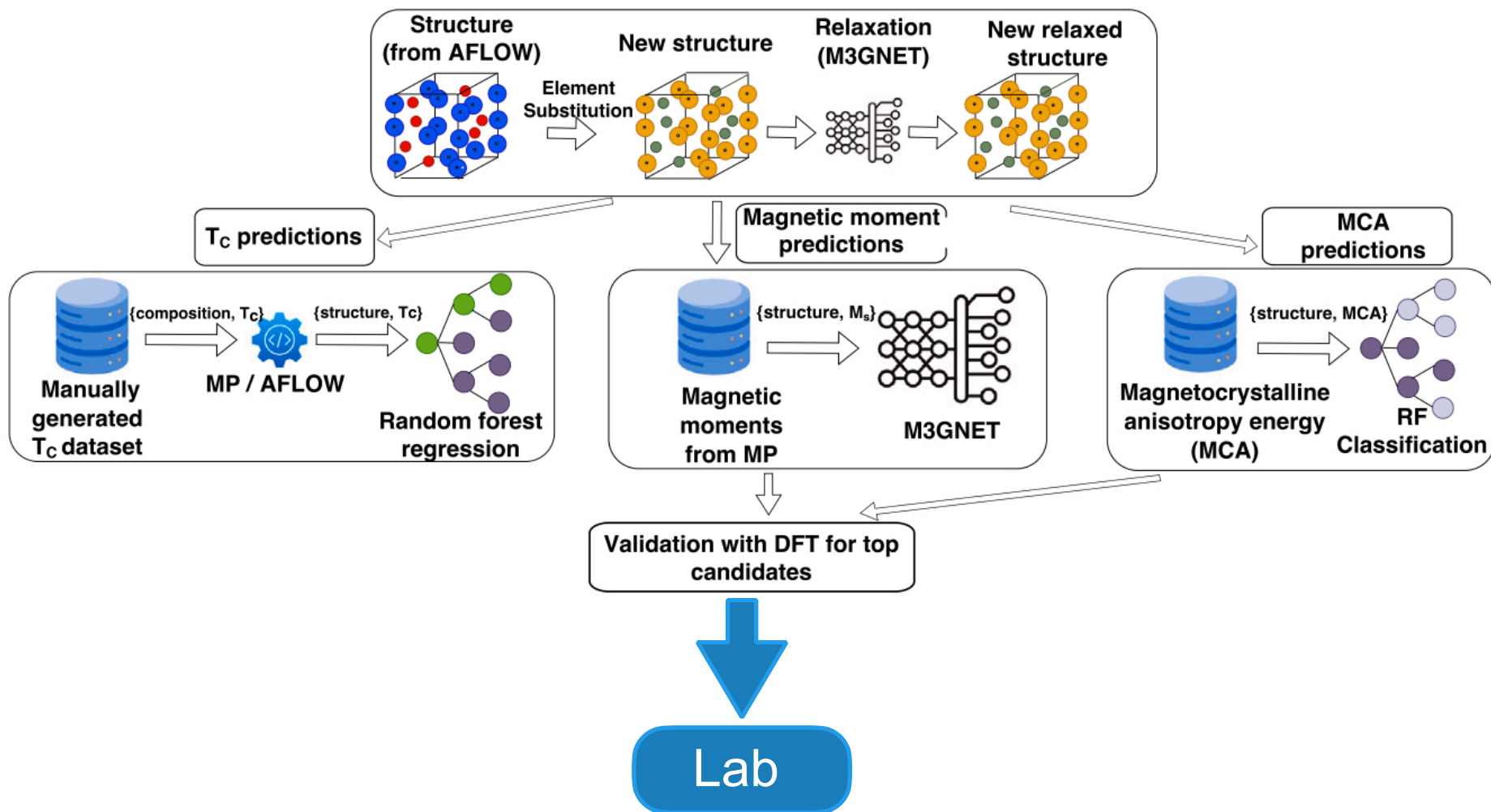
2026 Roadmap on Digitalising Materials Science,

JPhys Materials (2026);

10.1088/2515-7639/ae9010



How does a materials discovery workflow look like ?





COMPUTATIONAL
SPINTRONICS

SANVITO RESEARCH GROUP
TRINITY COLLEGE, DUBLIN



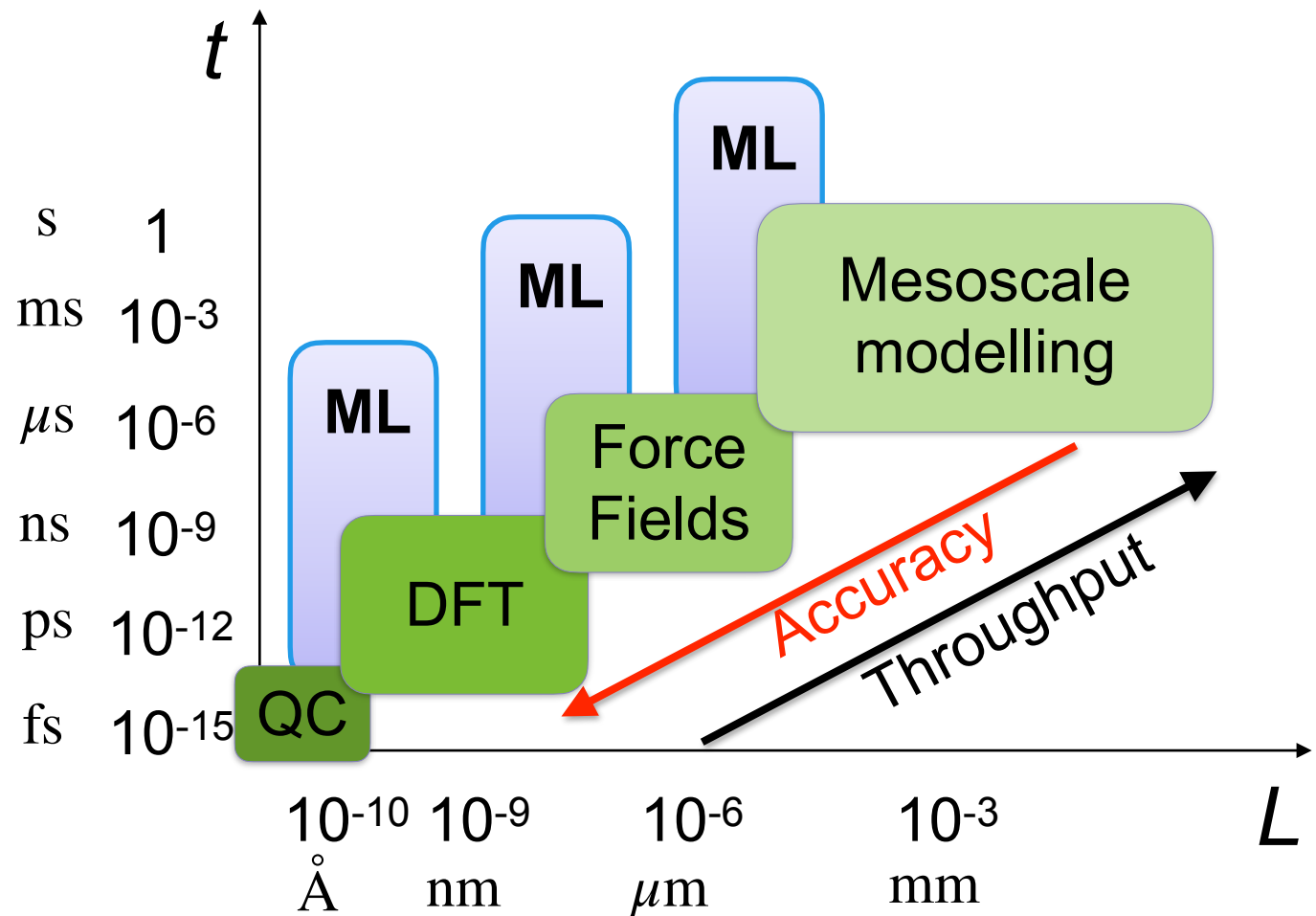
CRANN

Materials
theory across
the scales

Experimental data



ML Models





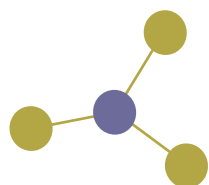
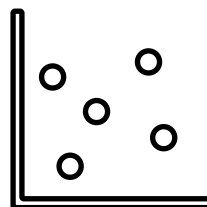
COMPUTATIONAL SPINTRONICS

SANVITO RESEARCH GROUP
TRINITY COLLEGE, DUBLIN

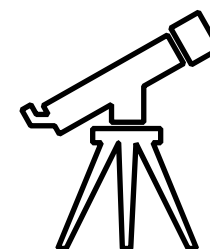
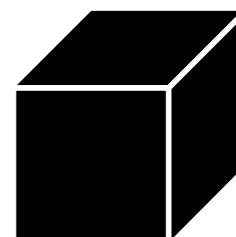


Data

Experimental
Synthetic
Theory



$\vec{v}$



Representation

ML model

Observable

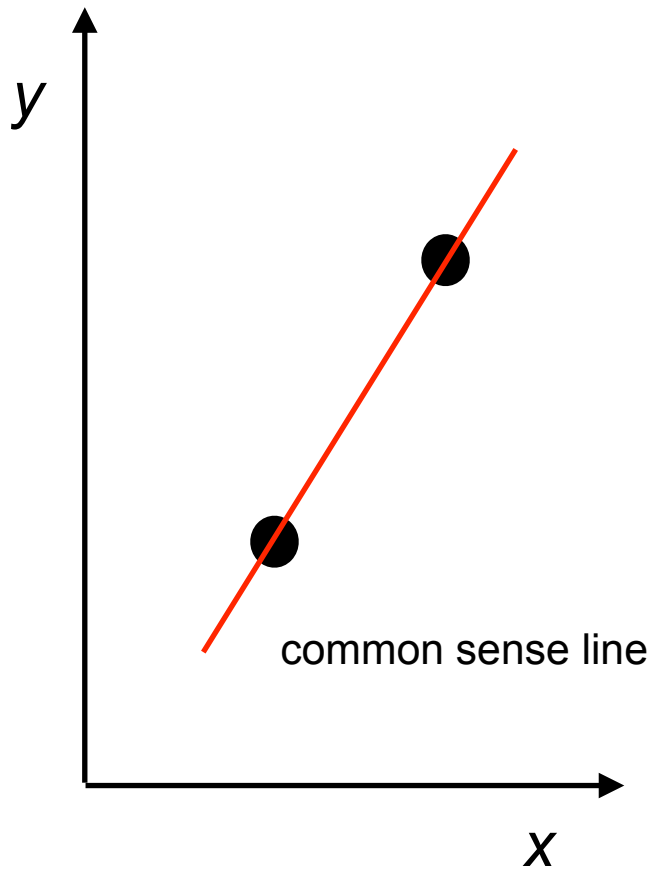
Chemistry
Structure
Time

Regression
N Network
Kernels
Transformers

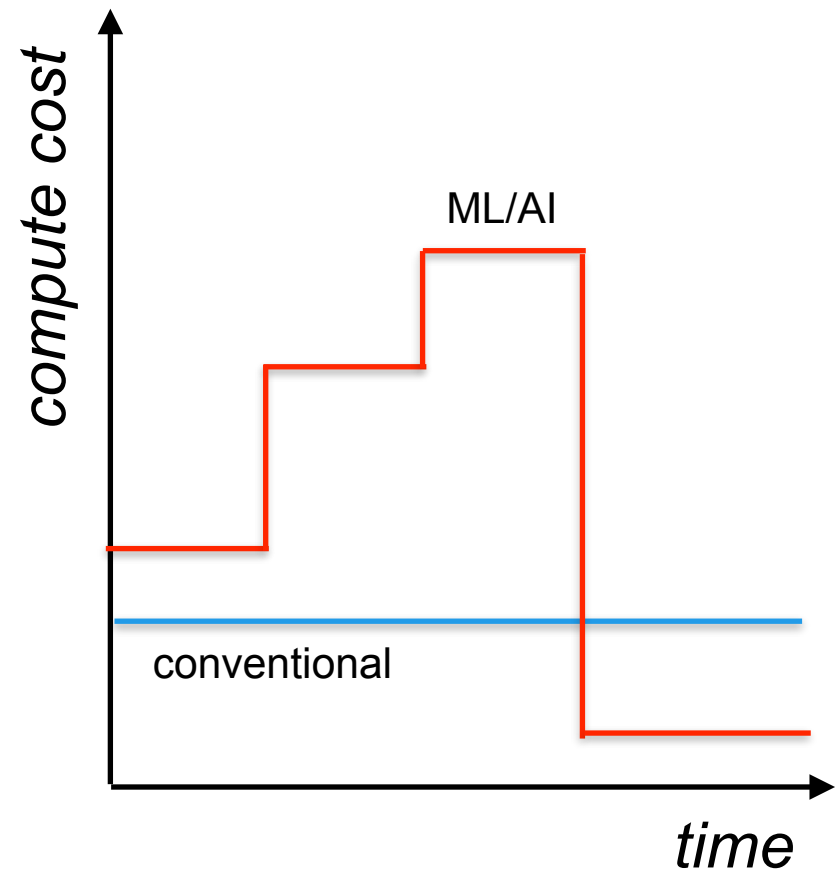
Scalar
Vectors
Tensors



Don't use it when it is simple



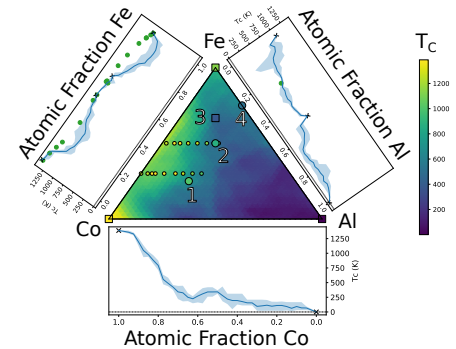
Don't use it when it is expensive



Outline

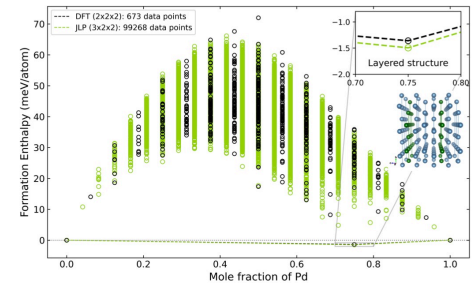
Curie temperature from experimental data

Moving to full AI agentic materials modelling



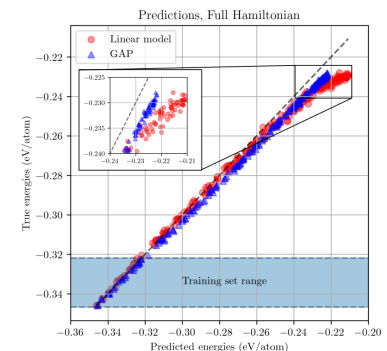
A cluster expansion for phase diagrams

ML to enable massive throughput



Spin-dependent force field

A model-agnostic tool for magnetic excitations



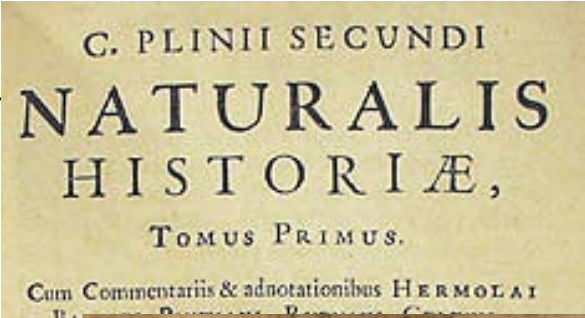
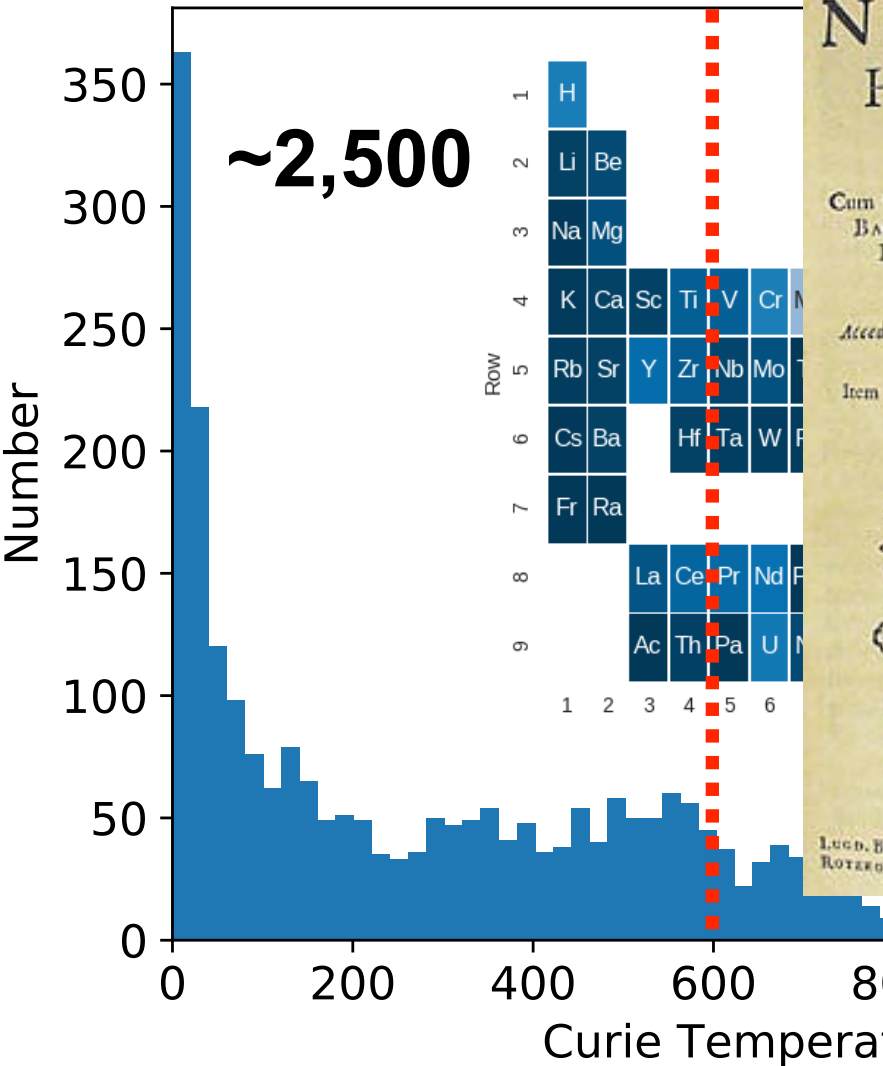


