

ML-guided Materials Discovery & Synthesis



Matbench Discovery

Can ML predict thermodynamic stability?

If yes, which ML method works best?

What discovery acceleration can we expect?

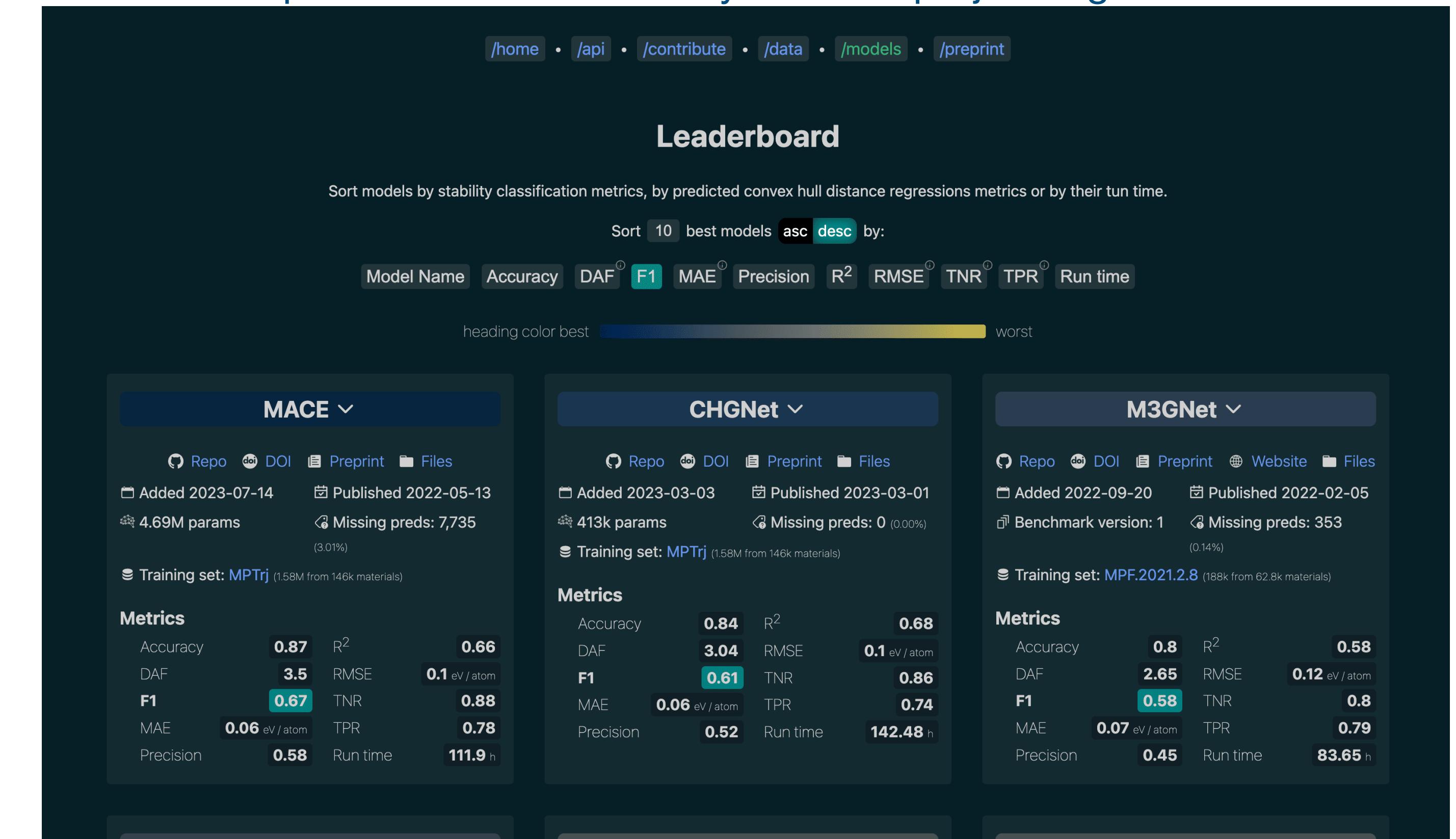
arXiv:2308.14920 (cond-mat)

[Submitted on 28 Aug 2023]

Matbench Discovery -- An evaluation framework for machine learning crystal stability prediction

Janosh Riebesell, Rhys E. A. Goodall,
Anubhav Jain, Philipp Benner, Kristin A.
Persson, Alpha A. Lee

<https://matbench-discovery.materialsproject.org/models>



Dielectric Materials Discovery

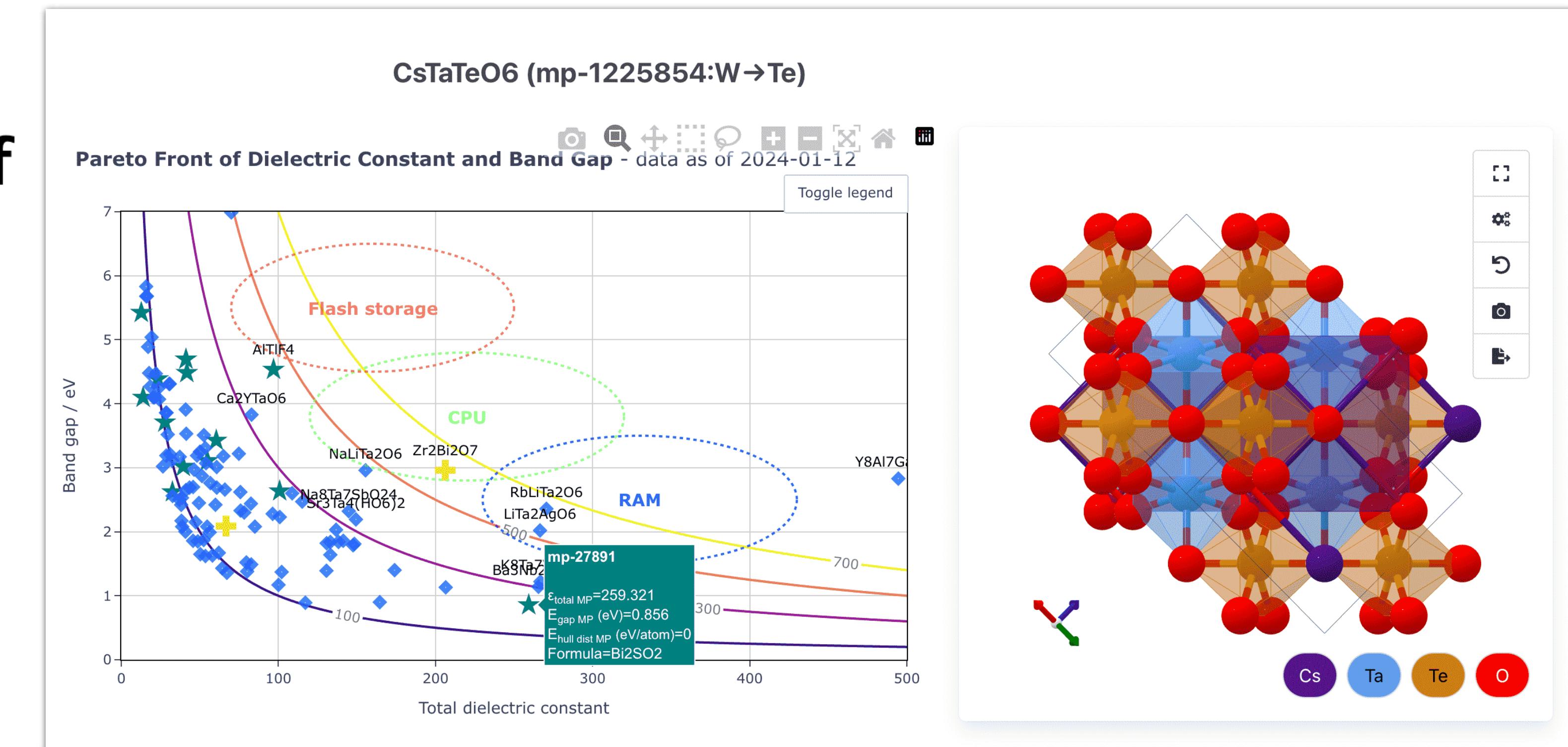
Can we deploy ML-guided discovery into practice?

arXiv:2401.05848 (cond-mat)

[Submitted on 11 Jan 2024]

Pushing the Pareto front of band gap and permittivity: ML-guided search for dielectric materials

Janosh Riebesell, T. Wesley Surta, Rhys Goodall, Michael Gaultois, Alpha A Lee



MACE-MP

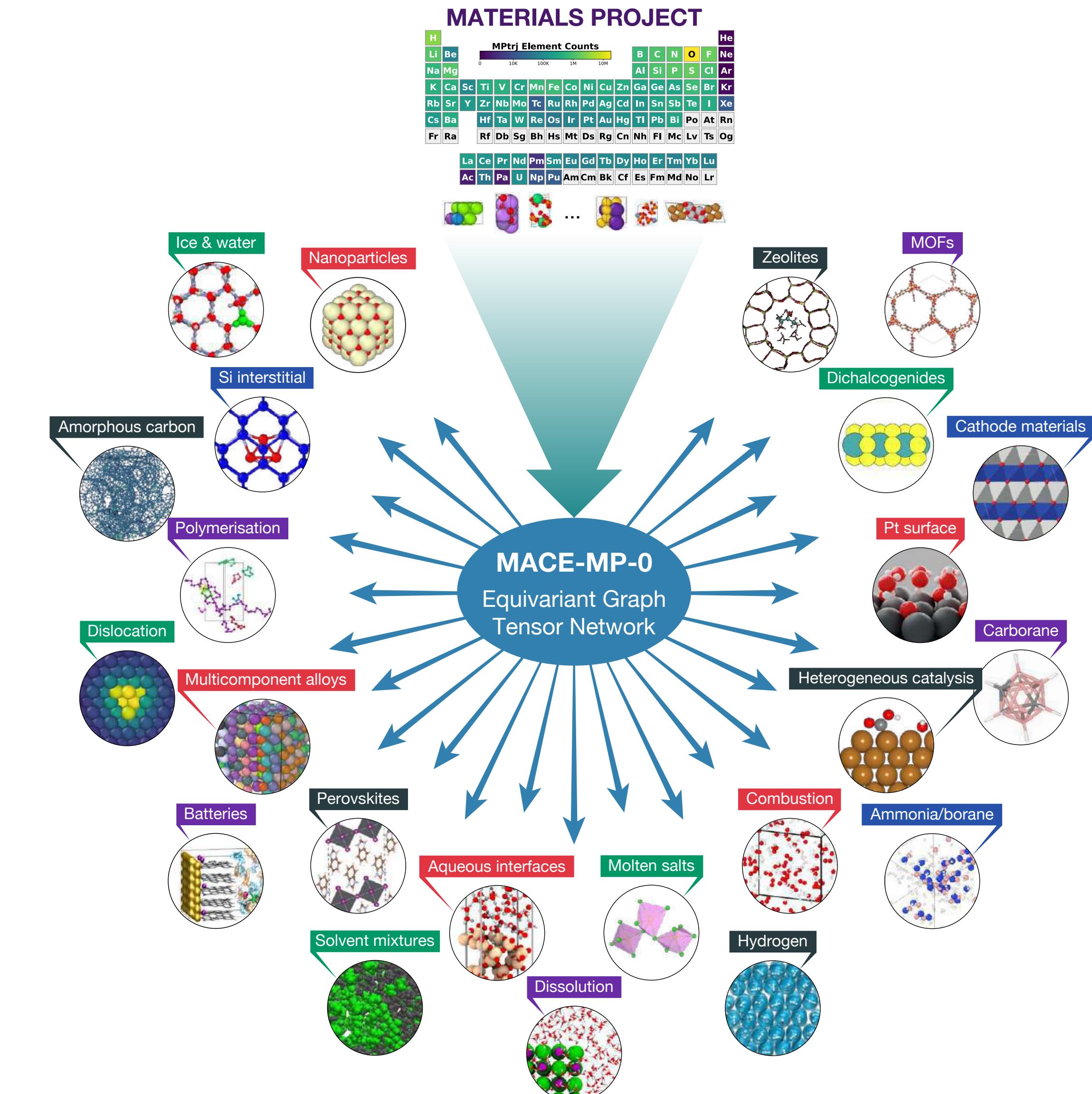
What else can foundation models trained for discovery do?

arXiv:2401.00096 (physics)

[Submitted on 29 Dec 2023]

A foundation model for atomistic materials chemistry

Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M. Elena, Dávid P. Kovács, Janosh Riebesell, Xavier R. Advincula, Mark Asta, William J. Baldwin, Noam Bernstein, Arghya Bhowmik, Samuel M. Blau, Vlad Cărare, James P. Darby, Sandip De, Flaviano Della Pia, Volker L. Deringer, Rokas Elijošius, Zakariya El-Machachi, Edvin Fako, Andrea C. Ferrari, Annalena Genreith-Schriever, Janine George, Rhys E. A. Goodall, Clare P. Grey, Shuang Han, Will Handley, Hendrik H. Heenen, Kersti Hermansson, Christian Holm, Jad Jaafar, Stephan Hofmann, Konstantin S. Jakob, Hyunwook Jung, Venkat Kapil, Aaron D. Kaplan, Nima Karimitari, Namu Kroupa, Jolla Kullgren, Matthew C. Kuner, Domantas Kuryla, Guoda Liepuoniute, Johannes T. Margraf, Ioan-Bogdan Magdău, Angelos Michaelides, J. Harry Moore, Aakash A. Naik, Samuel P. Niblett, Sam Walton Norwood, Niamh O'Neill, Christoph Ortner, Kristin A. Persson, Karsten Reuter, Andrew S. Rosen, Lars L. Schaaf, Christoph Schran, Eric Sivonxay, Tamás K. Stenczel, Viktor Svahn, Christopher Sutton, Cas van der Oord, Eszter Varga-Umbrich, Tejs Vegge, Martin Vondrák, Yangshuai Wang, William C. Witt, Fabian Zills, Gábor Csányi



Why Matbench Discovery?

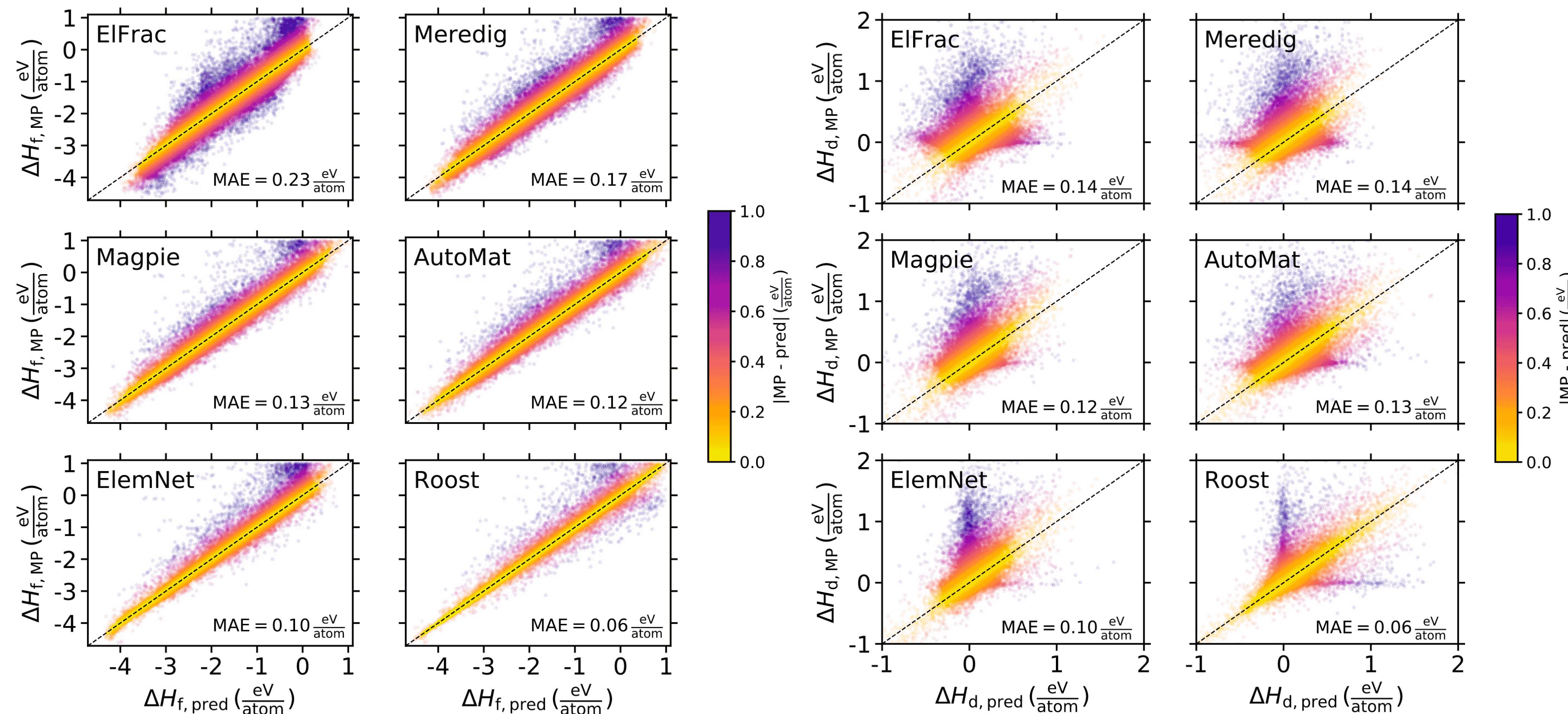
ARTICLE

OPEN

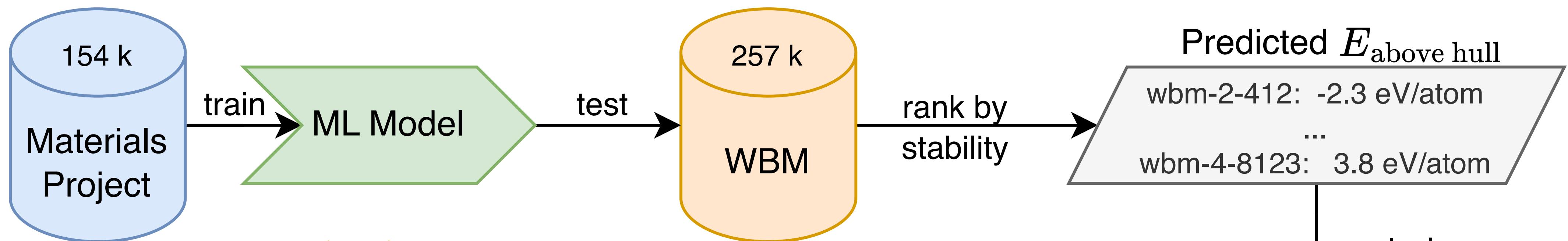


A critical examination of compound stability predictions from machine-learned formation energies

Christopher J. Bartel¹✉, Amalie Trewartha¹, Qi Wang², Alexander Dunn^{1,2}, Anubhav Jain² and Gerbrand Ceder^{1,3}✉

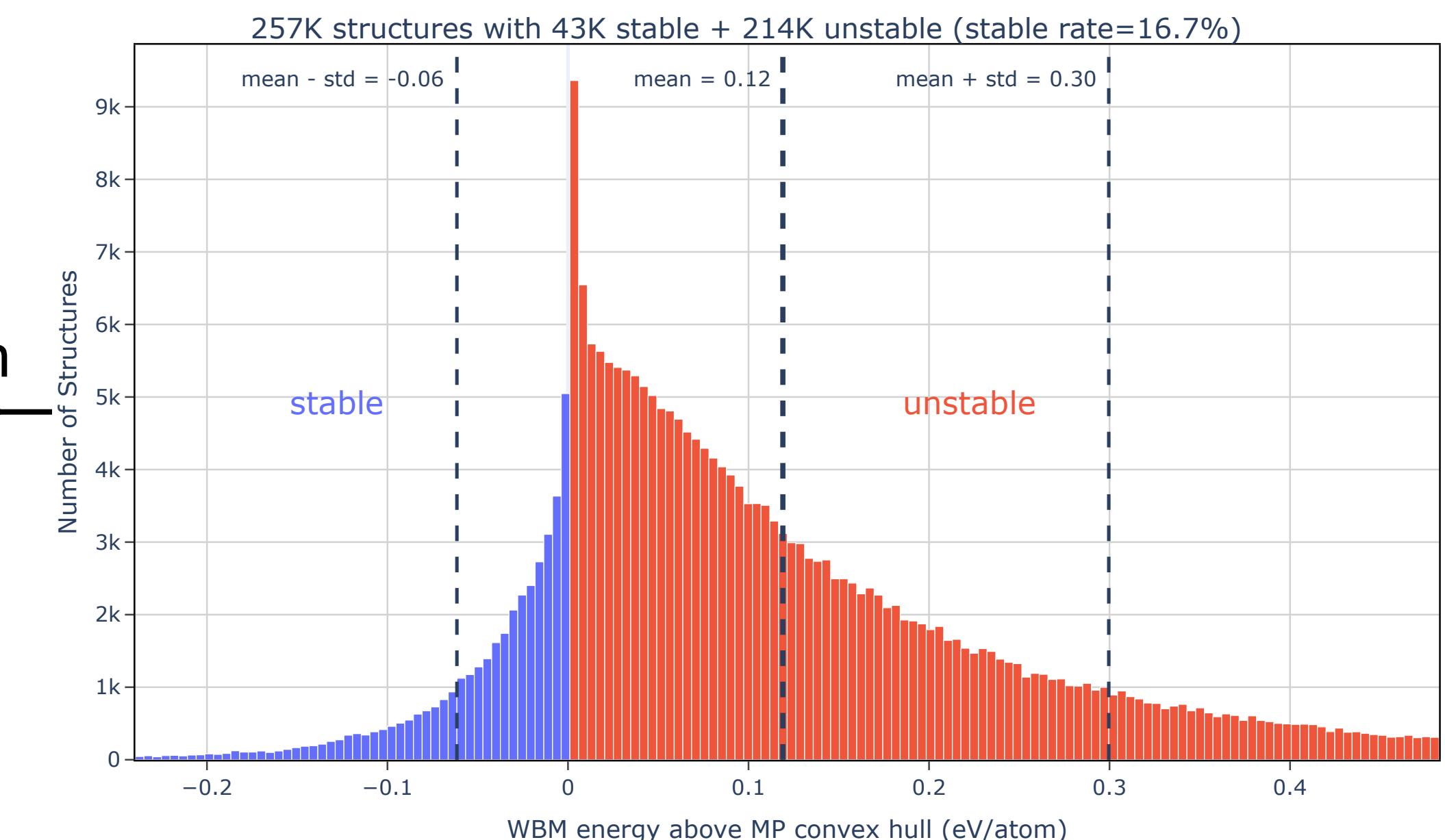


MP = Training Set WBM = Test Set



Leaderboard

Model	F1 ↑	DAF ↑	Prec ↑	Acc ↑	TPR ↑	TNR ↑	MAE ↓	RMSE ↓	R ² ↑
MACE	0.67	3.50	0.58	0.87	0.78	0.88	0.06	0.10	0.66
CHGNet	0.61	3.04	0.52	0.84	0.74	0.86	0.06	0.10	0.69
M3GNet	0.58	2.65	0.45	0.80	0.79	0.80	0.07	0.12	0.58
ALIGNN	0.56	2.92	0.50	0.83	0.65	0.87	0.09	0.15	0.27
MEGNet	0.51	2.70	0.46	0.81	0.57	0.86	0.13	0.20	-0.28
CGCNN	0.51	2.63	0.45	0.81	0.59	0.85	0.14	0.23	-0.62
CGCNN+P	0.51	2.40	0.41	0.78	0.67	0.80	0.11	0.18	0.03
Wrenformer	0.48	2.13	0.36	0.74	0.69	0.75	0.10	0.18	-0.04
BOWSR	0.44	1.91	0.32	0.68	0.74	0.67	0.12	0.16	0.14
Voronoi RF	0.34	1.51	0.26	0.67	0.51	0.70	0.14	0.21	-0.31
Dummy	0.19	1.00	0.17	0.68	0.23	0.77	0.12	0.18	0.00



Stability criterion:
 $E_{\text{above MP hull}} \leq 0$

MP = Training Set WBM = Test Set

npj Computational Materials

www.nature.com

ARTICLE OPEN
Predicting stable crystalline compounds using chemical similarity

Hai-Chen Wang¹, Silvana Botti¹ and Miguel A. L. Marques¹

WBM dataset = generated 256k MP-compatible crystals via element substitution

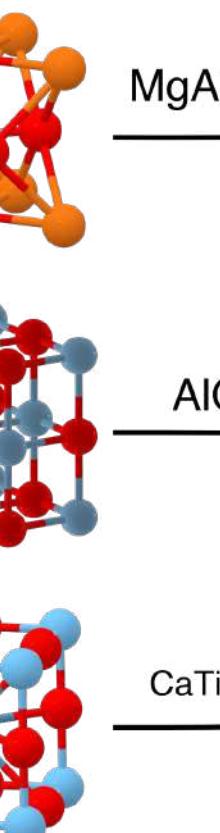
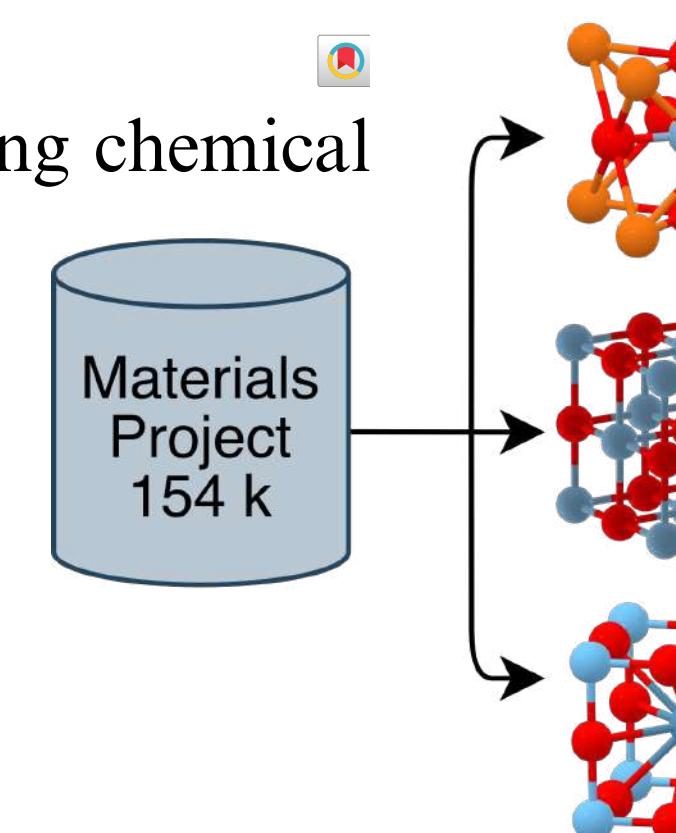
154 k
Materials Project