James Nelson, Matteo Cobelli, Luke Gilligan

ML with experimental data: T_c

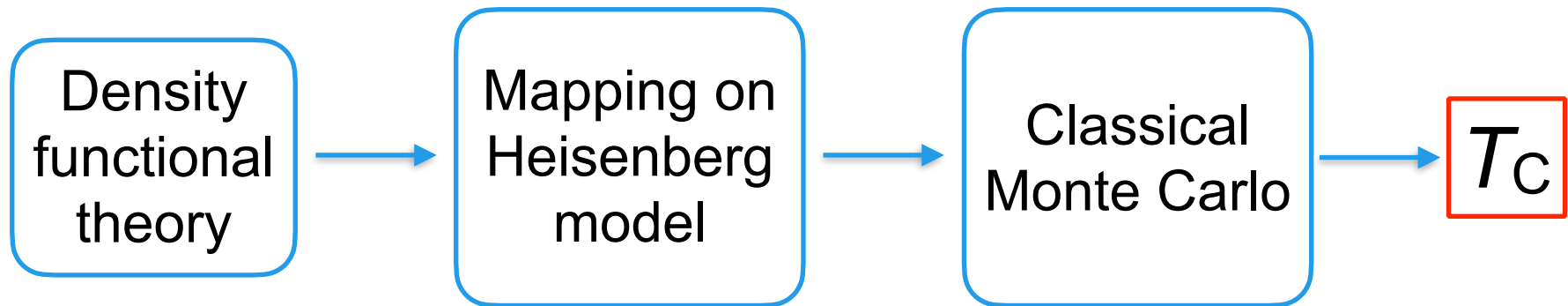
Goal: predict T_C of ferromagnets

Start from
experimental
data



Goal: predict T_C of ferromagnets

“Standard way”

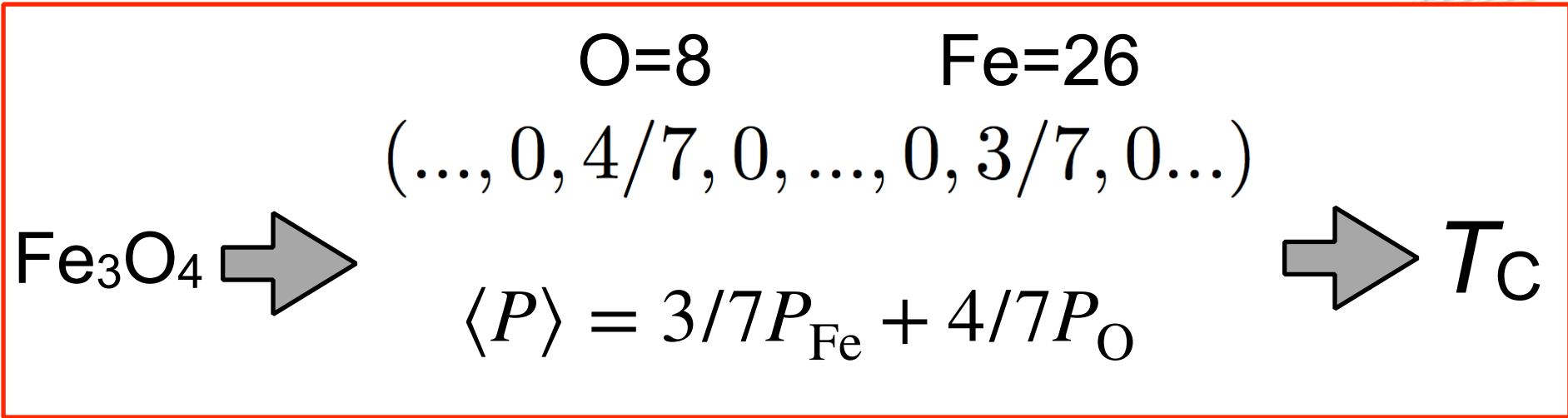


Approximation
Cell size
Broken symmetry
...

Non-magnons
Interaction range
....

Quantum effects
....

Often if we do not know the answer we cannot ‘predict’



1818 Training **746** Test

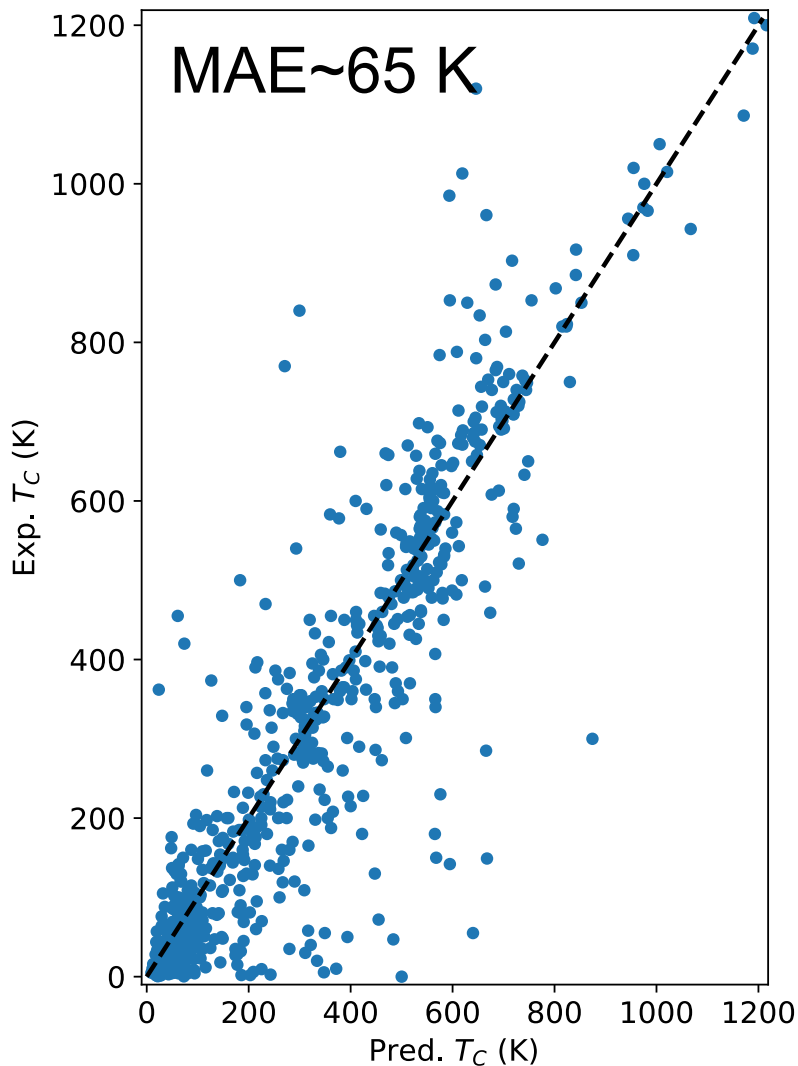
Some level of manual curation (e.g. paramagnets)

R^2	All	C10	C20	C40	C80	P10	P20	P40	P80
Ridge	0.50	0.48	0.50	0.51	0.50	0.27	0.30	0.32	0.37
KRR	0.70	0.72	0.76	0.74	0.71	0.70	0.71	0.70	0.69
NN	0.75	0.74	0.77	0.77	0.75	0.71	0.74	0.75	0.73
RF	0.83	0.78	0.80	0.80	0.81	0.73	0.75	0.74	0.73

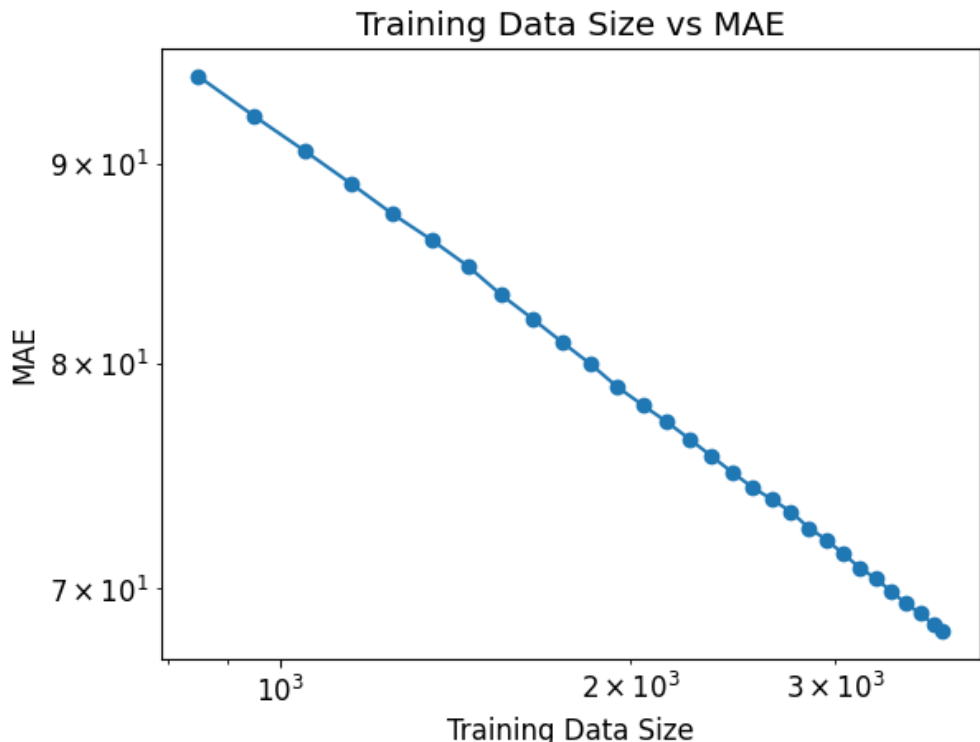
Ridge = Ridge Reg.; KRR = Kernel Ridge Reg.; NN = Neural Net; RF = Random Forest



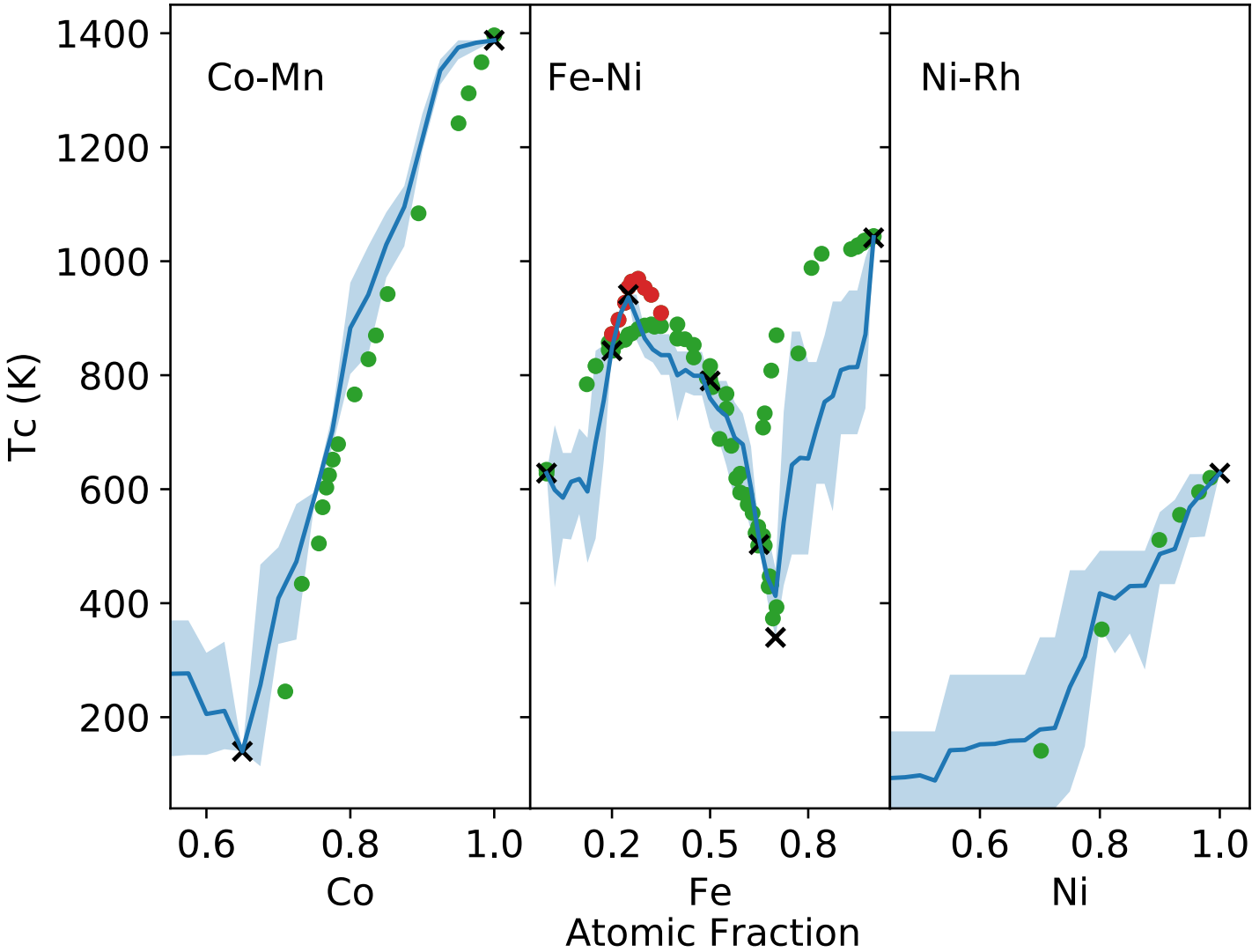
More accurate than *ab initio*

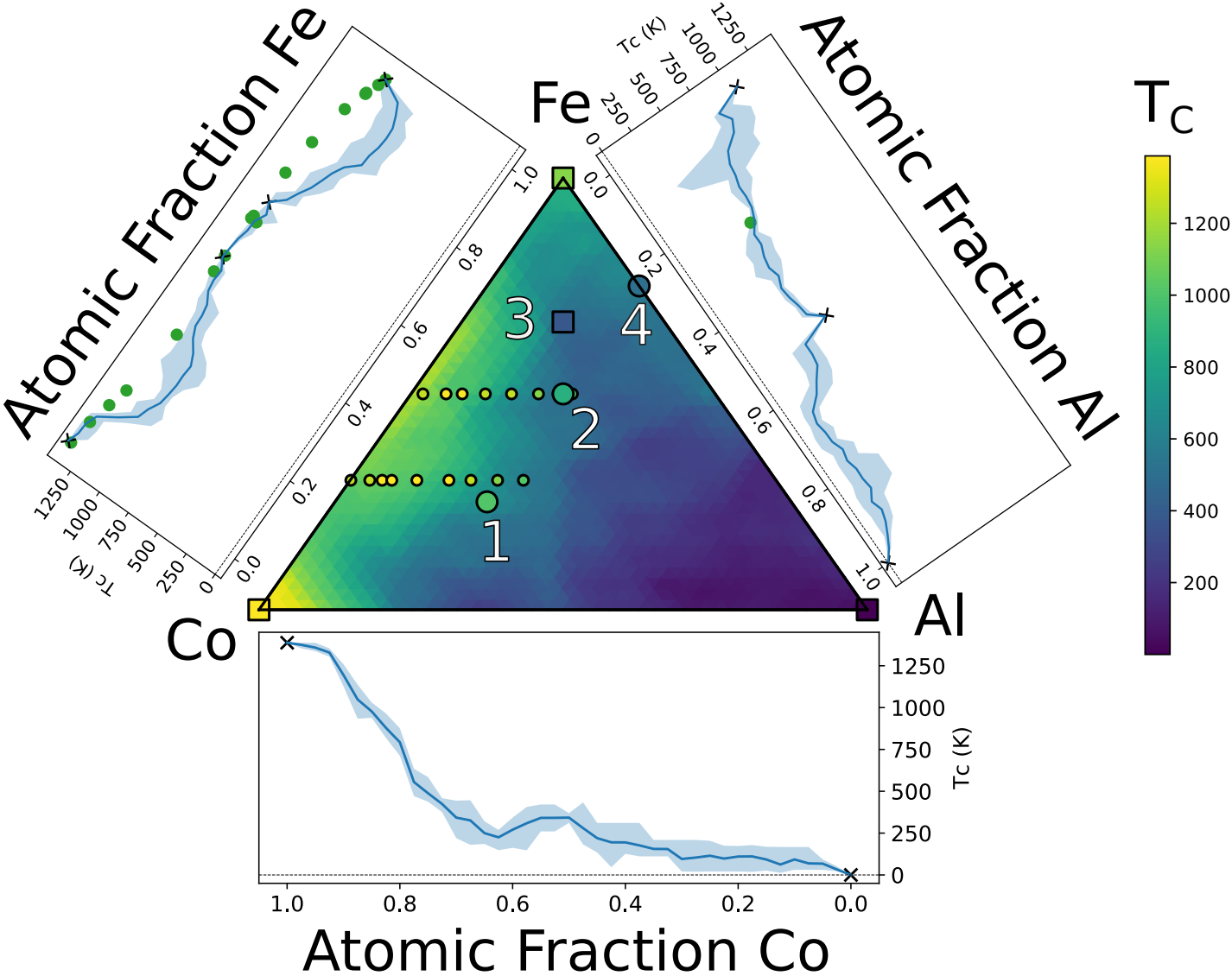


J.F. Belot, V. Taufour, SS and G.L.W. Hart,
Appl. Phys. Lett. **123**, 042405 (2023)



Gets better with volume





Data can be used to construct models !

Useful for an initial exploration

Extremely high throughput

Only experimental data are needed

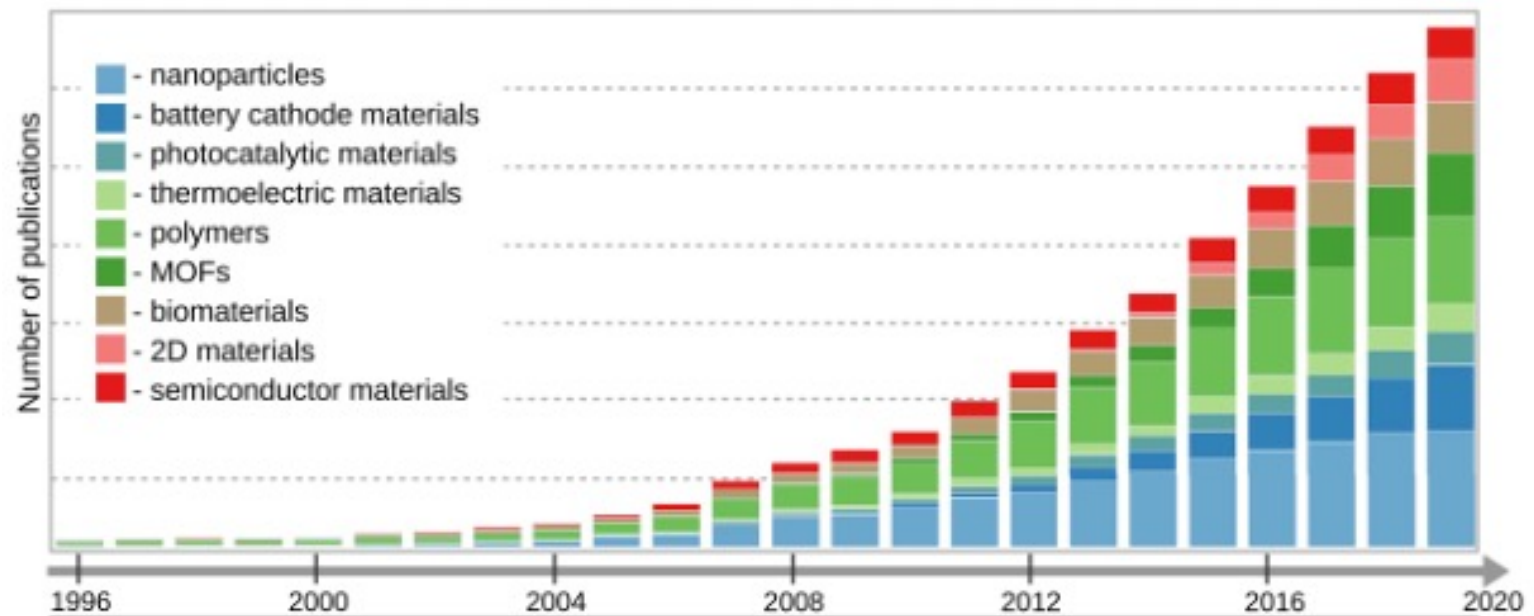
Structure difficult to introduce

Need to harvest more data

Need to accelerate models construction

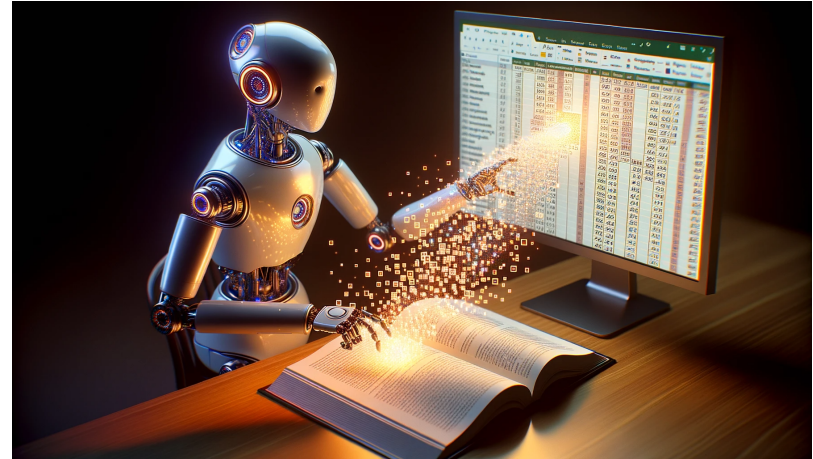
CRANN

- CRANN



Harvest more data

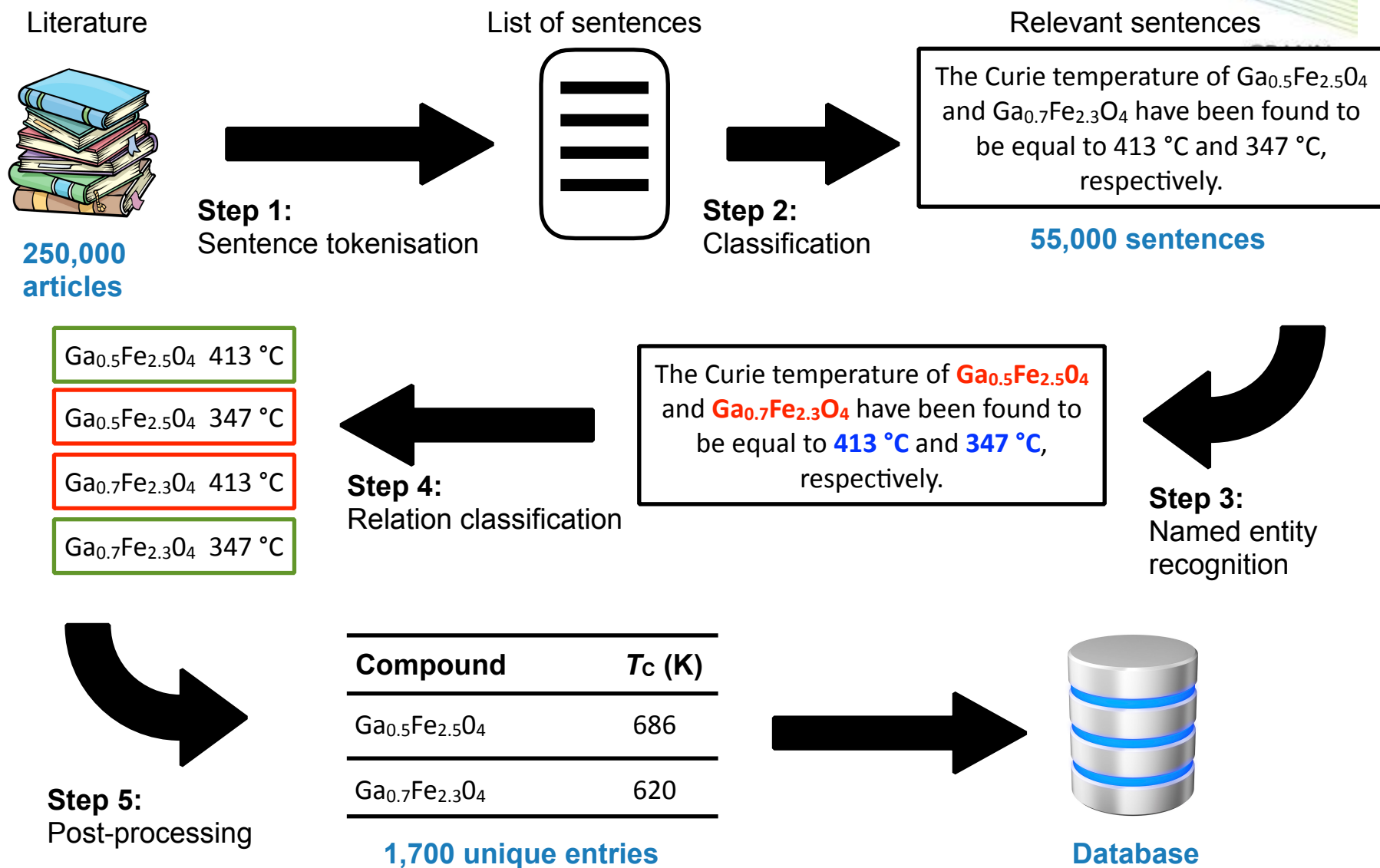
Concept: *construct an agent for precise information retrieval from scientific literature*



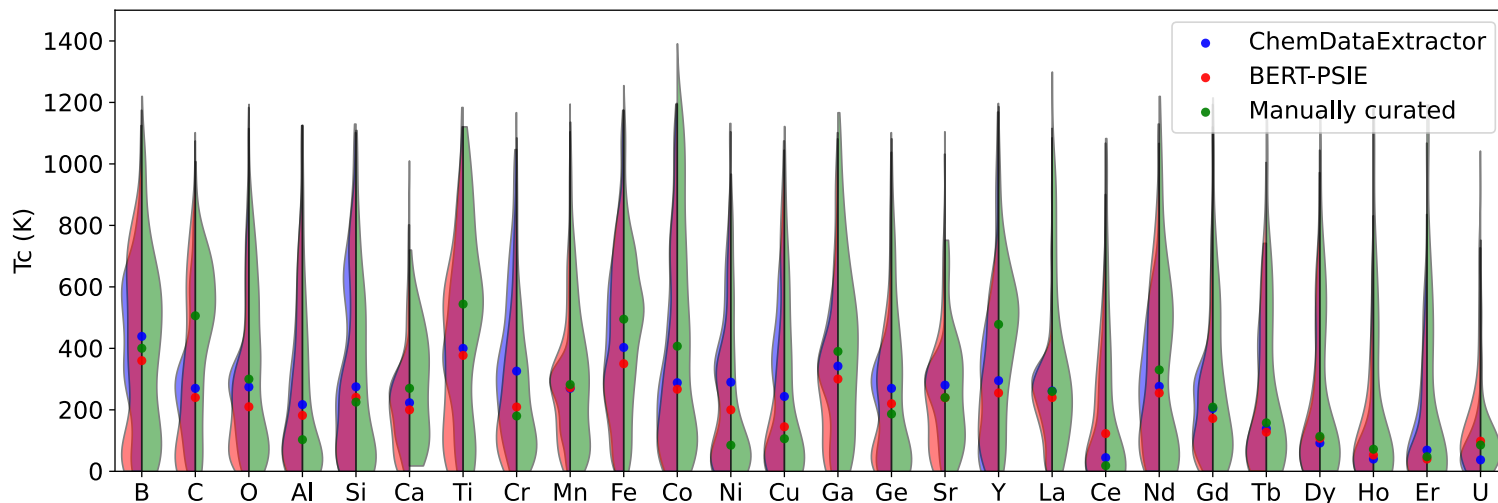
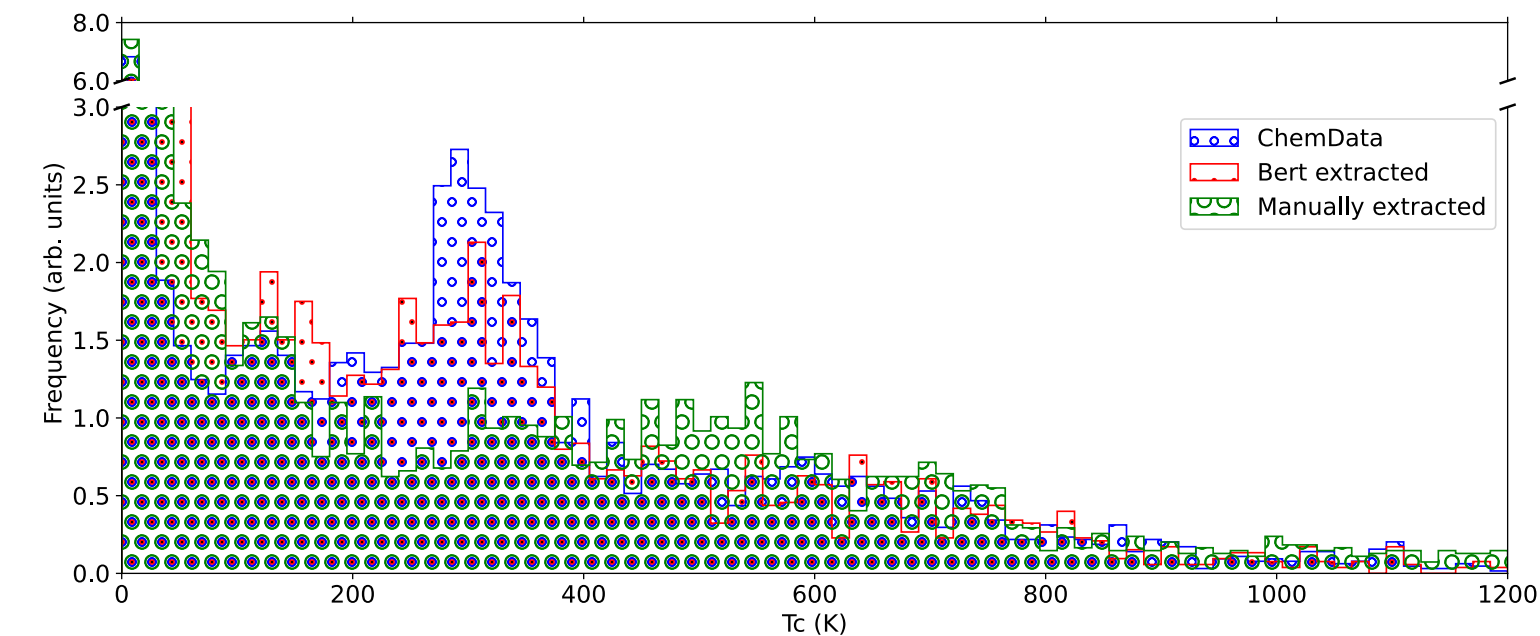
Requirements