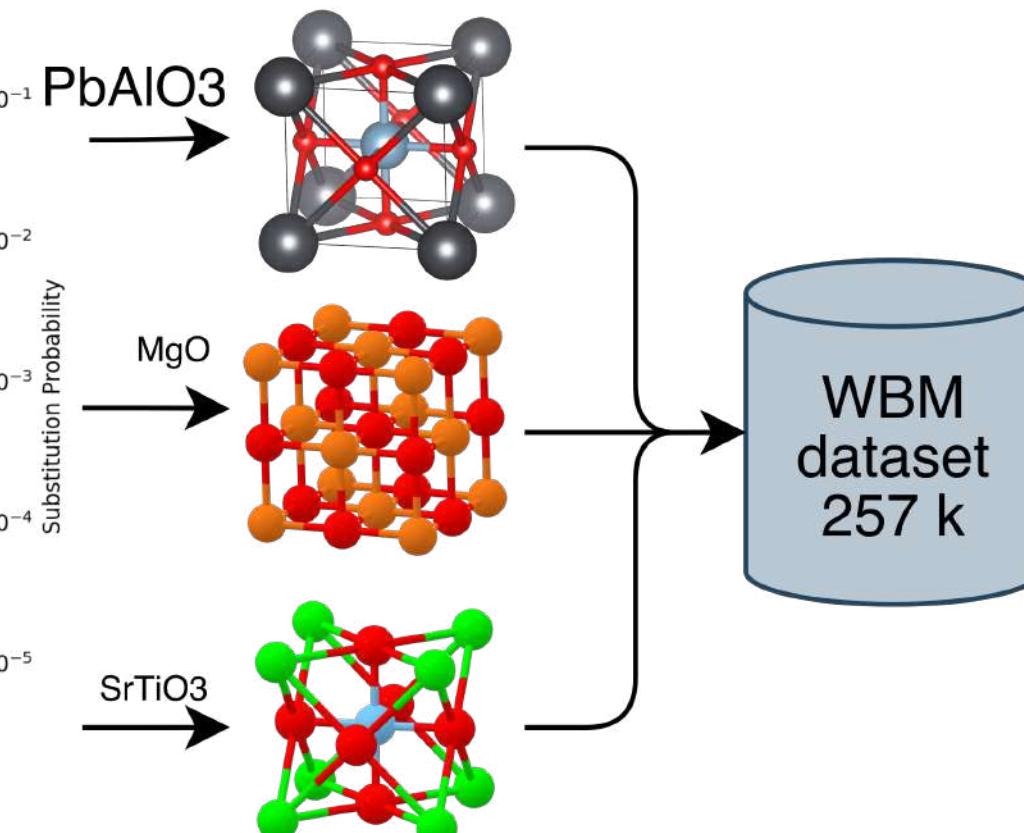
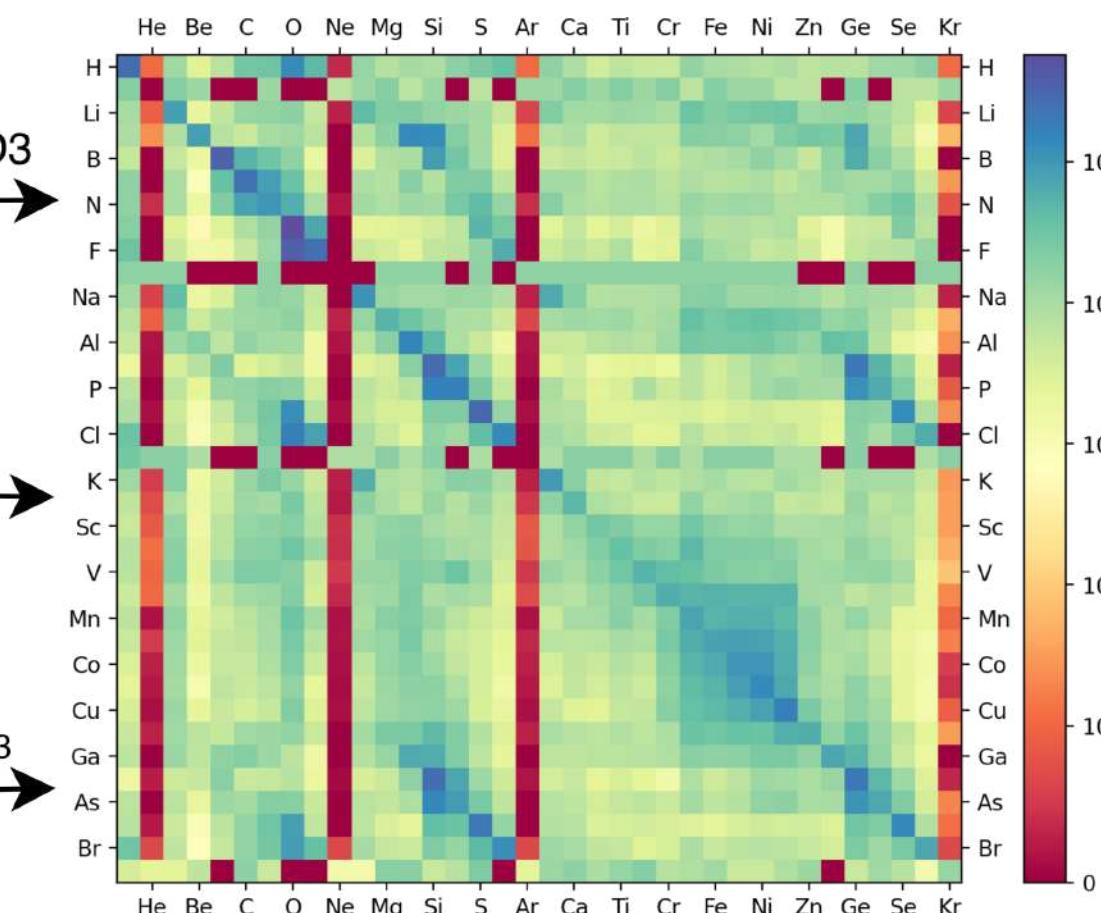
Leaderboard

Model	F1 ↑	DAF ↑	Prec ↑	Acc ↑	TPR ↑	TNR ↑	MAE ↓	RMSE ↓	R ² ↑
MACE	0.67	3.50	0.58	0.87	0.78	0.88	0.06	0.10	0.66
CHGNet	0.61	3.04	0.52	0.84	0.74	0.86	0.06	0.10	0.69
M3GNet	0.58	2.65	0.45	0.80	0.79	0.80	0.07	0.12	0.58
ALIGNN	0.56	2.92	0.50	0.83	0.65	0.87	0.09	0.15	0.27
MEGNet	0.51	2.70	0.46	0.81	0.57	0.86	0.13	0.20	-0.28
CGCNN	0.51	2.63	0.45	0.81	0.59	0.85	0.14	0.23	-0.62
CGCNN+P	0.51	2.40	0.41	0.78	0.67	0.80	0.11	0.18	0.03
Wrenformer	0.48	2.13	0.36	0.74	0.69	0.75	0.10	0.18	-0.04
BOWSR	0.44	1.91	0.32	0.68	0.74	0.67	0.12	0.16	0.14
Voronoi RF	0.34	1.51	0.26	0.67	0.51	0.70	0.14	0.21	-0.31
Dummy	0.19	1.00	0.17	0.68	0.23	0.77	0.12	0.18	0.00

www.nature.com



ICSD element similarity

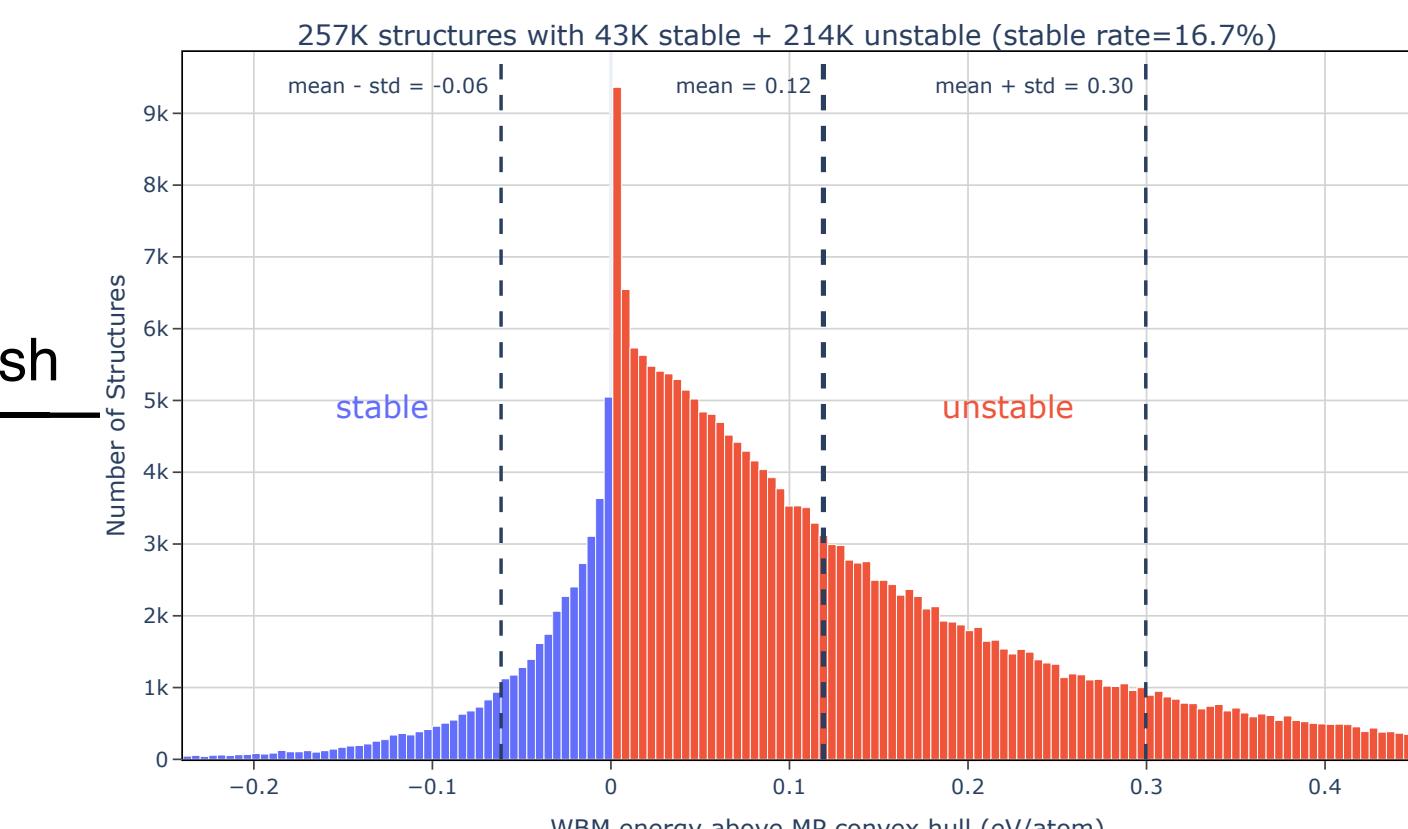


257 k
WBM

rank by stability

Predicted $E_{\text{above hull}}$
wbm-2-412: -2.3 eV/atom
...
wbm-4-8123: 3.8 eV/atom

analysis



merge
GitHub PR

push

10 models benchmarked to date

Model	F1 ↑	DAF ↑	Prec ↑	Acc ↑	TPR ↑	TNR ↑	MAE ↓	RMSE ↓	R² ↑	Training Size	Model Type
MACE	0.67	3.50	0.58	0.87	0.78	0.88	0.06	0.10	0.66	1.6M (145.9K)	UIP-GNN
CHGNet	0.61	3.04	0.52	0.84	0.74	0.86	0.06	0.10	0.69	1.6M (145.9K)	UIP-GNN
M3GNet	0.58	2.65	0.45	0.80	0.79	0.80	0.07	0.12	0.58	188.3K (62.8K)	UIP-GNN
ALIGNN	0.56	2.92	0.50	0.83	0.65	0.87	0.09	0.15	0.27	154.7K	GNN
MEGNet	0.51	2.70	0.46	0.81	0.57	0.86	0.13	0.20	-0.28	133.4K	GNN
CGCNN	0.51	2.63	0.45	0.81	0.59	0.85	0.14	0.23	-0.62	154.7K	GNN
CGCNN+P	0.51	2.40	0.41	0.78	0.67	0.80	0.11	0.18	0.03	154.7K	GNN
Wrenformer	0.48	2.13	0.36	0.74	0.69	0.75	0.10	0.18	-0.04	154.7K	Transformer
BOWSR	0.44	1.91	0.32	0.68	0.74	0.67	0.12	0.16	0.14	133.4K	BO-GNN
Voronoi RF	0.34	1.51	0.26	0.67	0.51	0.70	0.14	0.21	-0.31	154.7K	Fingerprint
Dummy	0.19	1.00	0.17	0.68	0.23	0.77	0.12	0.18	0.00		

UIP = universal interatomic potentials GNN = graph neural network RF = random forest

DAF = discovery acceleration factor BO = Bayesian optimization

WIP submissions for: Nequip, Allegro, PFP from Matlantis, GNoMe from DeepMind

Main Take Aways

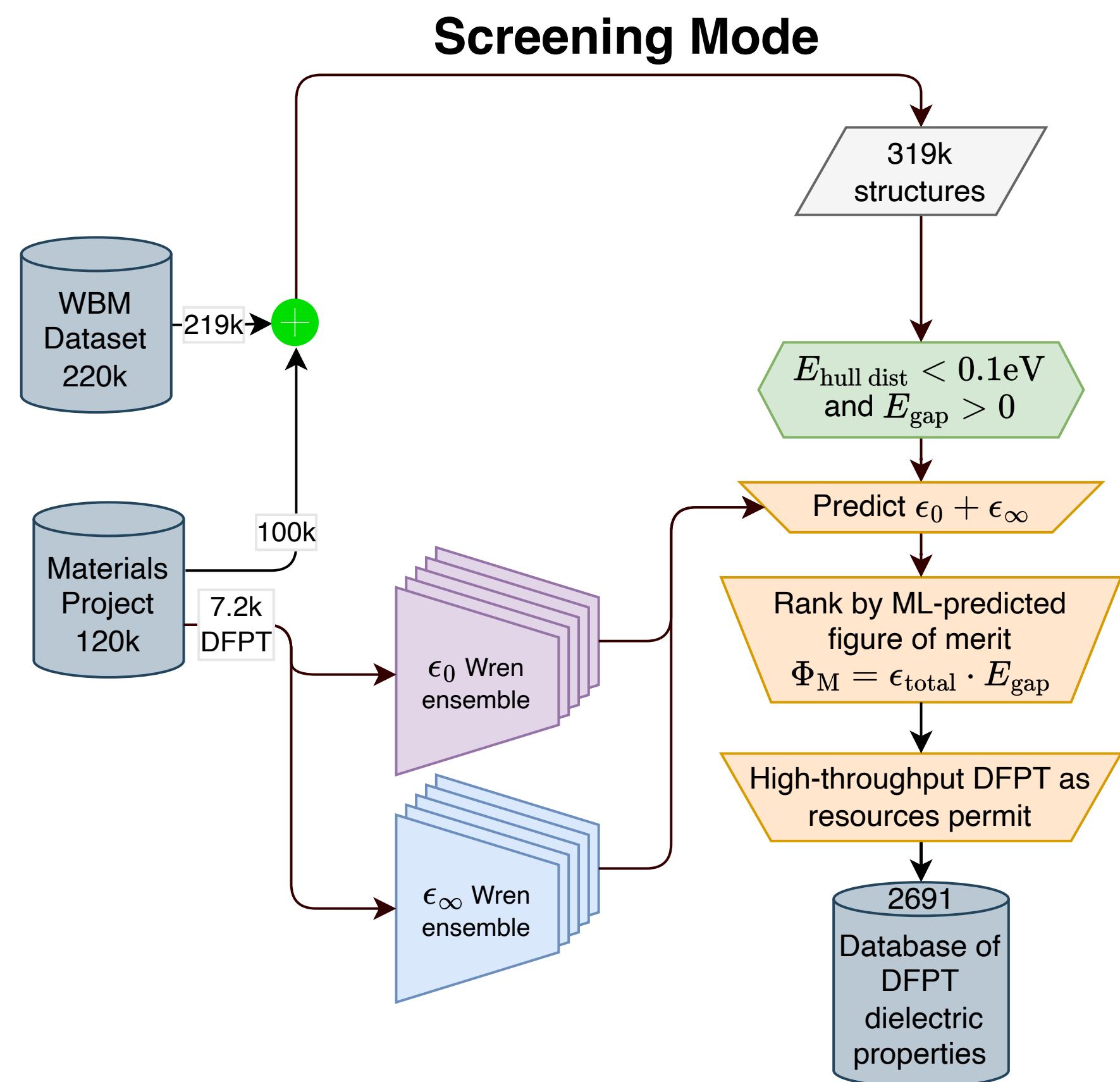
- ML force fields = winning ML methodology for materials discovery 
- ML stability predictions have improved a lot since Bartel et al. 2021!
- ML4Mat might be at similar inflection point as NLP after transformer paper
 - found the right architecture (MLIPs)
 - next: generate more data to build complete map of PES asap

Part 2

Pushing the Pareto front of band gap and dielectric constant

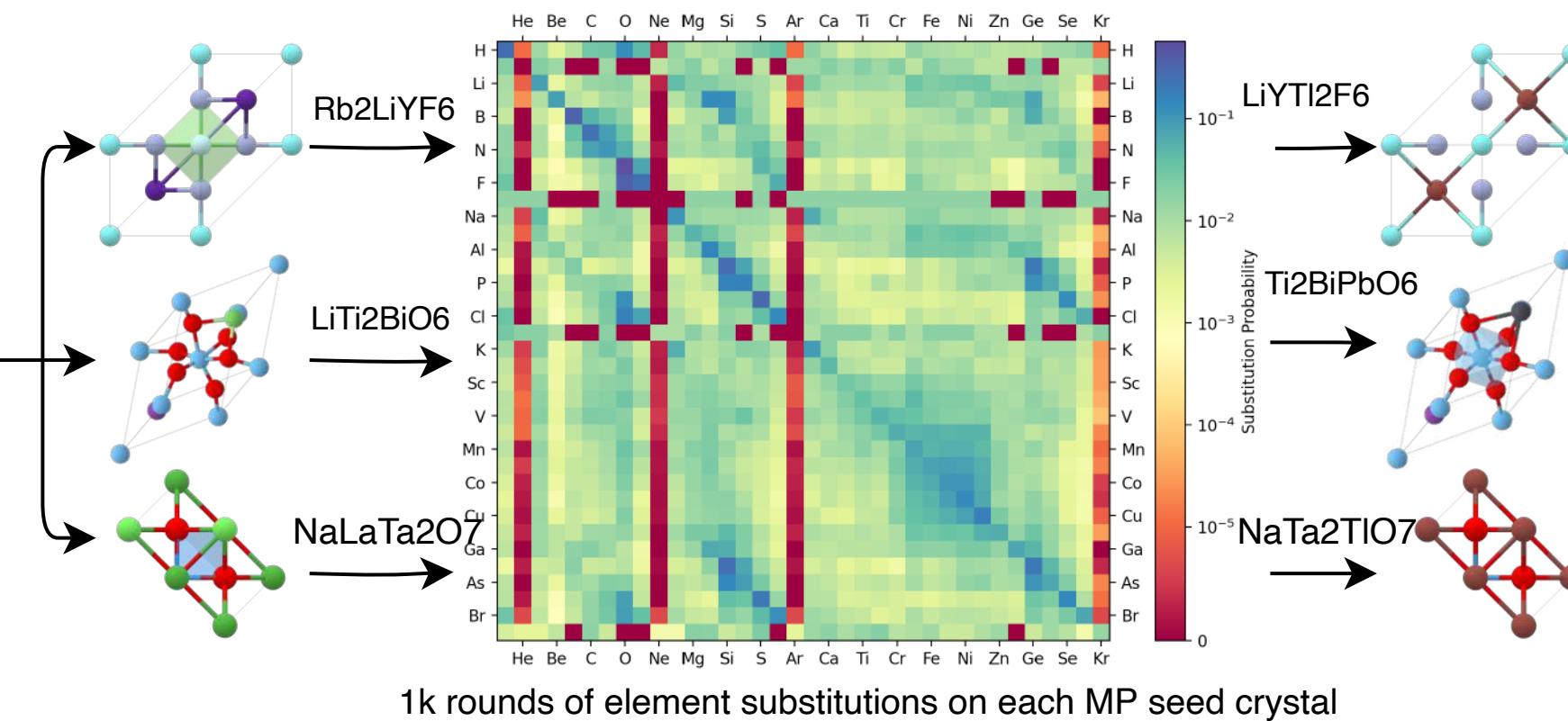
Does ML-guided discovery succeed in practice when applied to dielectrics discovery?