- Workflow must be domain non-specific
- Little training necessary
- Little labour intensive
- Accurate
- High-throughput

Collecting data automatically

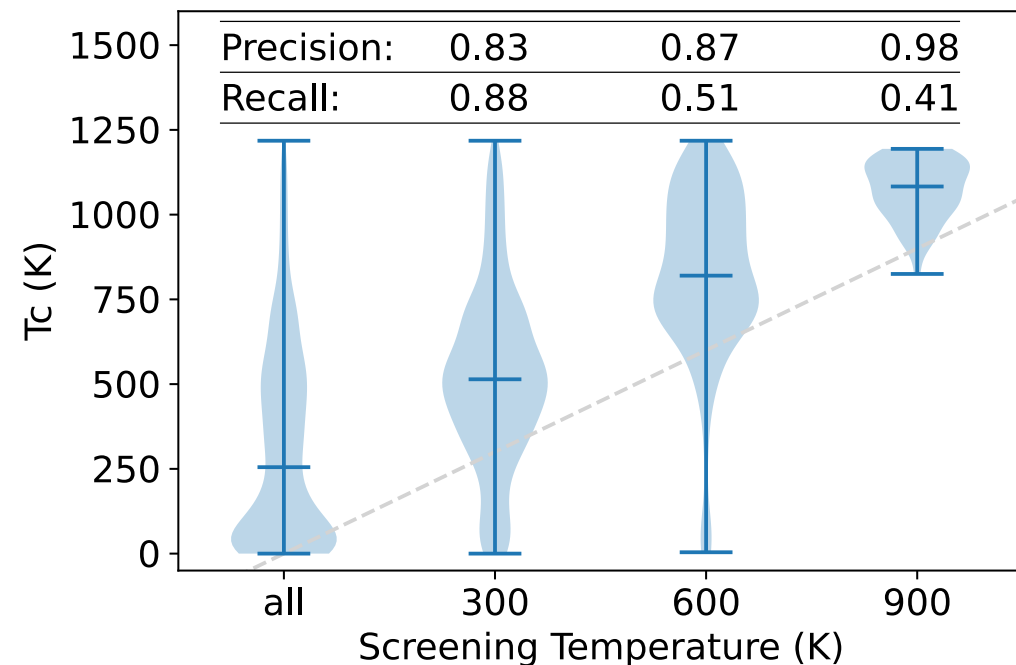


The quality of the database



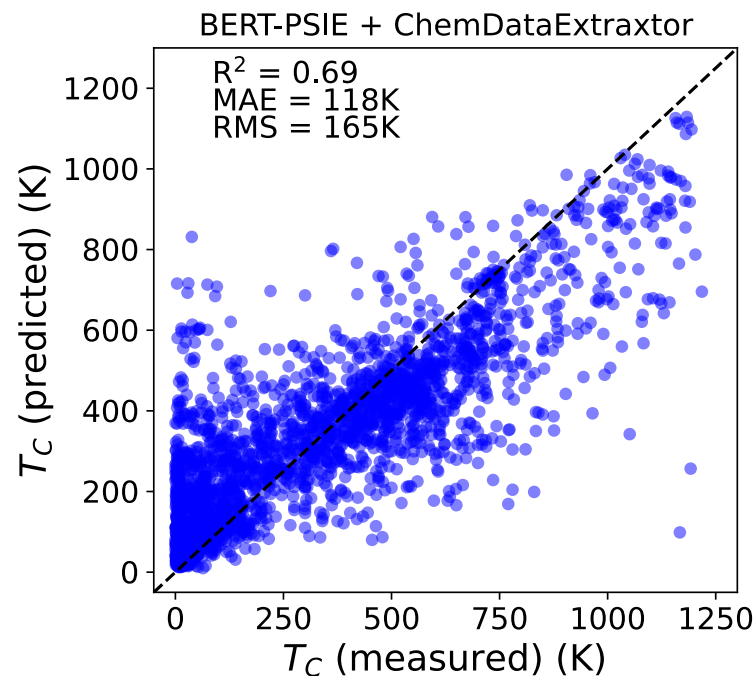
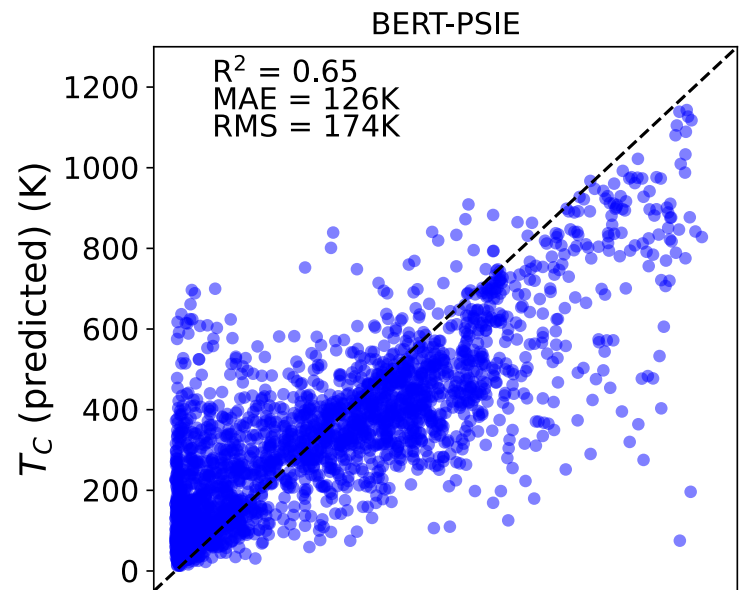
Database for T_c : comparison with manually curated and grammar-based databases

The database can be used



No curation this time

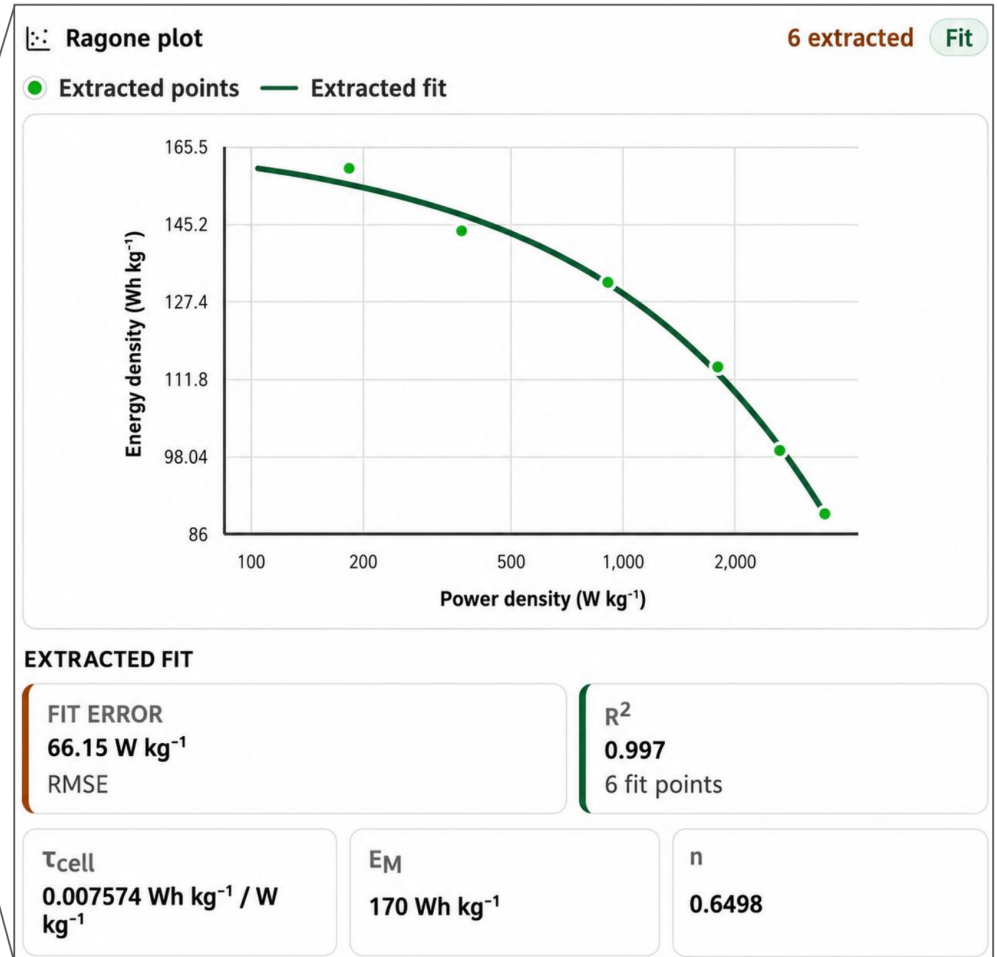
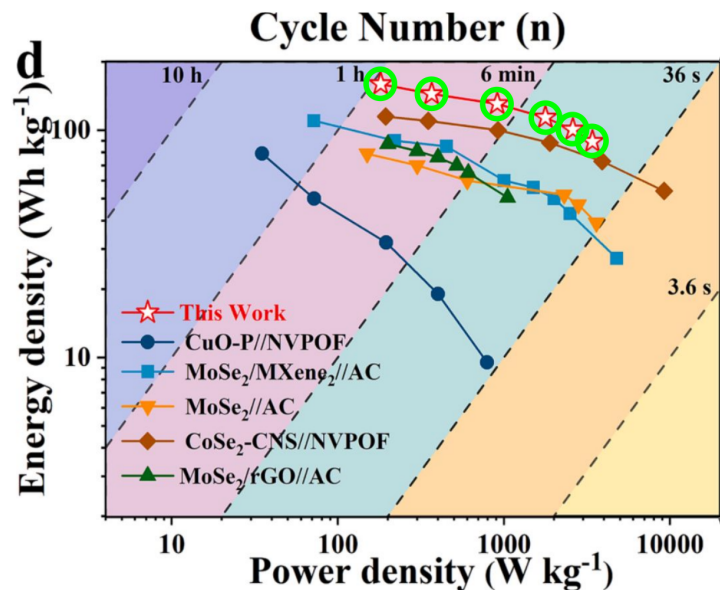
T_c : RF on the automatically extracted database



The database can be used



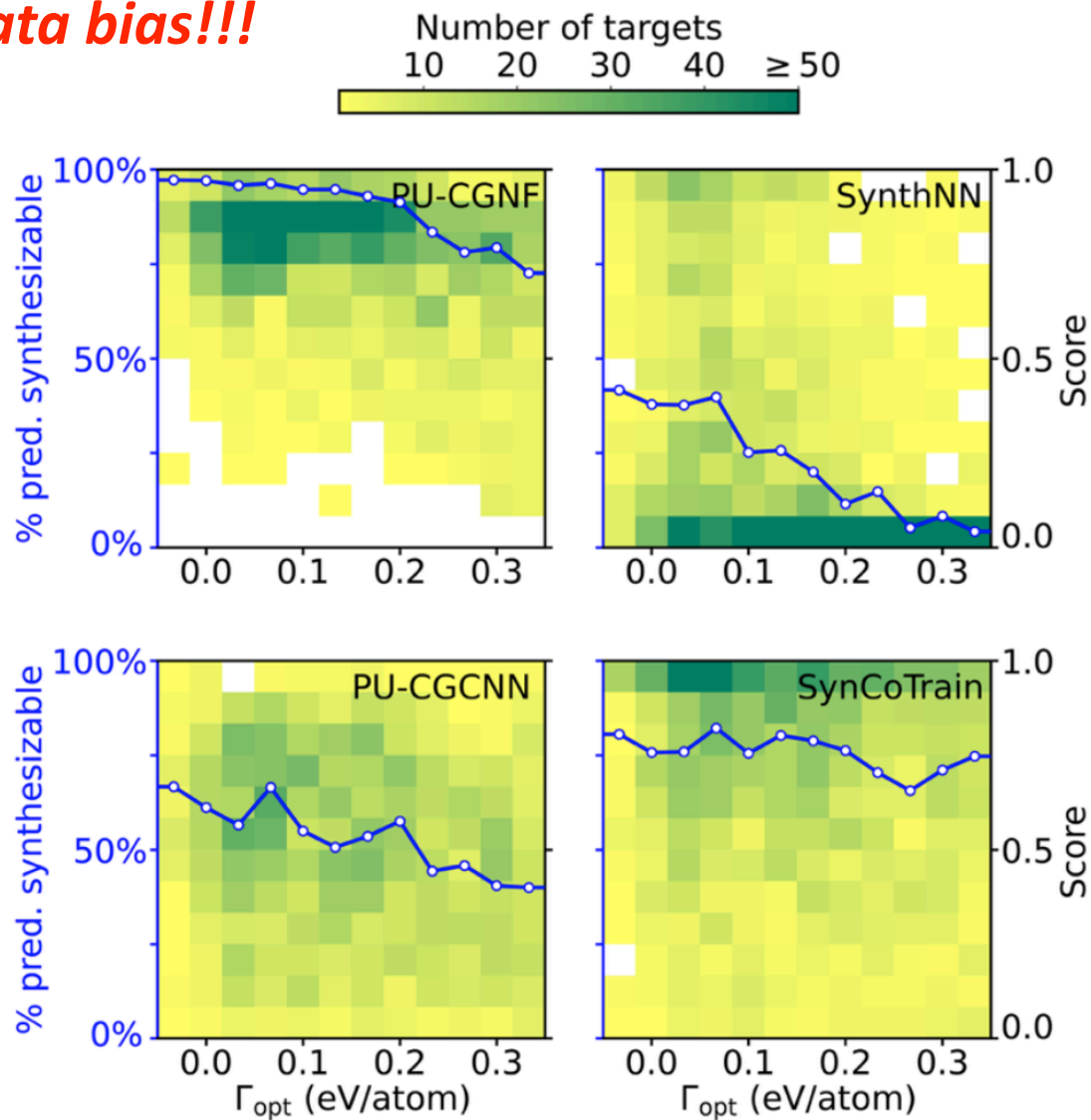
Now everything can be done with LLMs,
*enabling much more complex
extraction tasks (multimodal)*



The database can be used



Careful with data bias!!!