Discovery Pipeline

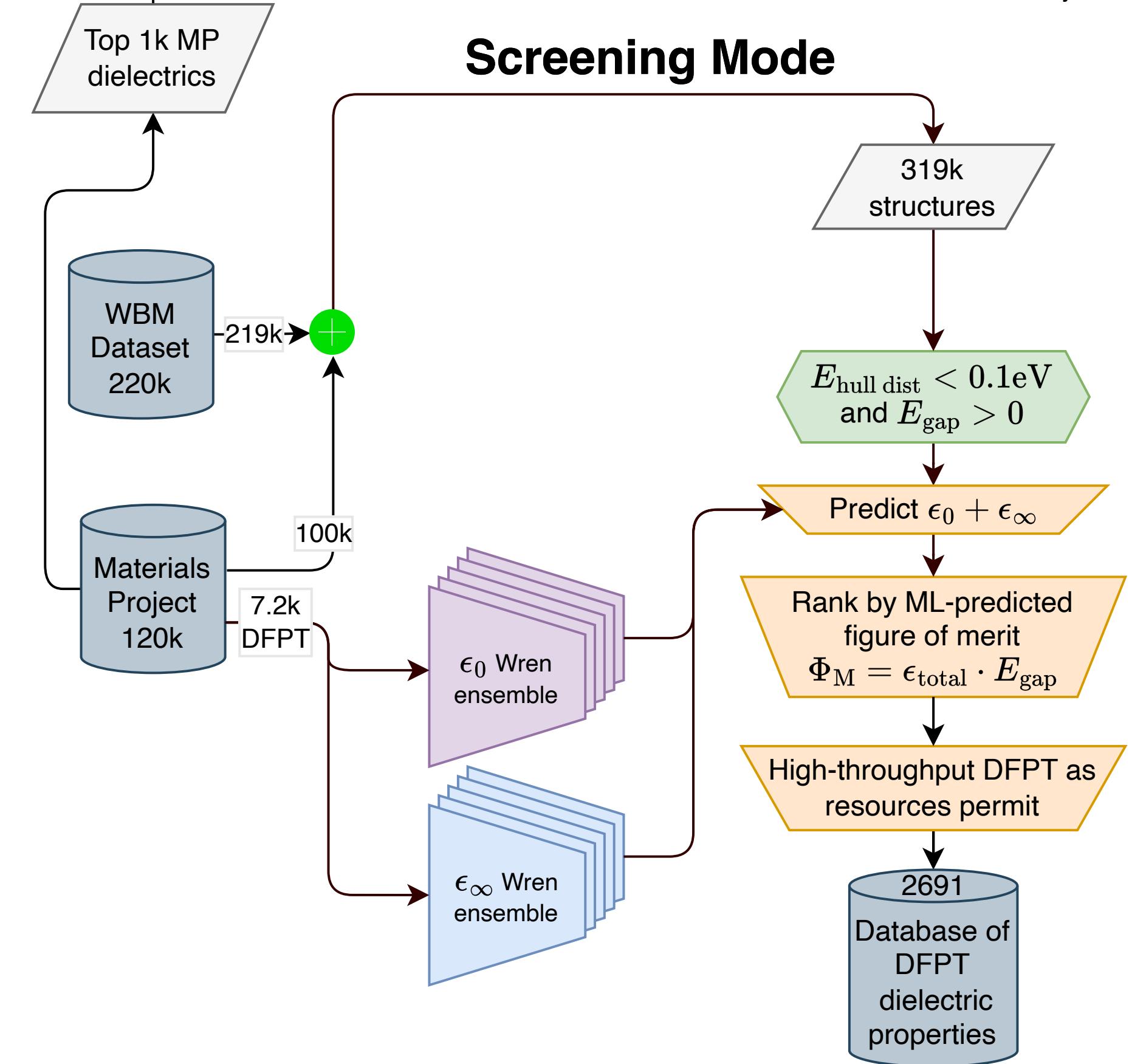


Discovery Pipeline

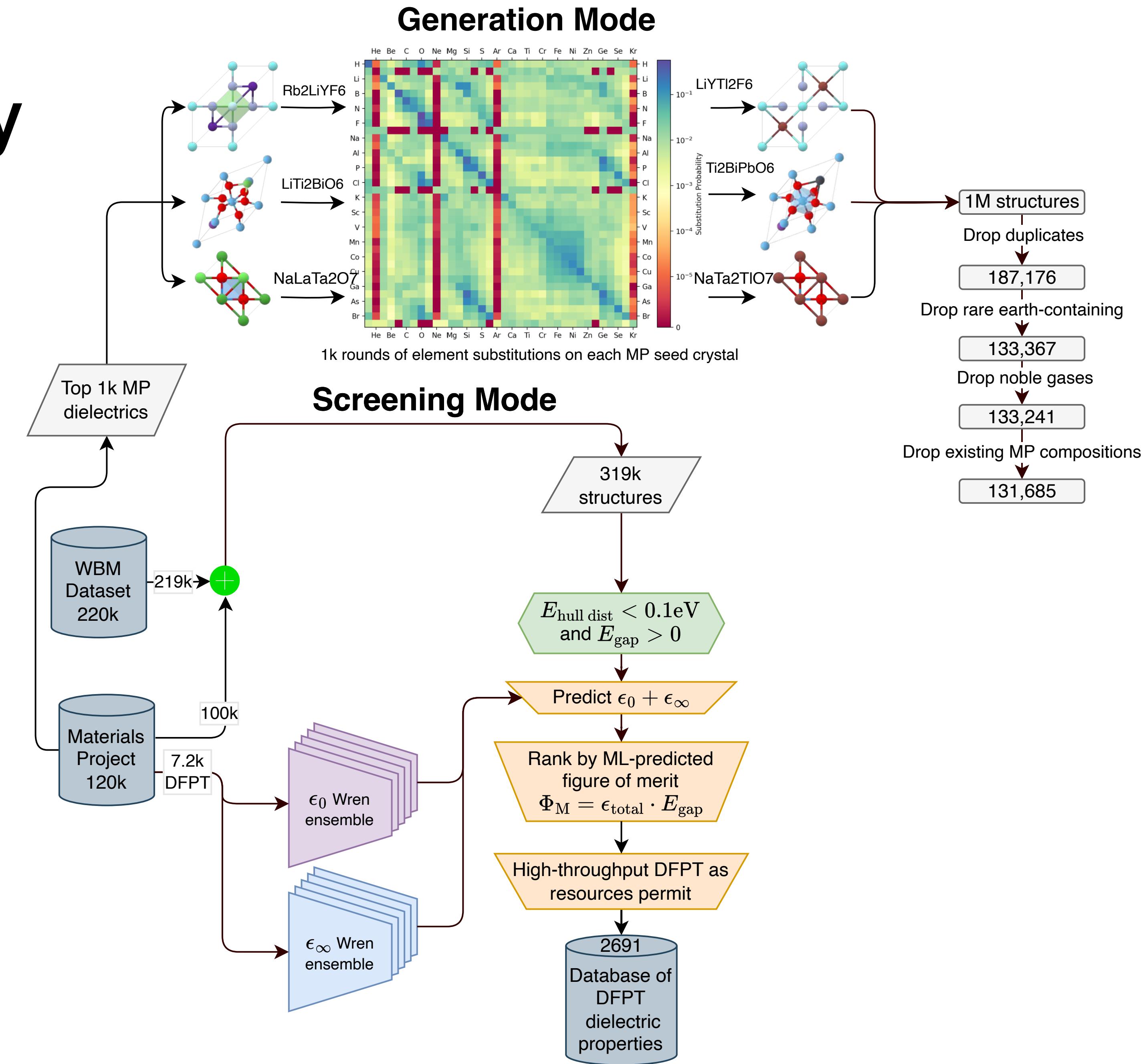
Generation Mode



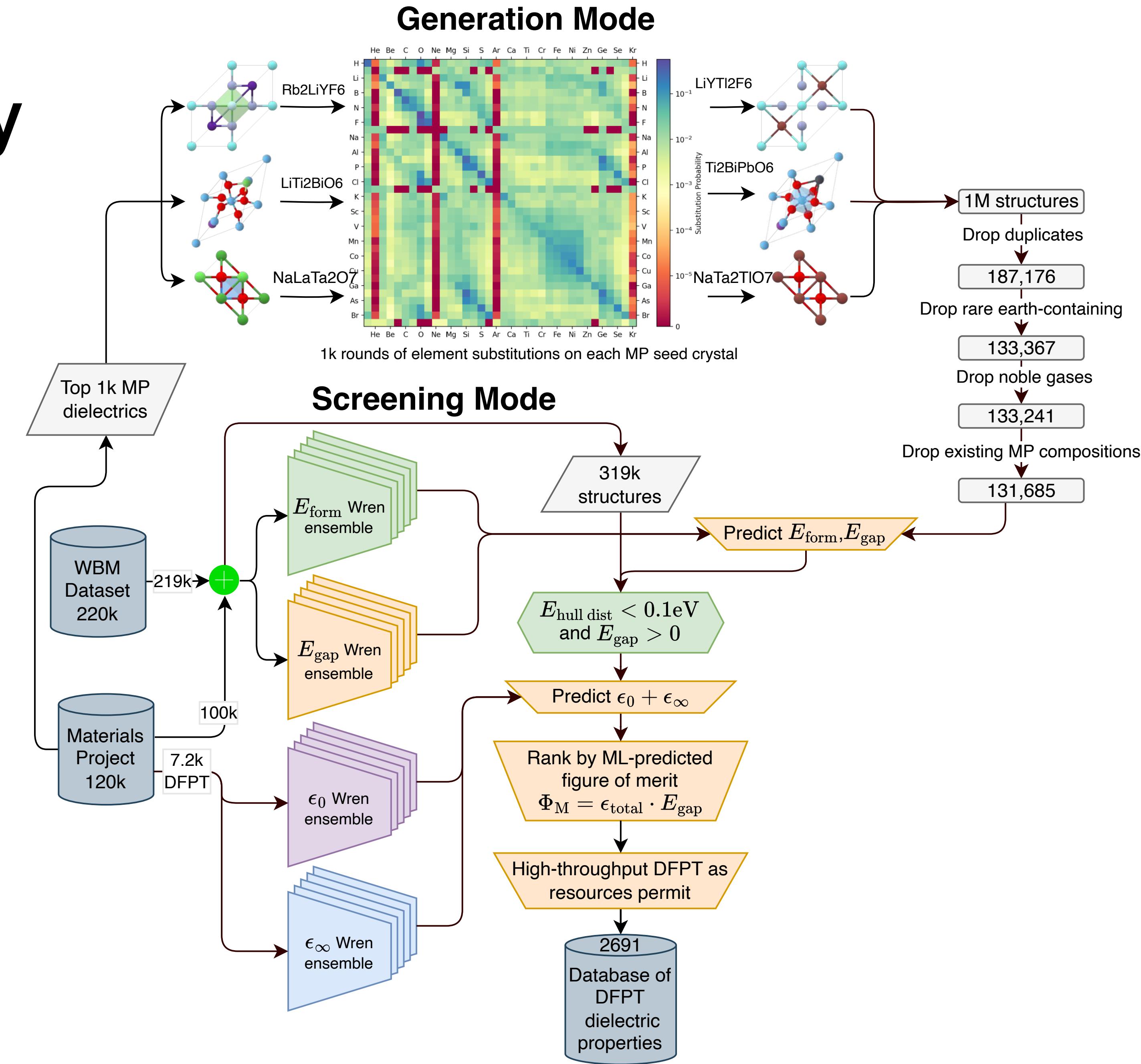
Screening Mode



Discovery Pipeline



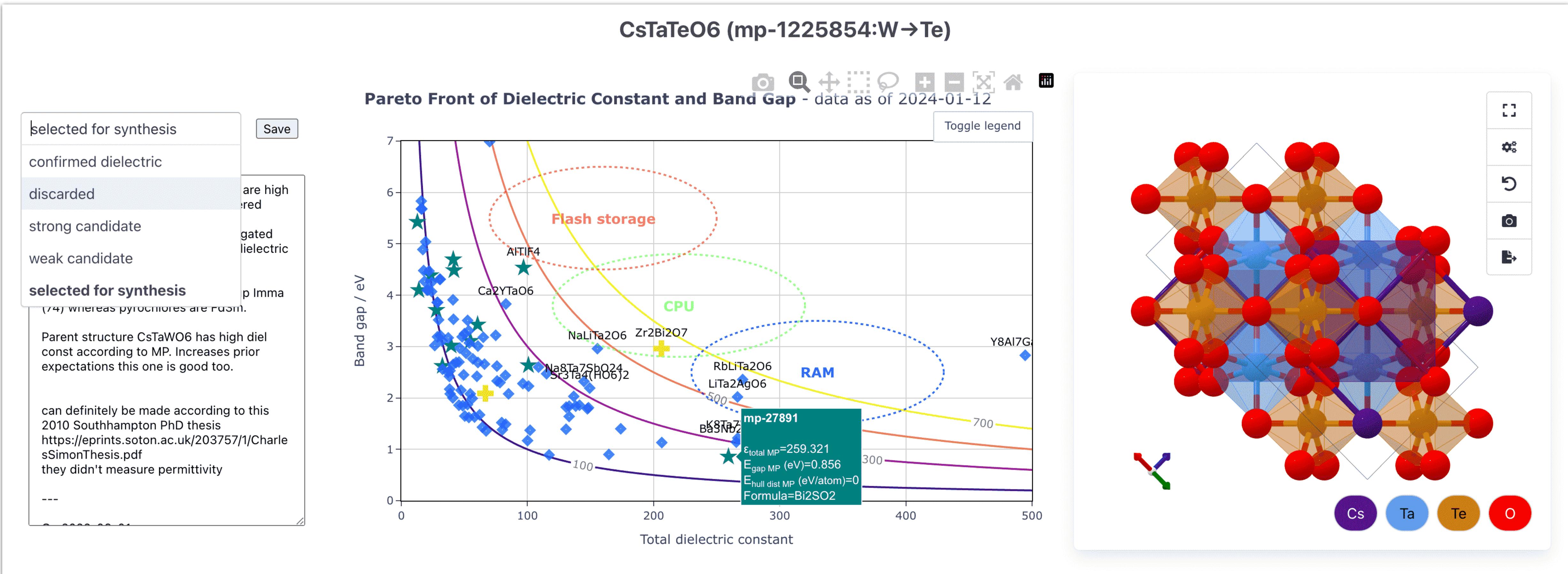
Discovery Pipeline



Custom web app for real-time collaborative synthesis selection

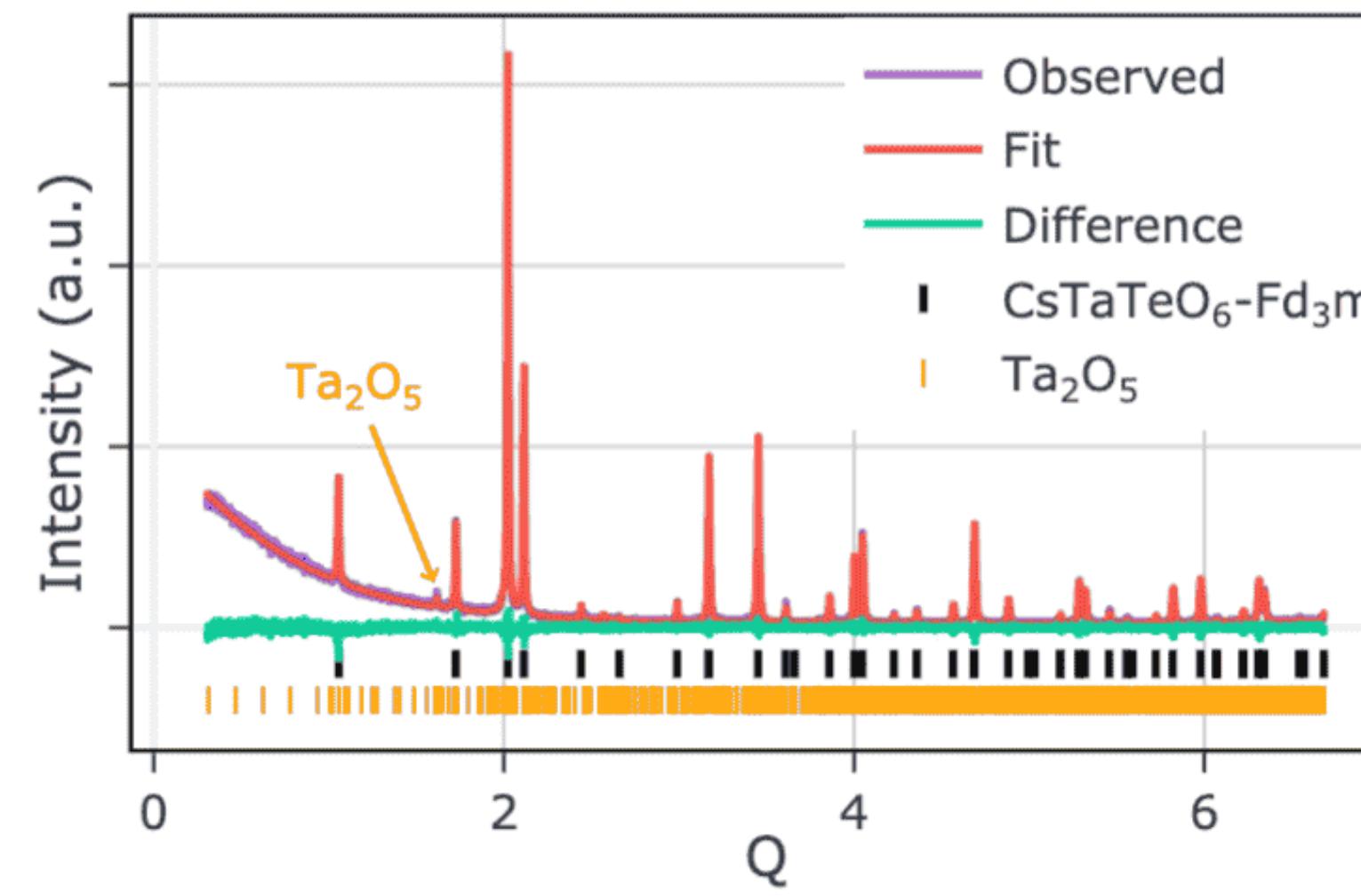
Dielectric constant ϵ_{total} and band gap E_{gap} = anti-correlated but both essential in electronic applications

We optimize figure of merit $\Phi_M = \epsilon_{\text{total}} \cdot E_{\text{gap}}$

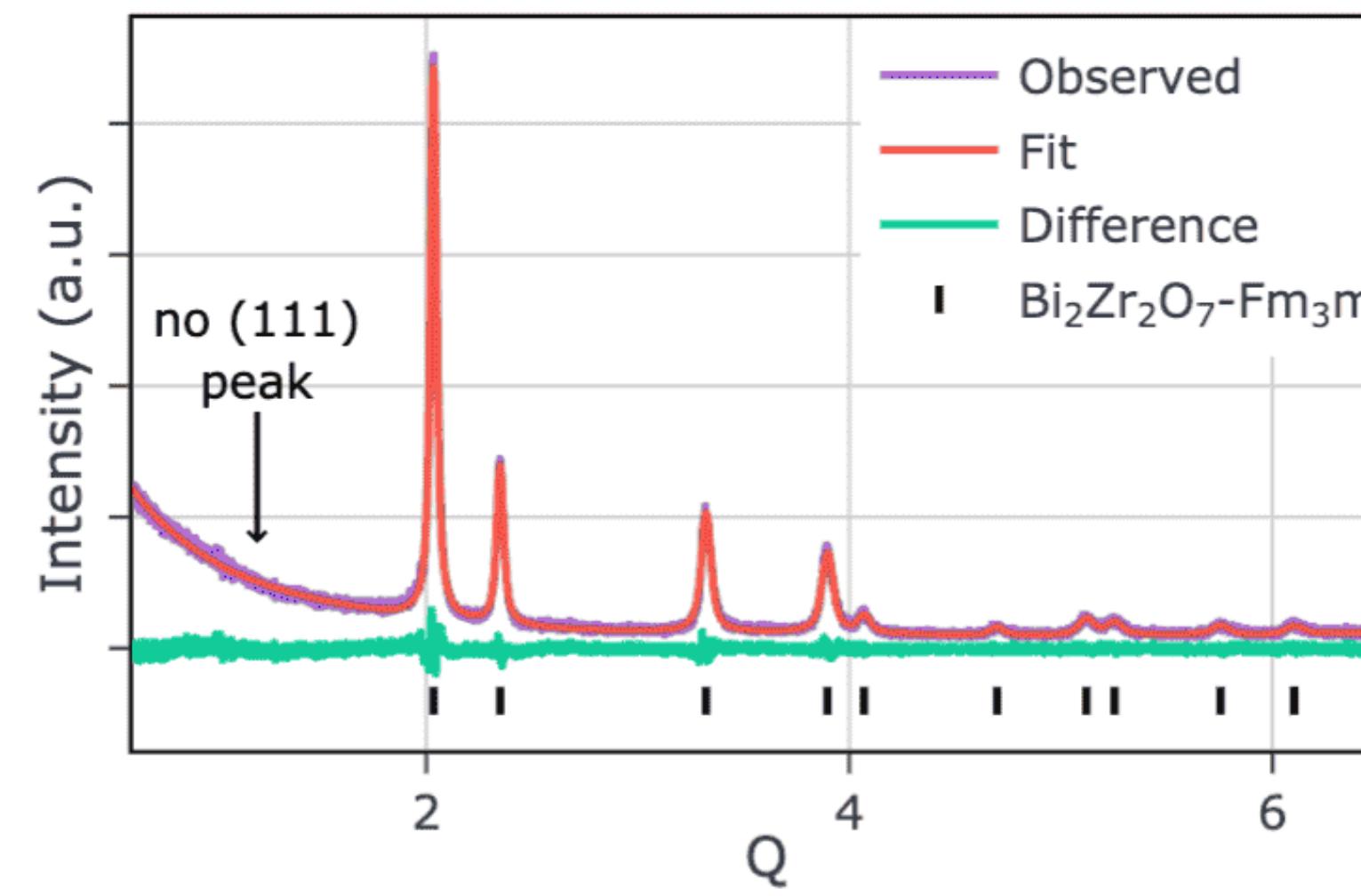
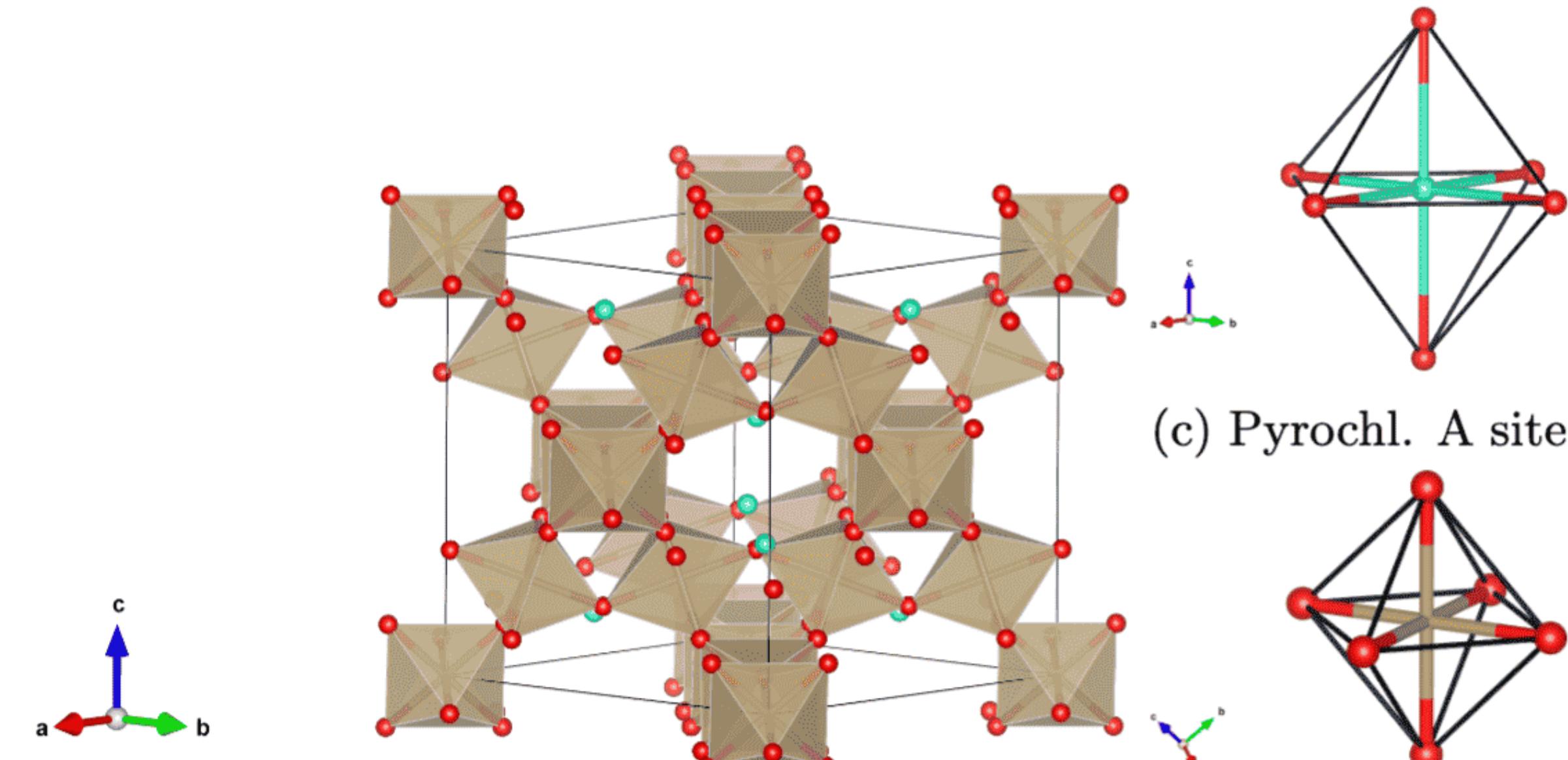


<https://janosh.github.io/dielectrics>

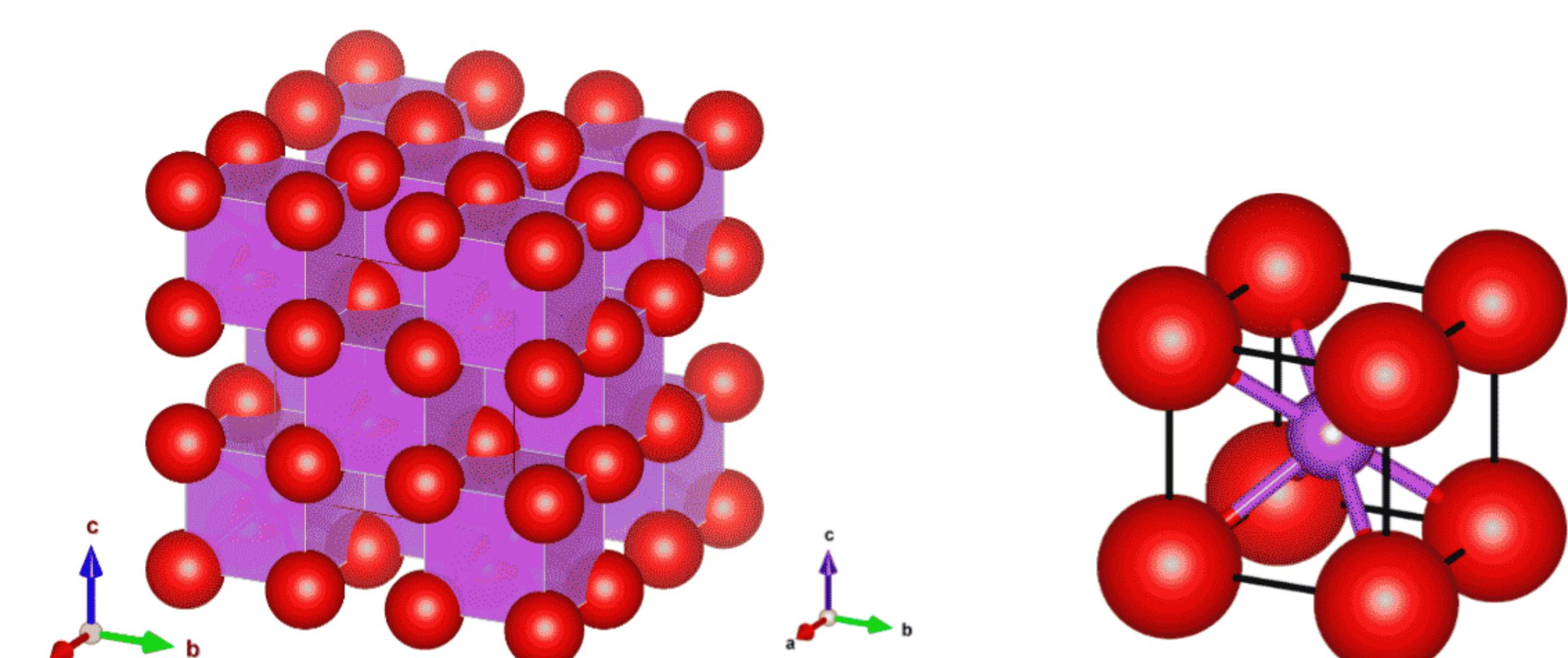
XRD Fits of $\text{Bi}_2\text{Zr}_2\text{O}_7$ and CsTaTeO_6