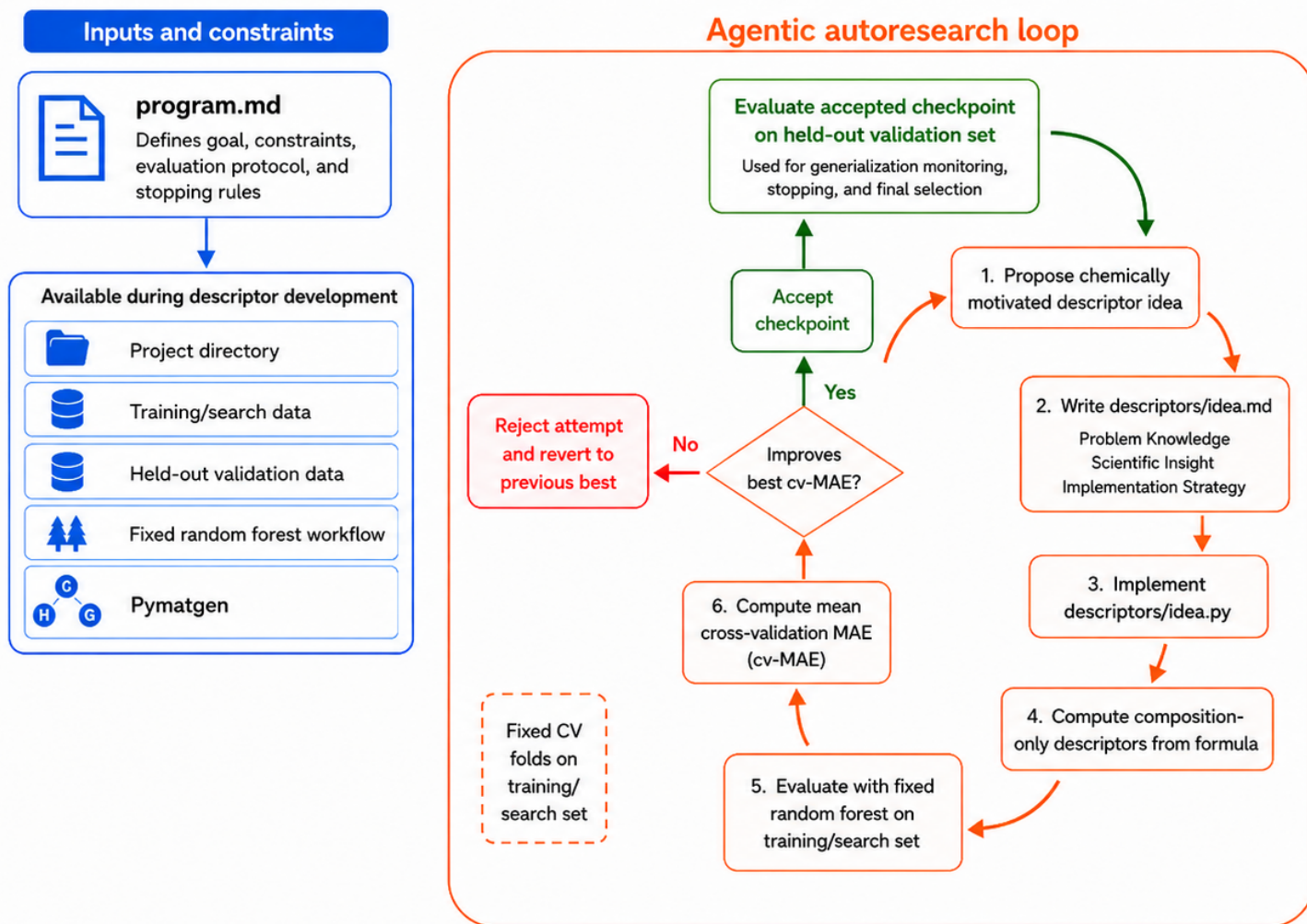




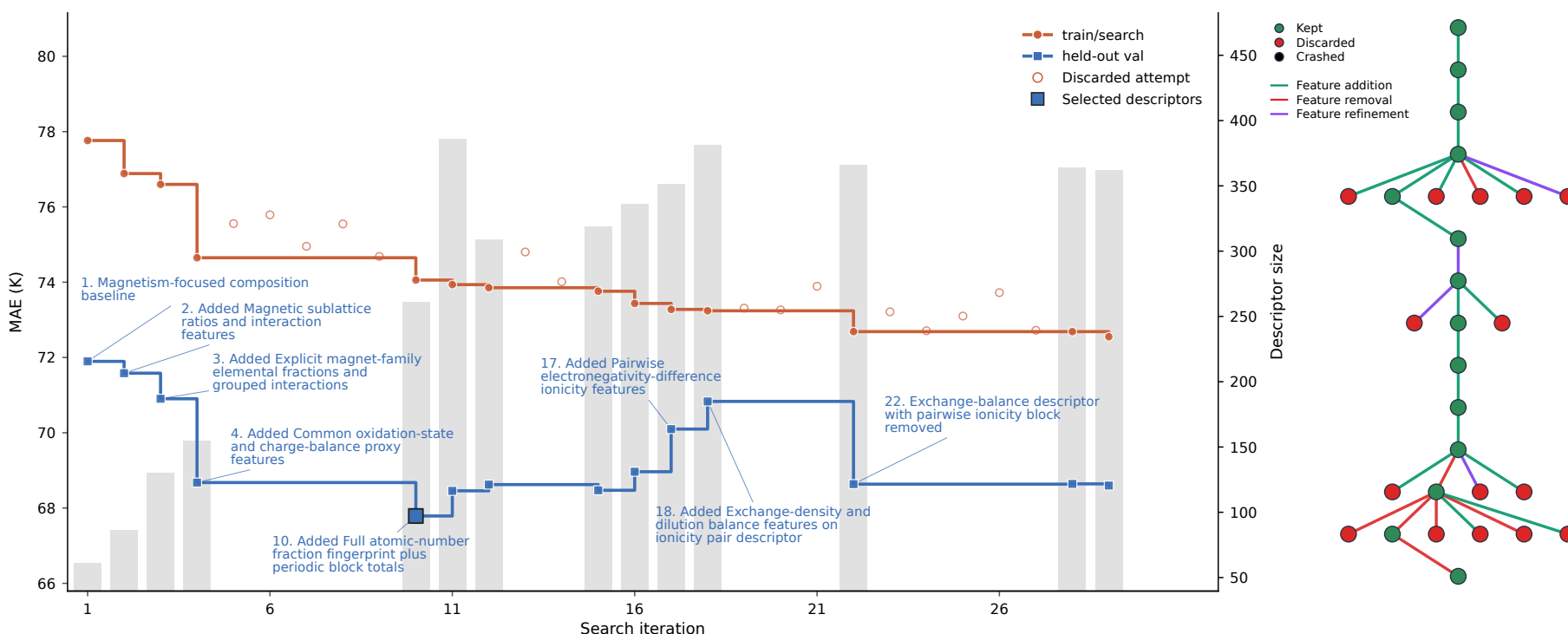
***Model construction acceleration
needs more sophisticated agents***

Model construction acceleration

Automat autoresearch workflow



Model construction acceleration





Data can be used to construct models !

Useful for an initial exploration

Extremely high throughput

Only experimental data are needed

Structure difficult to introduce

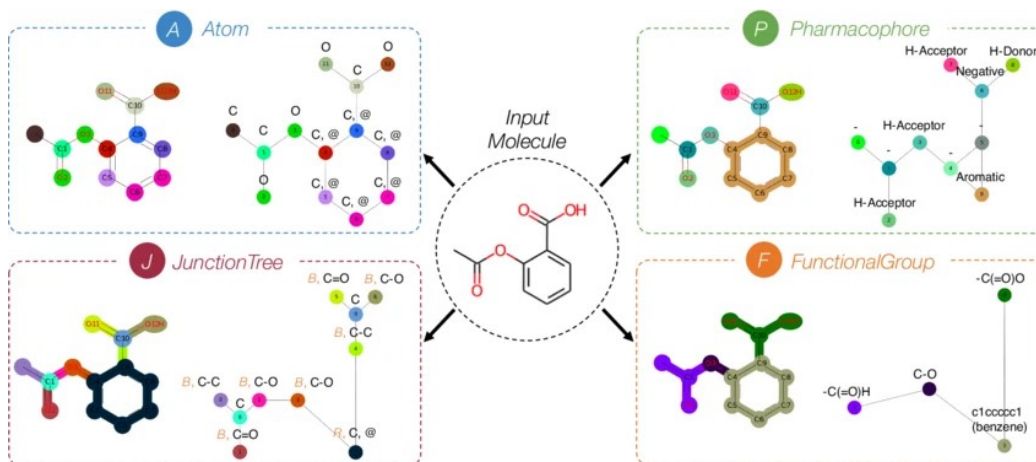
Need to harvest more data

High-throughput NLP method now developed

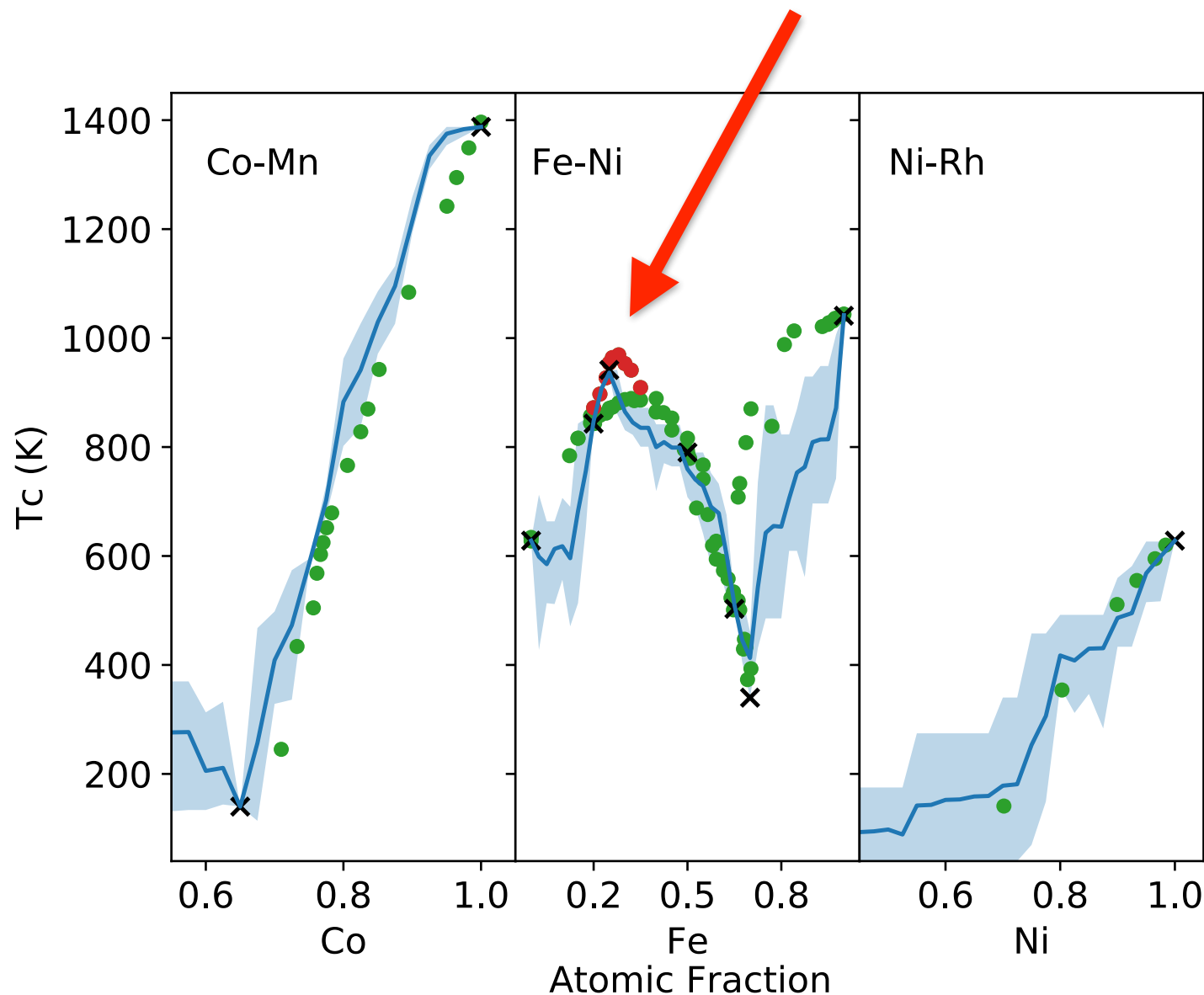
Need to accelerate models construction

Agents AI can help here

Structure Representations



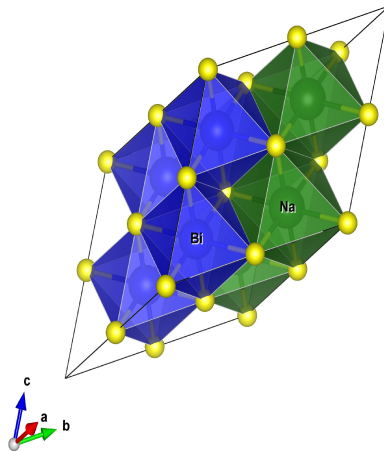
What do we do when the structure is important?



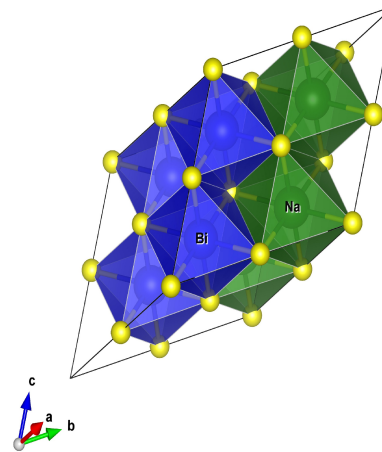
Structural information more difficult to include

Many conventional methods are not good for ML,
e.g. space groups

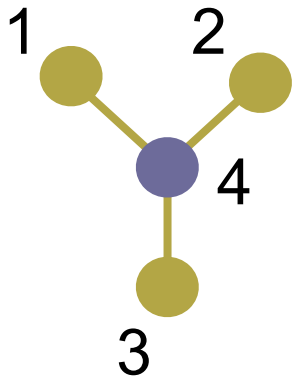
No. 225



No. 166

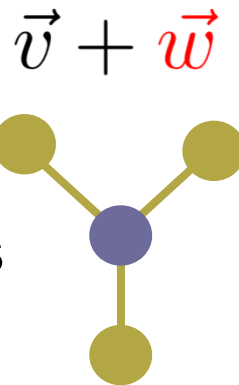


Local Structural Representations



Cartesian coordinates

$$\vec{v} = (\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4)$$

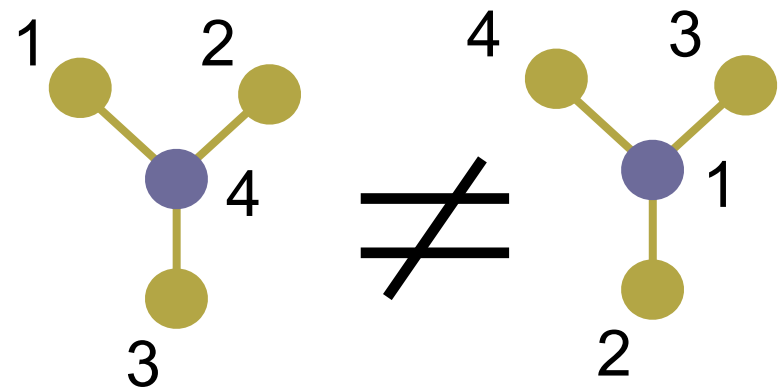


Invariant by translations

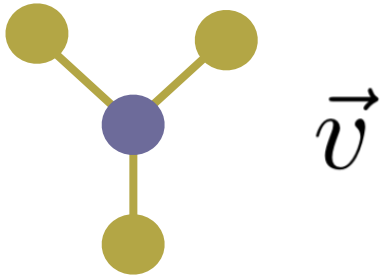
Locality fixes the problem

$$Q_{\text{local}}(\mathbf{r}^N) = \sum_{i=1}^N Q_i(\mathbf{r}'_i)$$

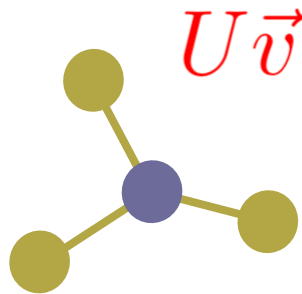
Locality fixes also invariance by permutation



Local Structural Representations

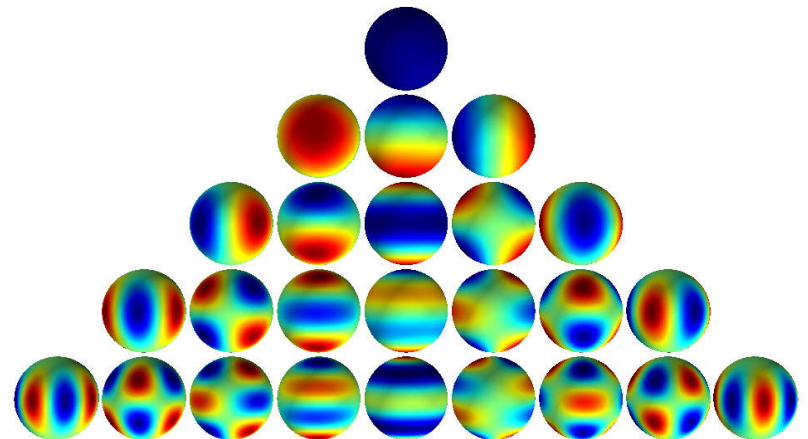
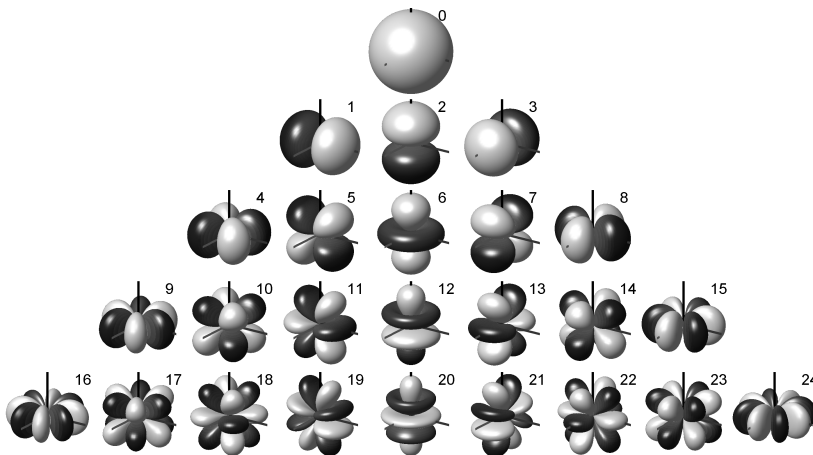


Invariant by rotations



Needs to be built in

Usually spherical harmonics expansion



Problems with local representations

1) Tend to be large (many features even for simple structures)

2) They are not invertible

Molecule	→	Representation
Representation	→	????

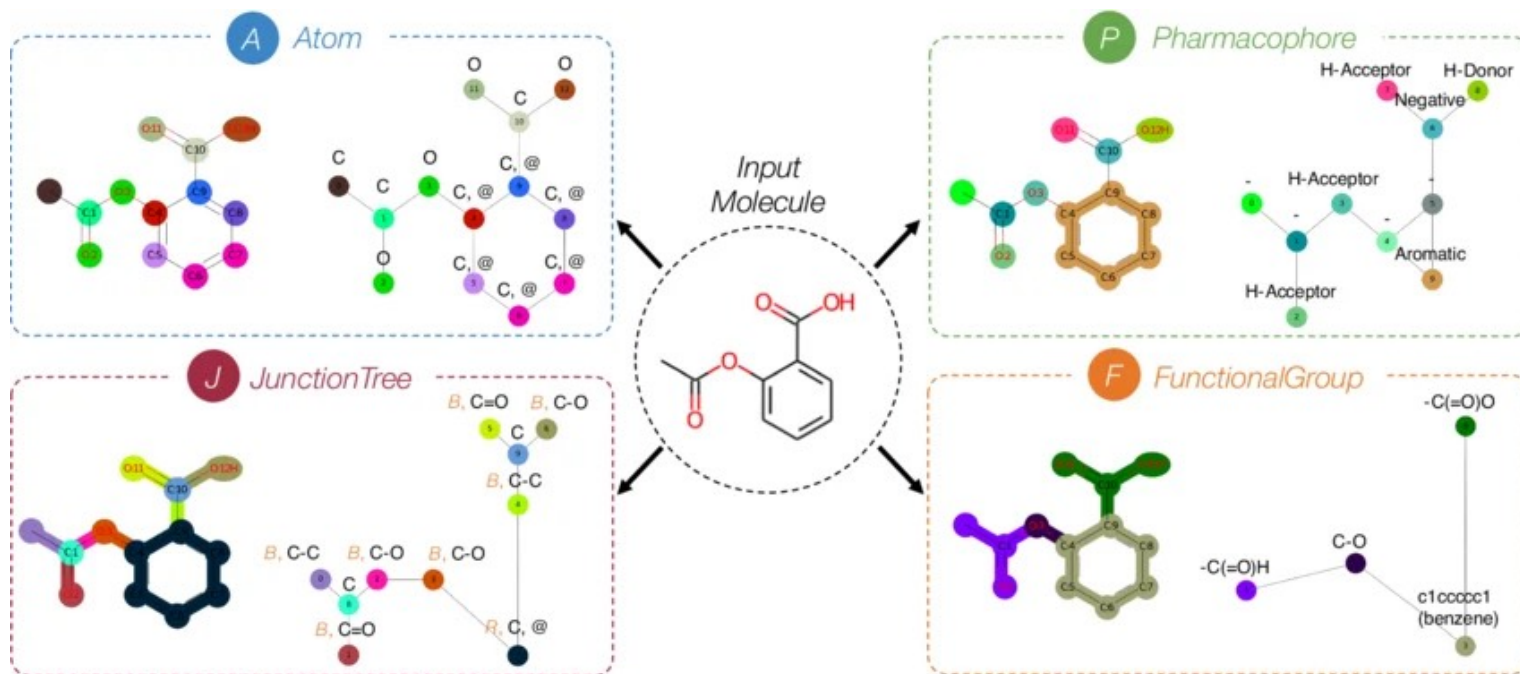
3) Not so easy to implement

4) Often not complete (same representation to different molecules)

5) Cannot describe global quantities

Global representations

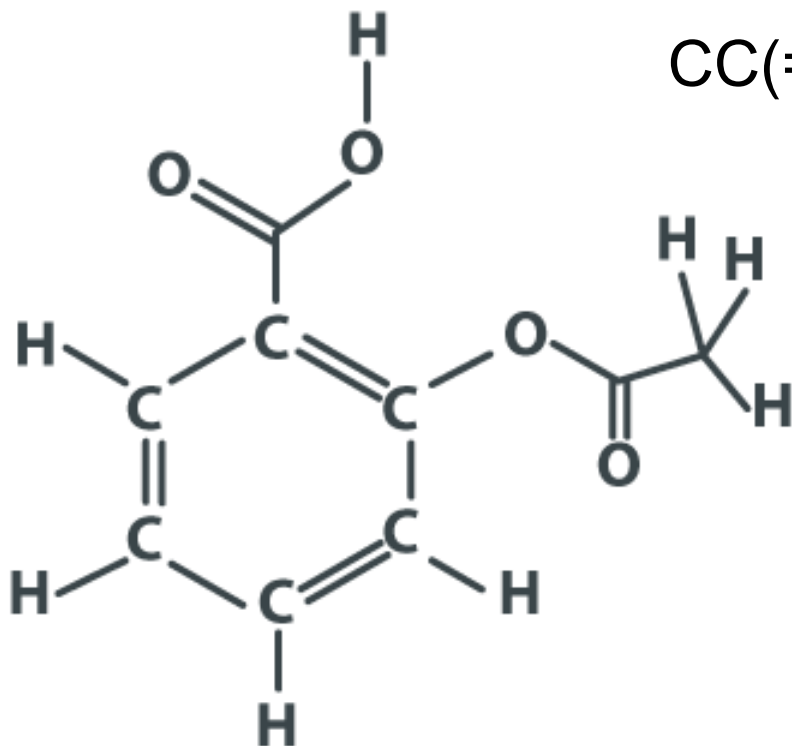
Graphs



- Very effective and universal
- Equivariance can be built in
- Tend to be complex
- Need to be used with deep models and lots of data

Global representations

SMILES / SELFIES

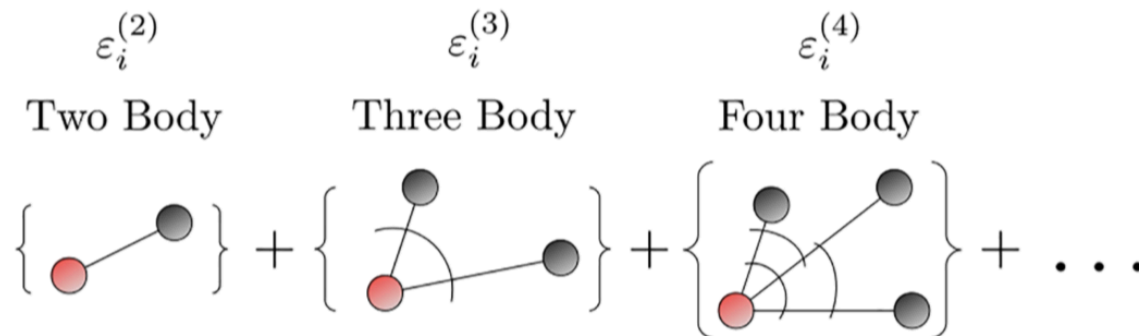


CC(=O)OC1=CC=CC=C1C(=O)O

Lexical representation (good for generative models)

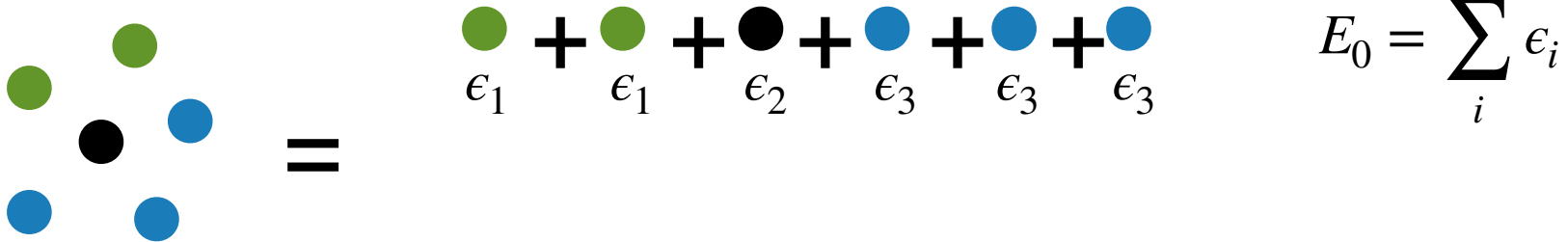
No 3D info (polymorphs?)

A cluster expansion for phase diagrams





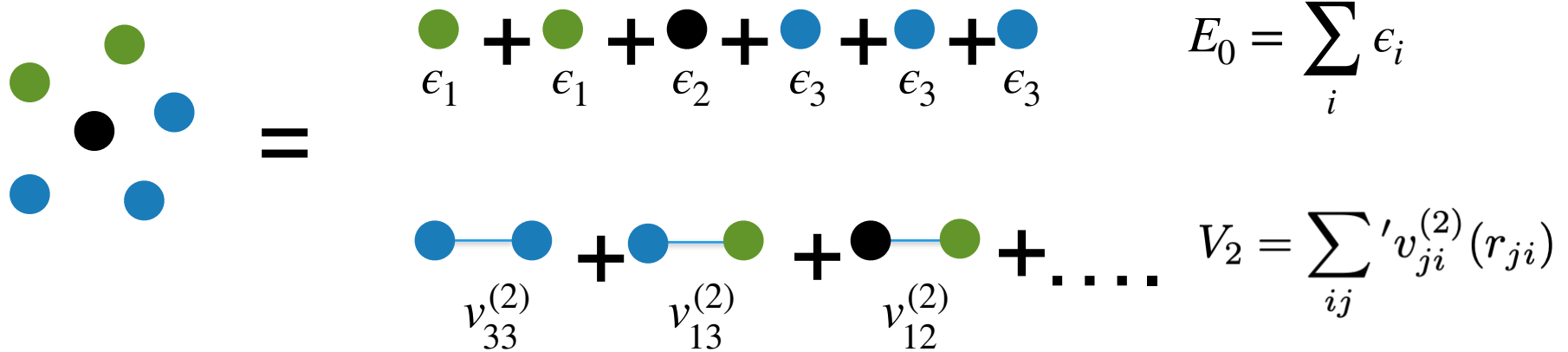
Jacobi-Legendre potential (cluster expansion)



$$E_{\text{short}} = E_0$$

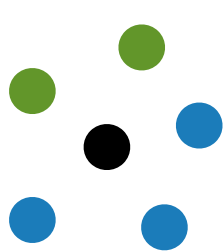


Jacobi-Legendre potential (cluster expansion)



$$E_{\text{short}} = E_0 + V_2$$

Jacobi-Legendre potential (cluster expansion)



=

$$\epsilon_1 + \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_3 + \epsilon_3$$

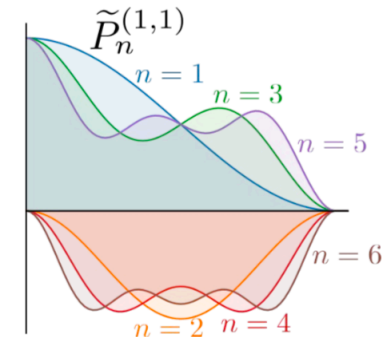
$$E_0 = \sum_i \epsilon_i$$

$$v_{33}^{(2)} + v_{13}^{(2)} + v_{12}^{(2)} + \dots$$

$$V_2 = \sum_{ij} v_{ji}^{(2)}(r_{ji})$$

$$v_{ji}^{(2)}(r_{ji}) = \sum_{n>}^{n_{\max}} a_n^{ji} \tilde{P}_n^{(\alpha,\beta)} \left(\gamma \cos \left(\pi \frac{r_{ji}}{r_{\text{cut}}} \right) \right)$$

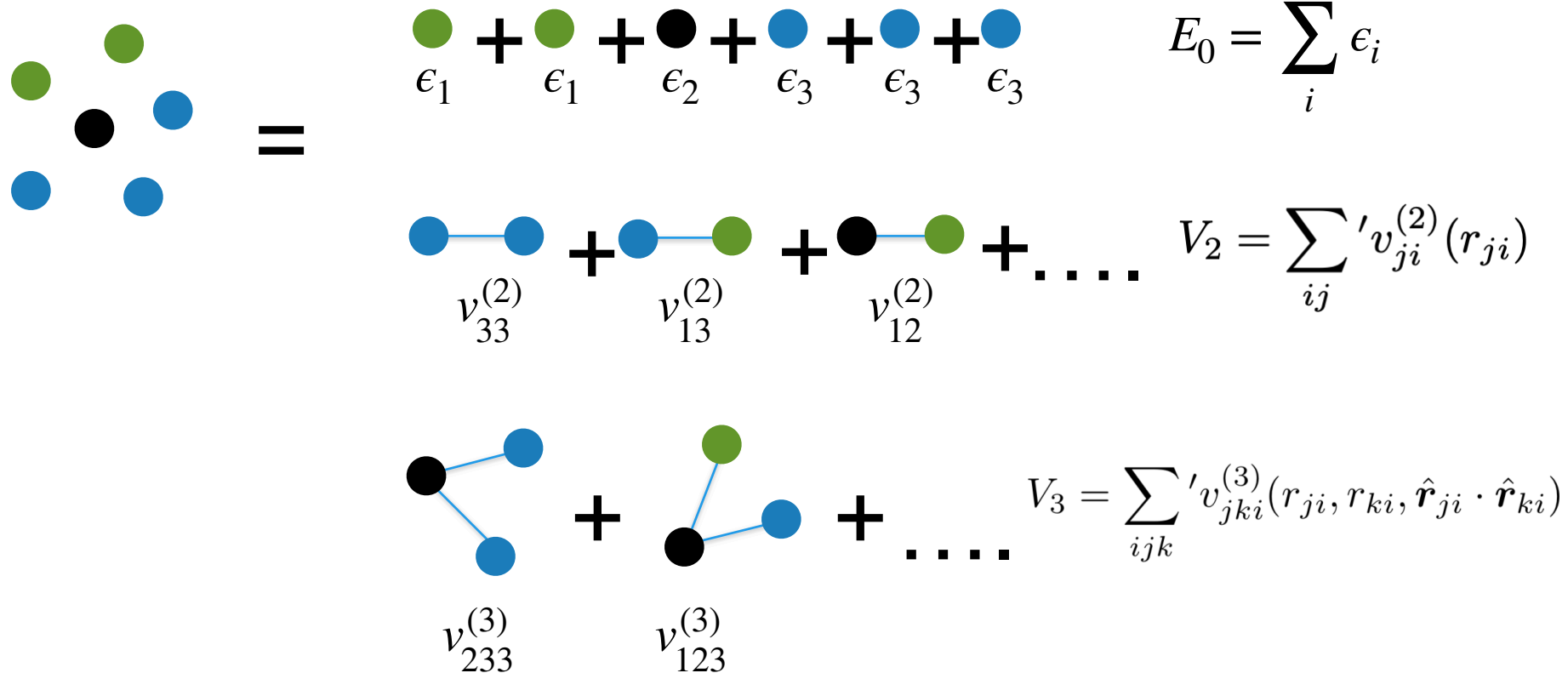
Shifted Jacobi polynomials



$$E_{\text{short}} = E_0 + V_2$$



Jacobi-Legendre potential (cluster expansion)



$$E_{\text{short}} = E_0 + V_2 + V_3$$



Jacobi-Legendre potential (cluster expansion)



$$E_0 = \sum_i \epsilon_i$$

$$V_2 = \sum_{ij} v_{ji}^{(2)}(r_{ji})$$

$$V_3 = \sum_{ijk} v_{jki}^{(3)}(r_{ji}, r_{ki}, \hat{\mathbf{r}}_{ji} \cdot \hat{\mathbf{r}}_{ki})$$

$$v_{jki}^{(3)}(r_{ji}, r_{ki}, s_{jki}) = \sum_{n_1, n_2=2}^{n_{\max}} \sum_{l=0}^{l_{\max}} a_{n_1 n_2 l}^{jki} \overline{P}_{n_1 j i}^{(\alpha, \beta)} \overline{P}_{n_2 k i}^{(\alpha, \beta)} P_l$$

“Doubly vanishing” Jacobi polynomials

Legendre polynomials

Jacobi-Legendre polynomials to spherical harmonics

Addition theorem for spherical harmonics

$$P_l(\hat{\mathbf{r}}_{ji} \cdot \hat{\mathbf{r}}_{ki}) = \frac{4\pi}{2l+1} \sum_{m=-l}^l Y_l^{m*}(\hat{\mathbf{r}}_{ji}) Y_l^m(\hat{\mathbf{r}}_{ki})$$