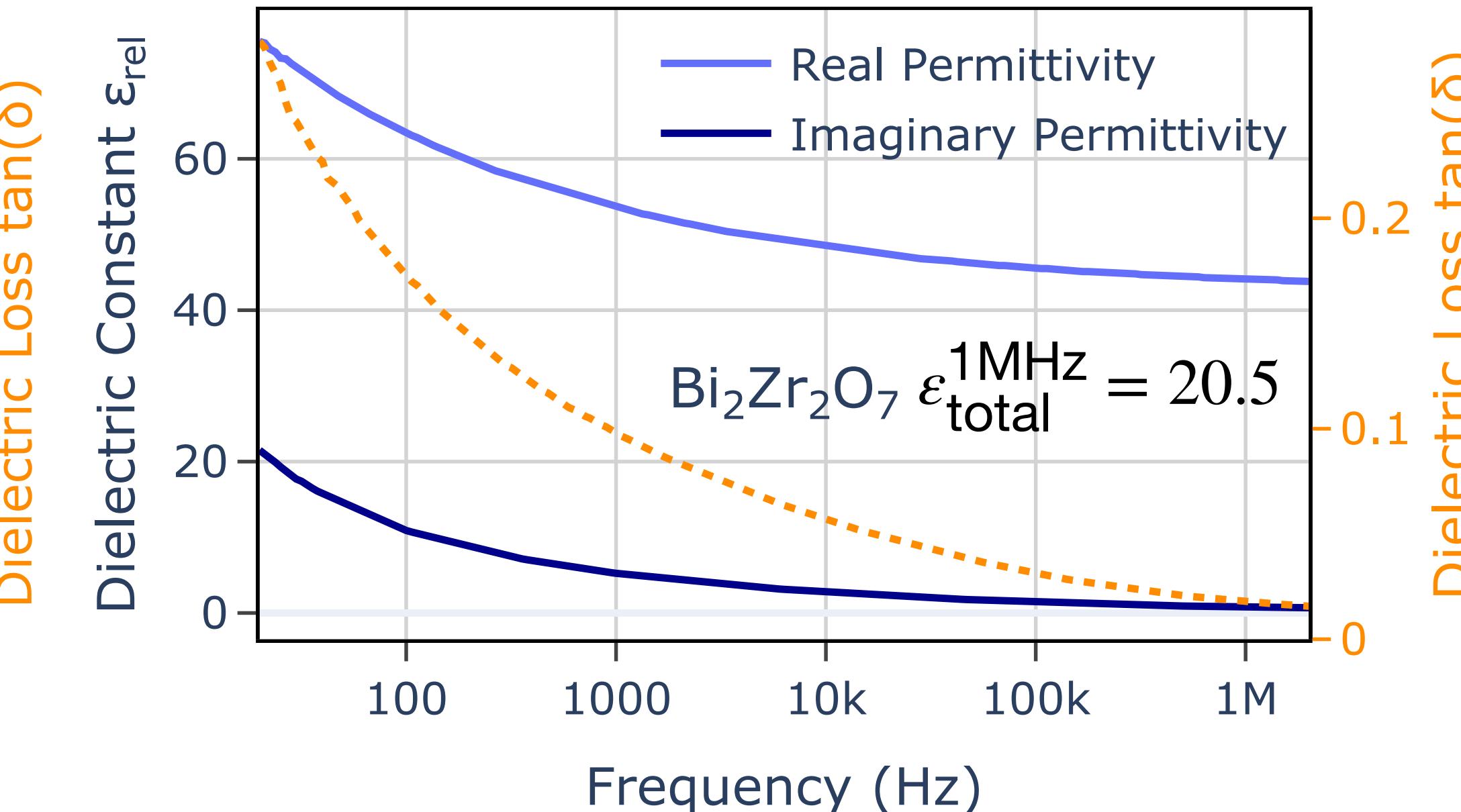
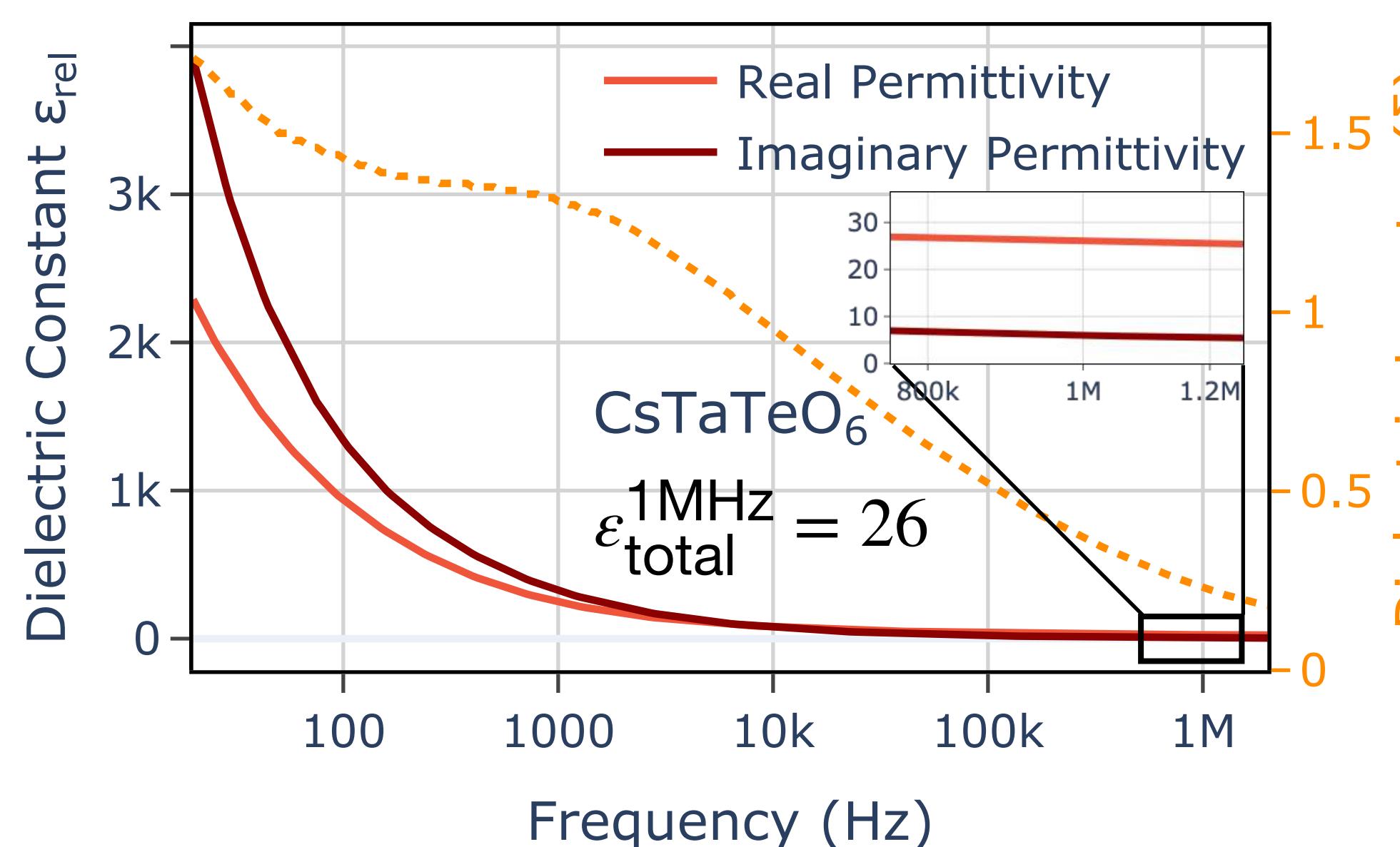
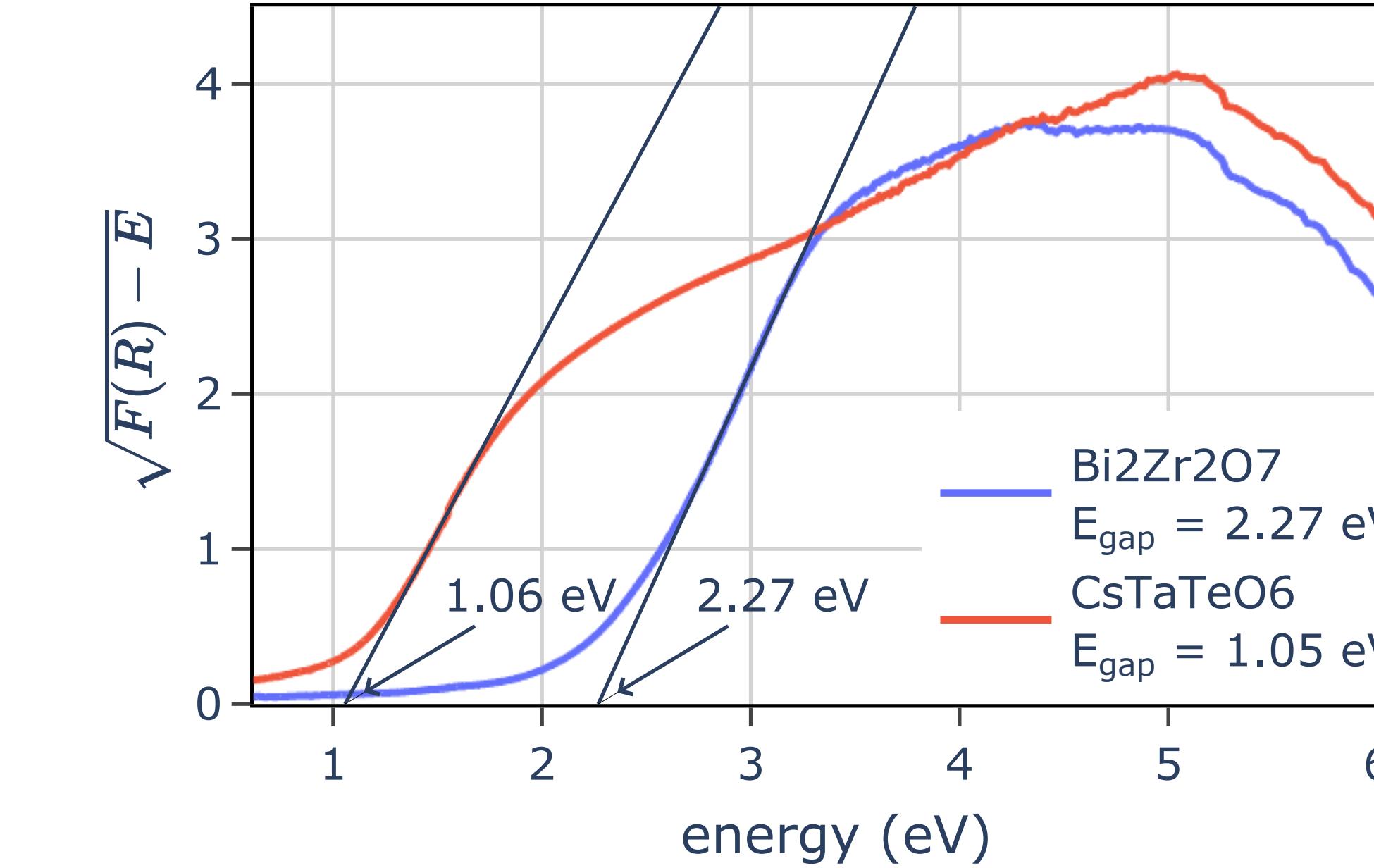
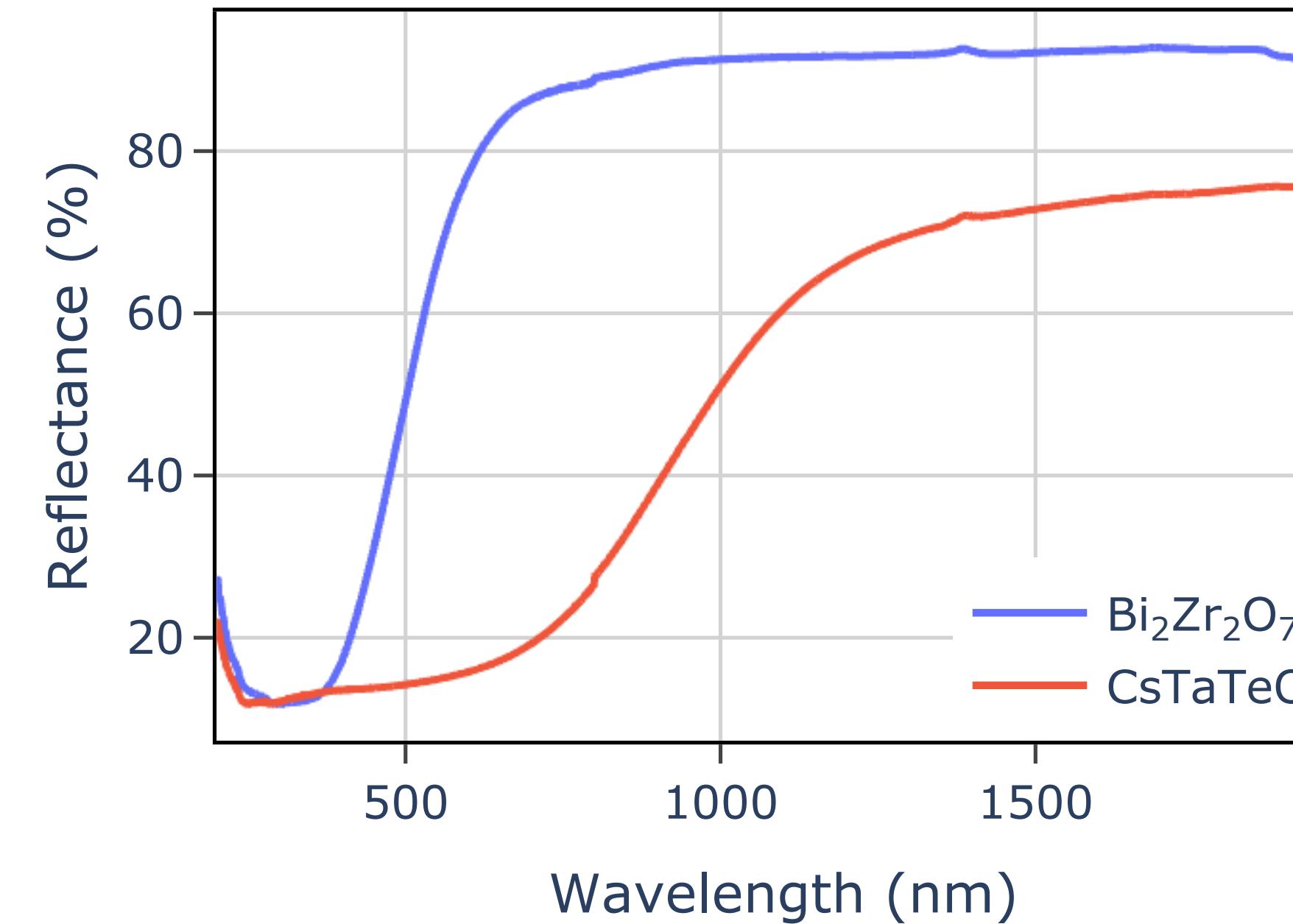
(a) Pyrochlore Fd3m Rietveld fit for CsTaTeO_6