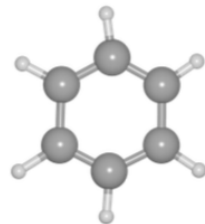
One can establish a connection between the JL cluster expansion and atom-centred representations

This can be generalised with multi-polar spherical harmonics

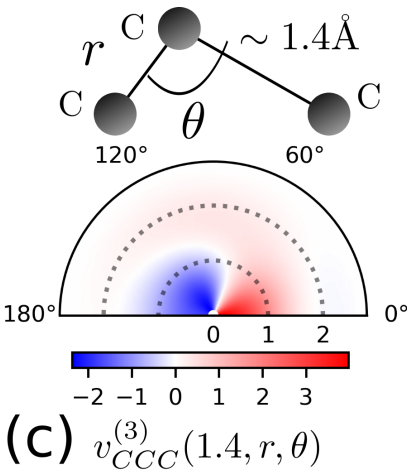
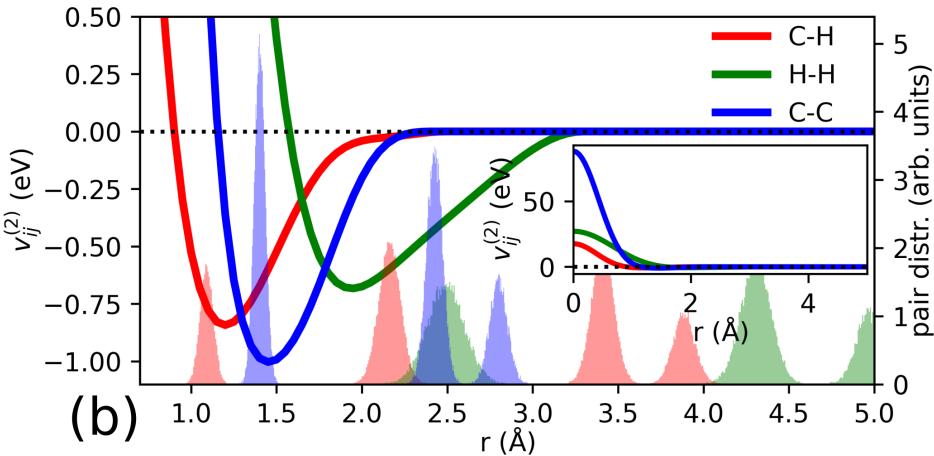
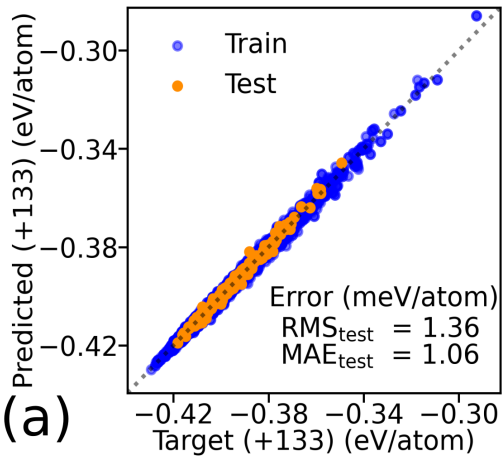


DFT-accurate force fields

$$E_{\text{short}} = E_0 + V_2 + V_3 + V_4$$

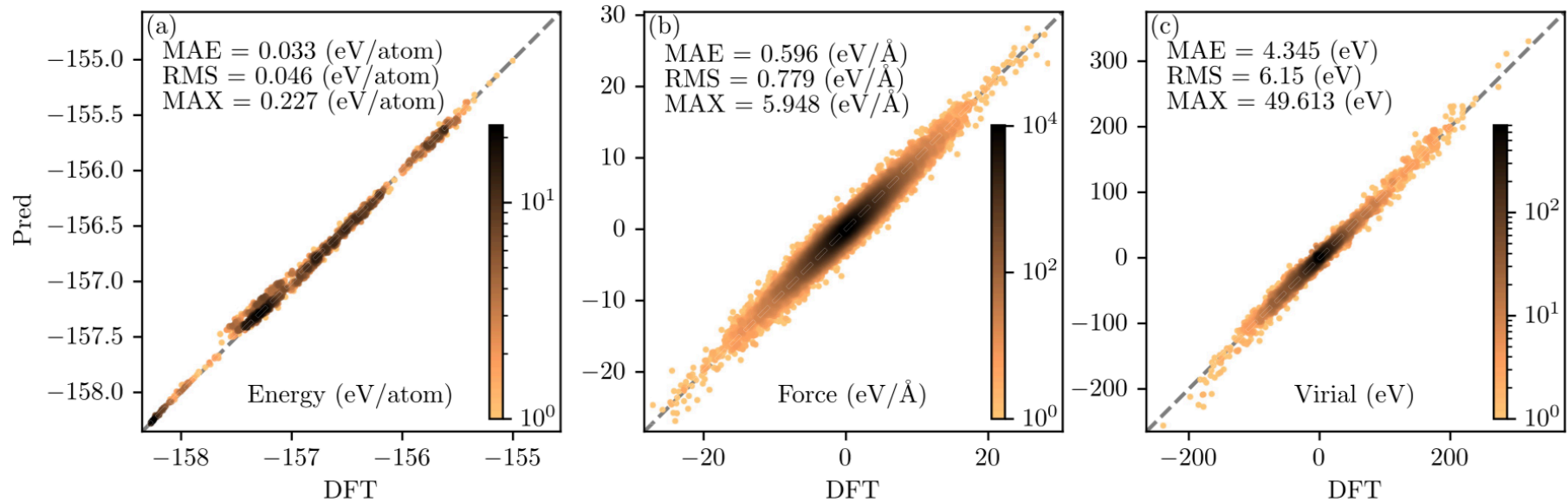


Benzene Molecule

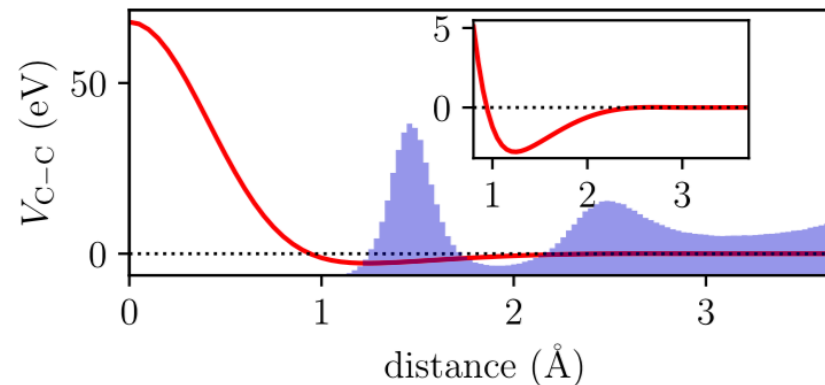


$$E_{\text{short}} = E_0 + V_2 + V_3 + V_4$$

Amorphous Carbon

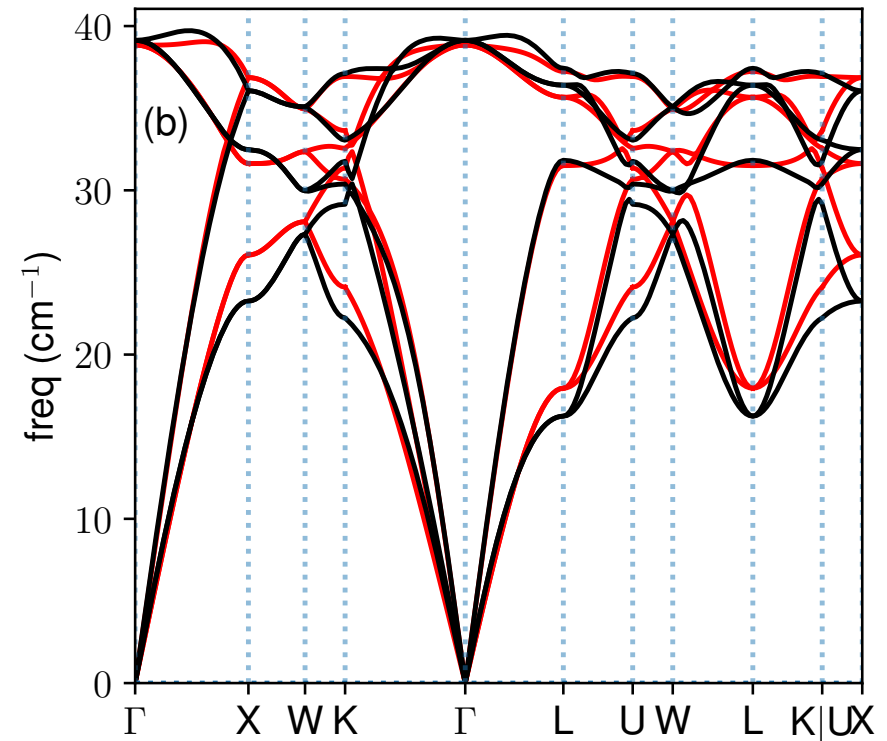
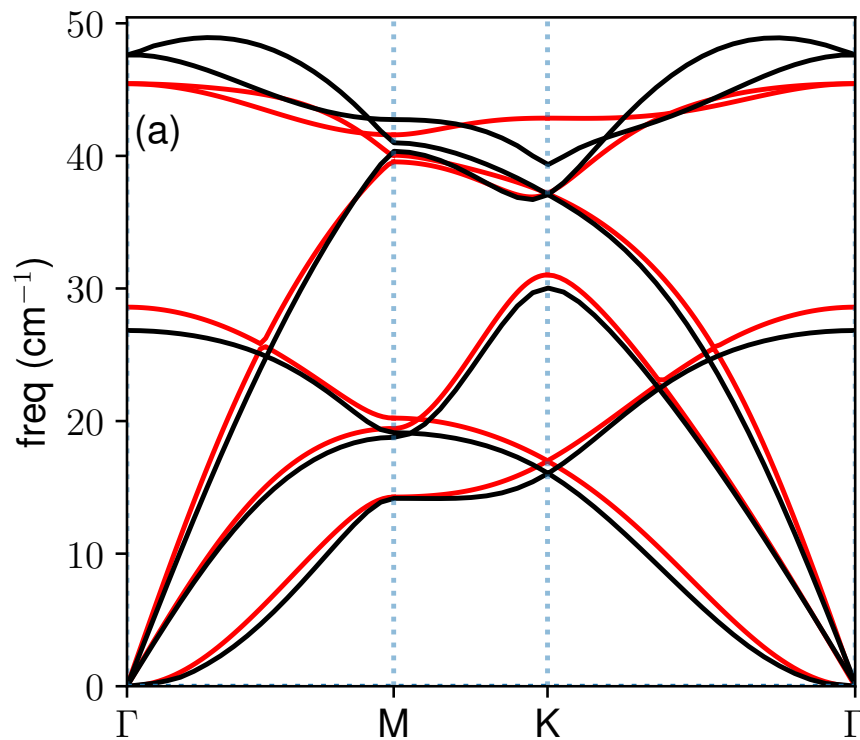


465 features

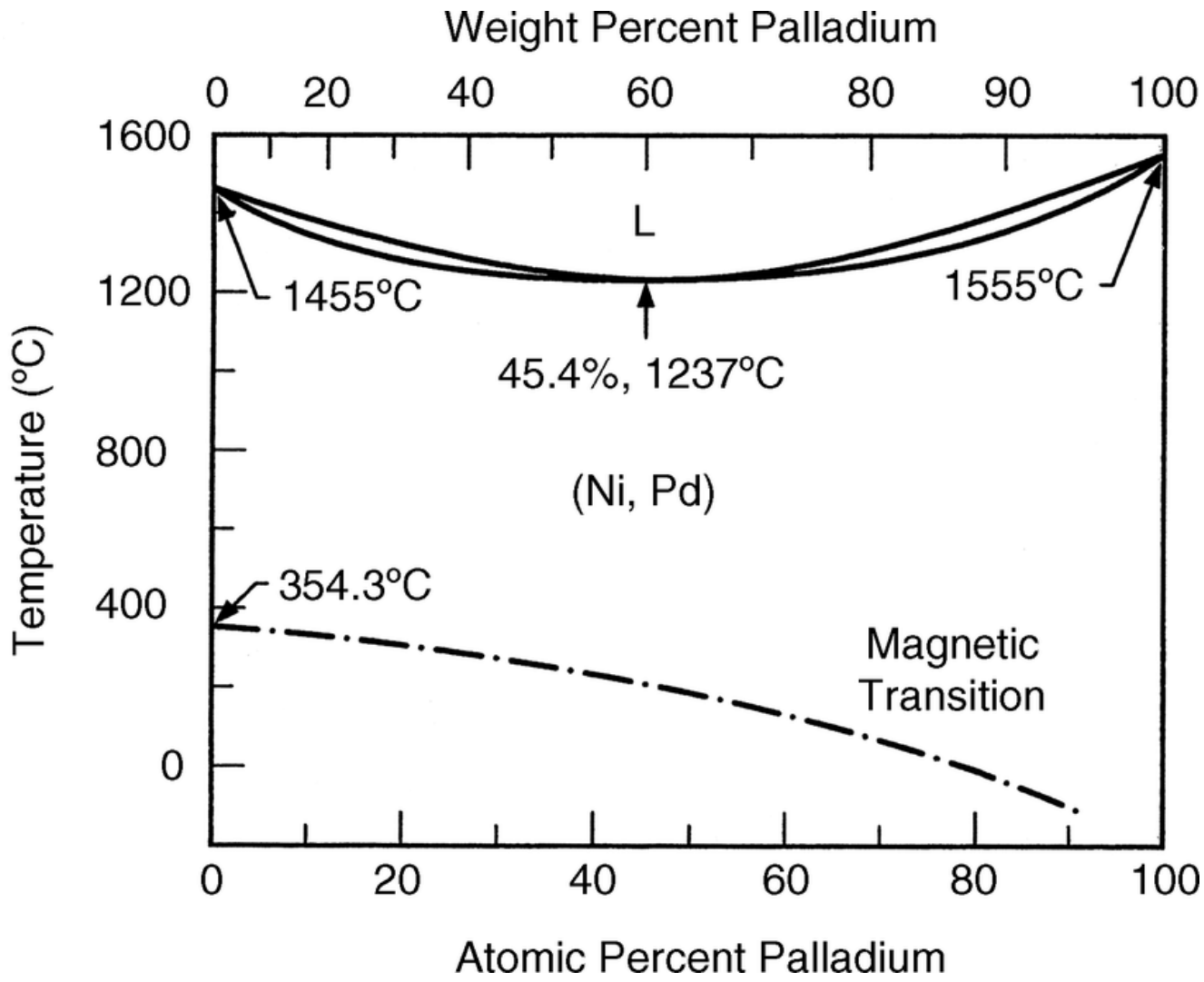


$$E_{\text{short}} = E_0 + V_2 + V_3 + V_4$$

Amorphous Carbon



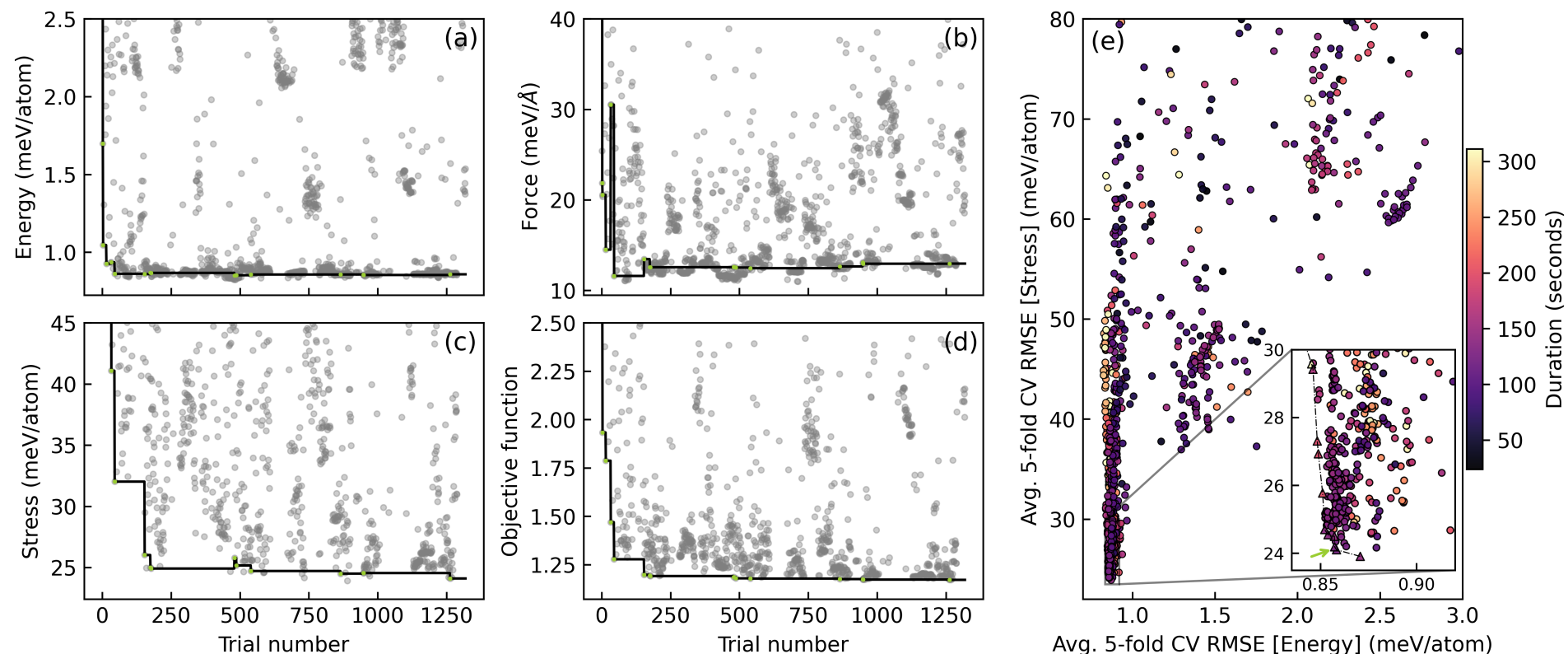
The **Ni-Pd** phase diagram



Very accurate JLP

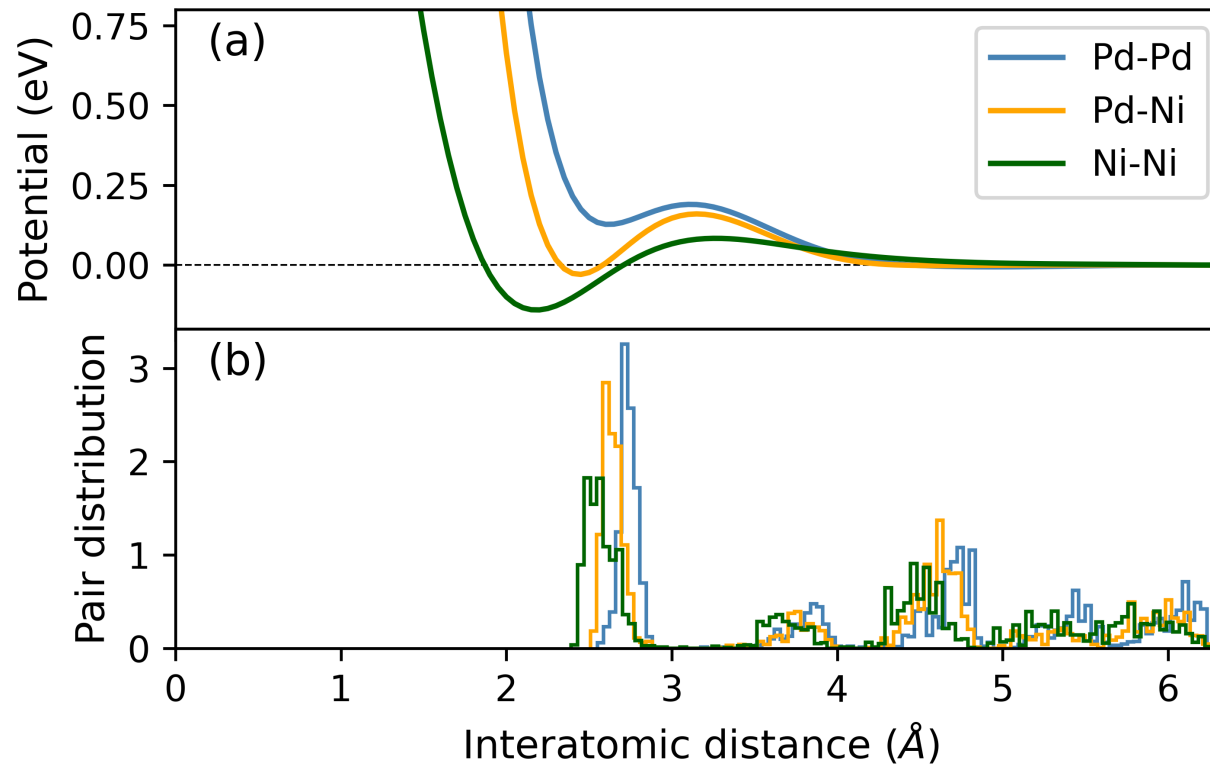


Multiple models possible: following Pareto front important



3B model: 873 parameters

Very accurate JL model

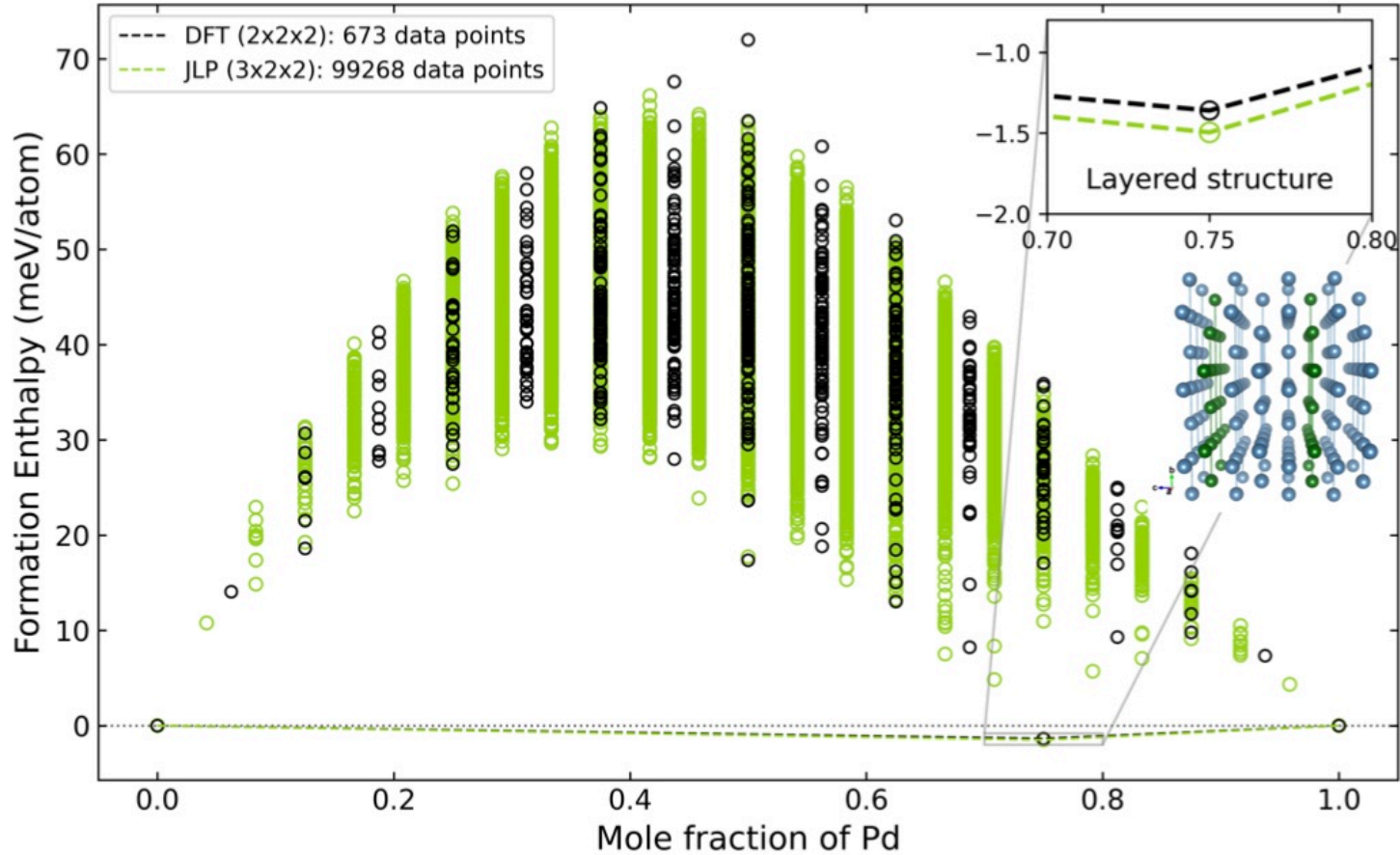


Massive throughput

Ni-Pd phase diagram



ΔH

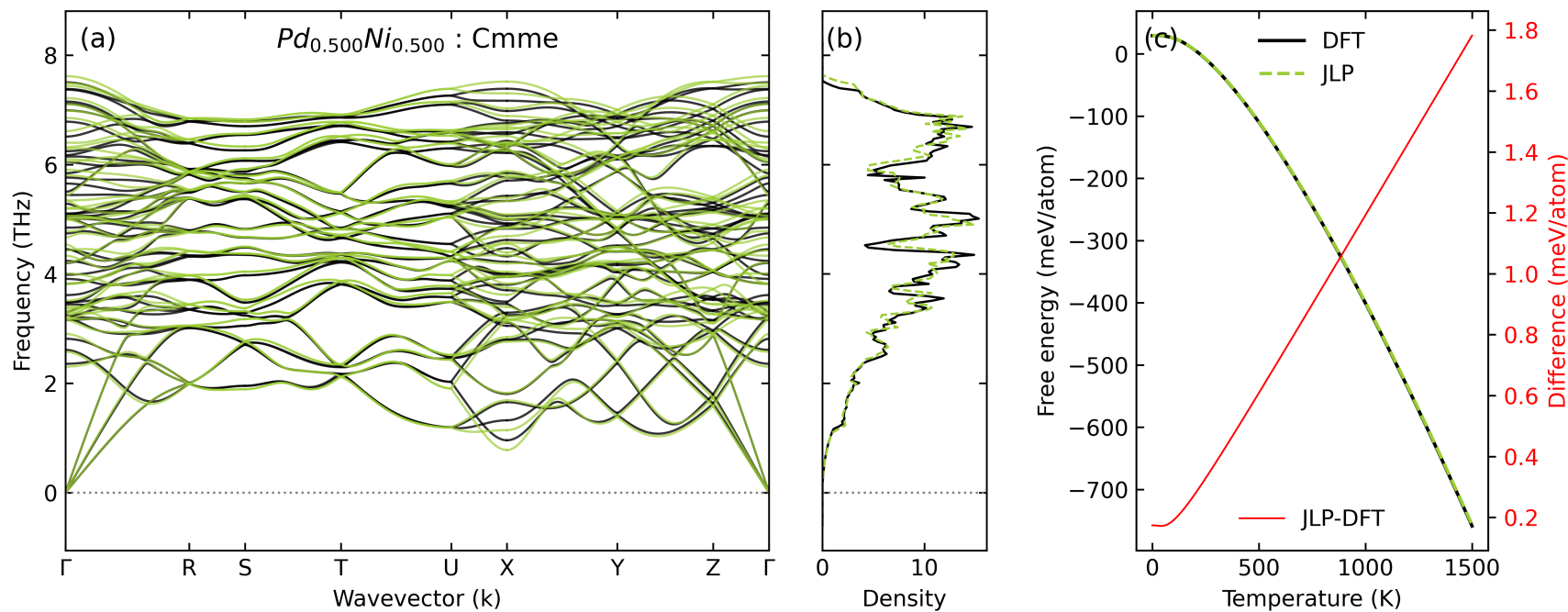


$$\Delta G = \Delta H - T\Delta S$$

$$Z^{\text{HA}}(x, T) = \sum_{\sigma \in \text{unique}} g(\sigma) \exp \left(-\frac{E(\sigma) + F^{\text{vib}}(\sigma, T)}{k_{\text{B}}T} \right)$$

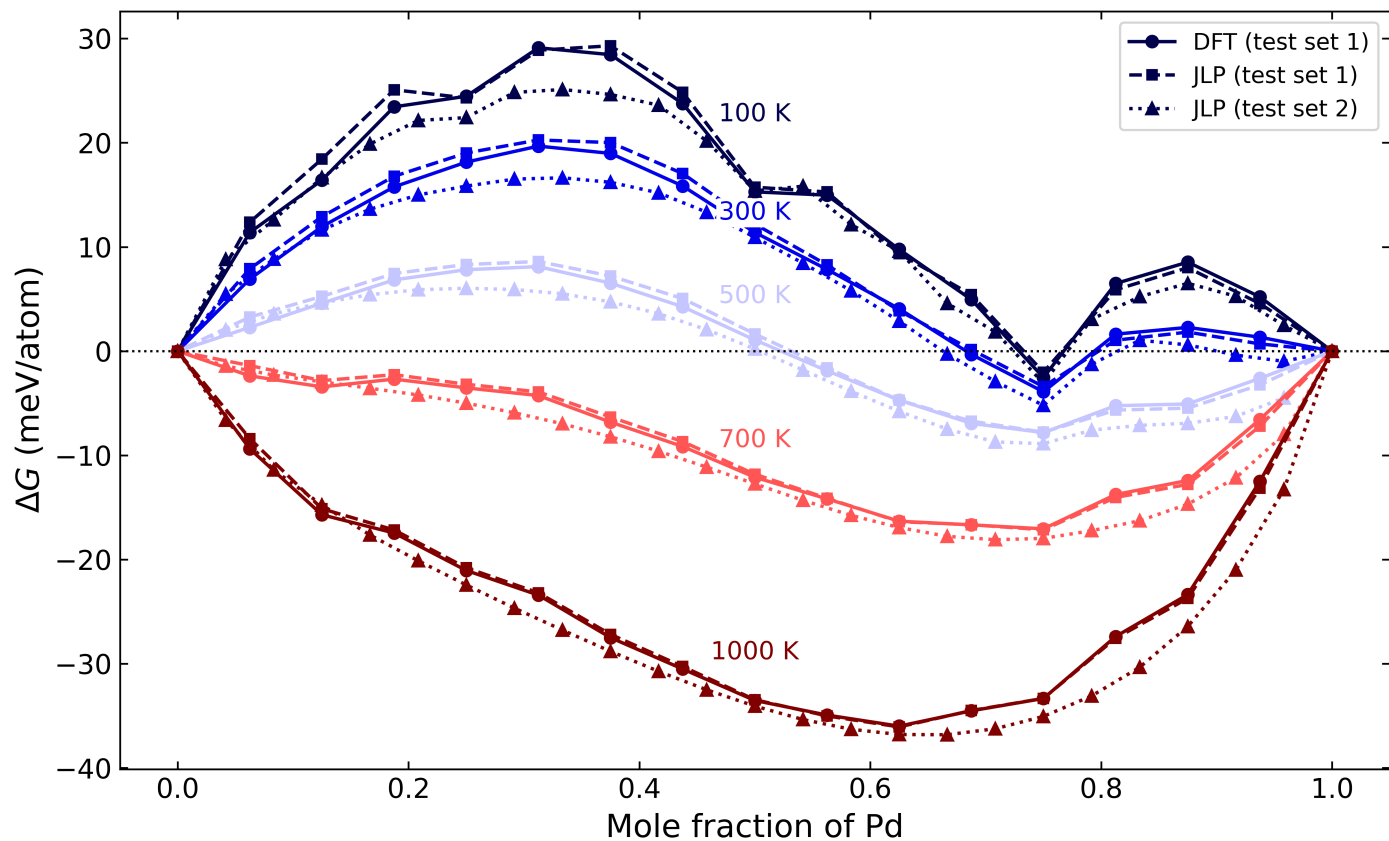
$$F^{\text{vib}}(\sigma, T) = \frac{1}{2} \sum_{\mathbf{q}j} \hbar \omega(\mathbf{q}j) + k_{\text{B}}T \sum_{\mathbf{q}j} \ln[1 - e^{-\hbar \omega(\mathbf{q}j)/k_{\text{B}}T}]$$

This is a big task: phonon calculations for 100,000 structures from supercells containing 24 atoms !!



Massive throughput

Ni-Pd phase diagram



>15,000,000 energy calculations
~100,000 phonon calculations

Very general framework for ML in computational materials science

Low-body orders interpretable

Very efficient and avoids deep learning

Extendible to vector and tensor fields and working with vector and tensor inputs



COMPUTATIONAL SPINTRONICS

SANVITO RESEARCH GROUP
TRINITY COLLEGE, DUBLIN