(e) Fluorite Fm3m Rietveld fit for $\text{Bi}_2\text{Zr}_2\text{O}_7$



Dielectric Characterization of $\text{Bi}_2\text{Zr}_2\text{O}_7$ and CsTaTeO_6



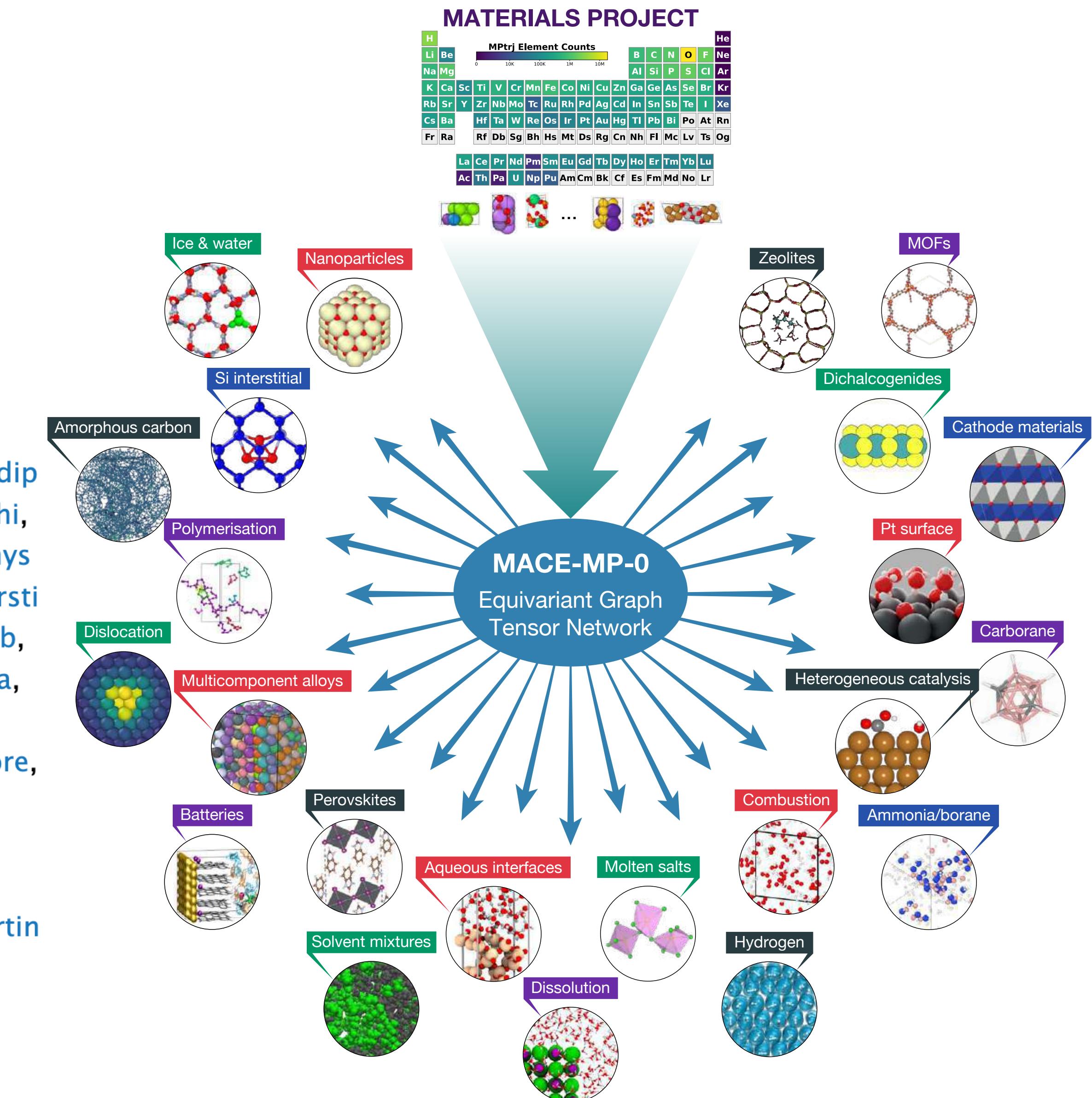
MACE-MP-0 Foundation Model

arXiv:2401.00096 (physics)

[Submitted on 29 Dec 2023]

A foundation model for atomistic materials chemistry

Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M. Elena, Dávid P. Kovács, Janosh Riebesell, Xavier R. Advincula, Mark Asta, William J. Baldwin, Noam Bernstein, Arghya Bhowmik, Samuel M. Blau, Vlad Cărare, James P. Darby, Sandip De, Flaviano Della Pia, Volker L. Deringer, Rokas Elijošius, Zakariya El-Machachi, Edvin Fako, Andrea C. Ferrari, Annalena Genreith-Schriever, Janine George, Rhys E. A. Goodall, Clare P. Grey, Shuang Han, Will Handley, Hendrik H. Heenen, Kersti Hermansson, Christian Holm, Jad Jaafar, Stephan Hofmann, Konstantin S. Jakob, Hyunwook Jung, Venkat Kapil, Aaron D. Kaplan, Nima Karimitari, Namu Kroupa, Jolla Kullgren, Matthew C. Kuner, Domantas Kuryla, Guoda Liepuoniute, Johannes T. Margraf, Ioan-Bogdan Magdău, Angelos Michaelides, J. Harry Moore, Aakash A. Naik, Samuel P. Niblett, Sam Walton Norwood, Niamh O'Neill, Christoph Ortner, Kristin A. Persson, Karsten Reuter, Andrew S. Rosen, Lars L. Schaaf, Christoph Schran, Eric Sivonxay, Tamás K. Stenczel, Viktor Svahn, Christopher Sutton, Cas van der Oord, Eszter Varga-Umbrich, Tejs Vegge, Martin Vondrák, Yangshuai Wang, William C. Witt, Fabian Zills, Gábor Csányi



Part 3

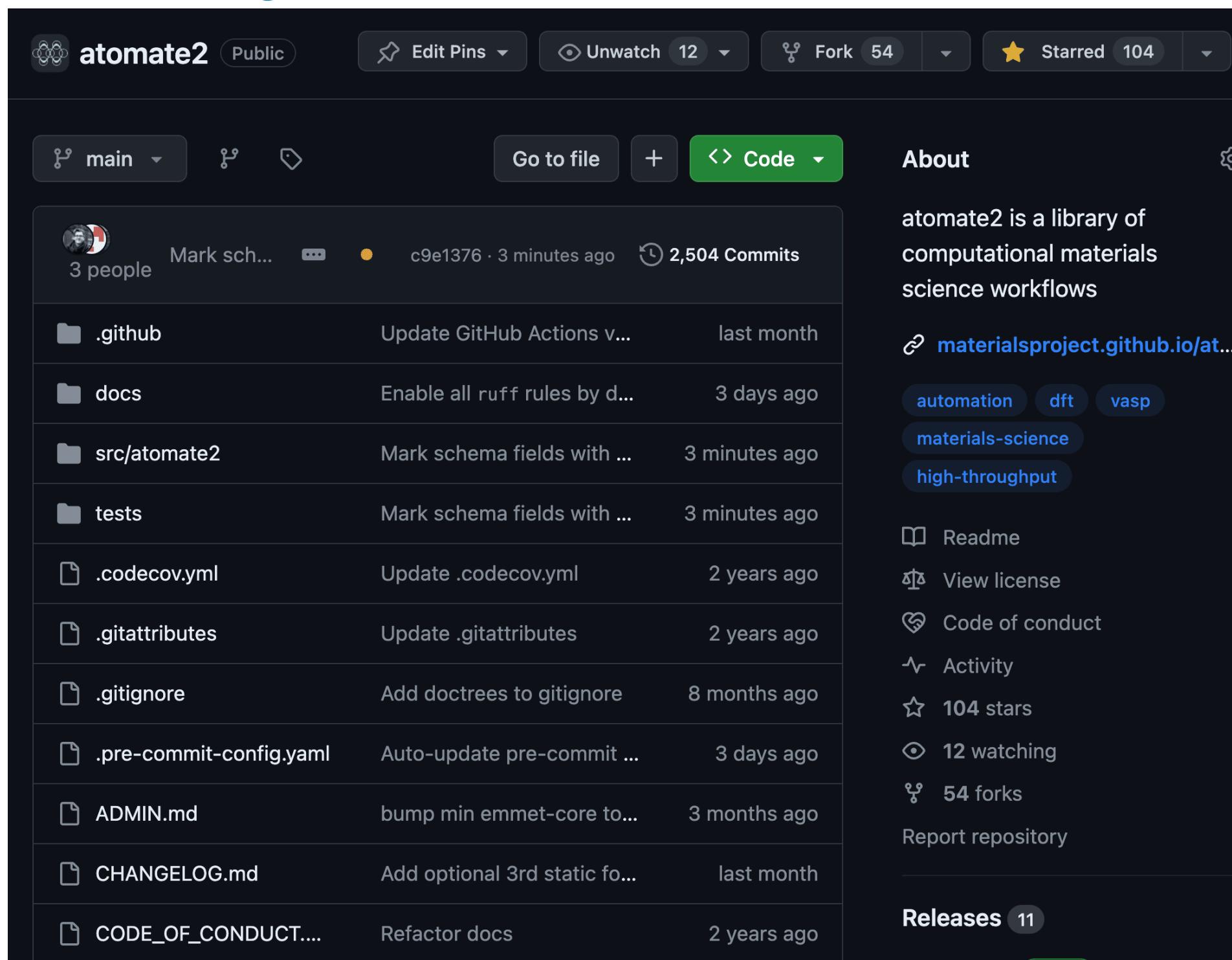
**Can ML potentials predict
DFT lattice vibrations (aka phonons)?**

Can ML potentials reproduce DFT lattice vibrations (aka phonons)?

Compare MACE/CHGNet with PhononDB PBE data using

 Atomate2 PhononMaker

<https://github.com/materialsproject/atomate2>



The screenshot shows the GitHub repository page for atomate2. The top navigation bar includes 'Edit Pins', 'Unwatch', 'Fork', and 'Starred'. The main content area displays a list of recent commits:

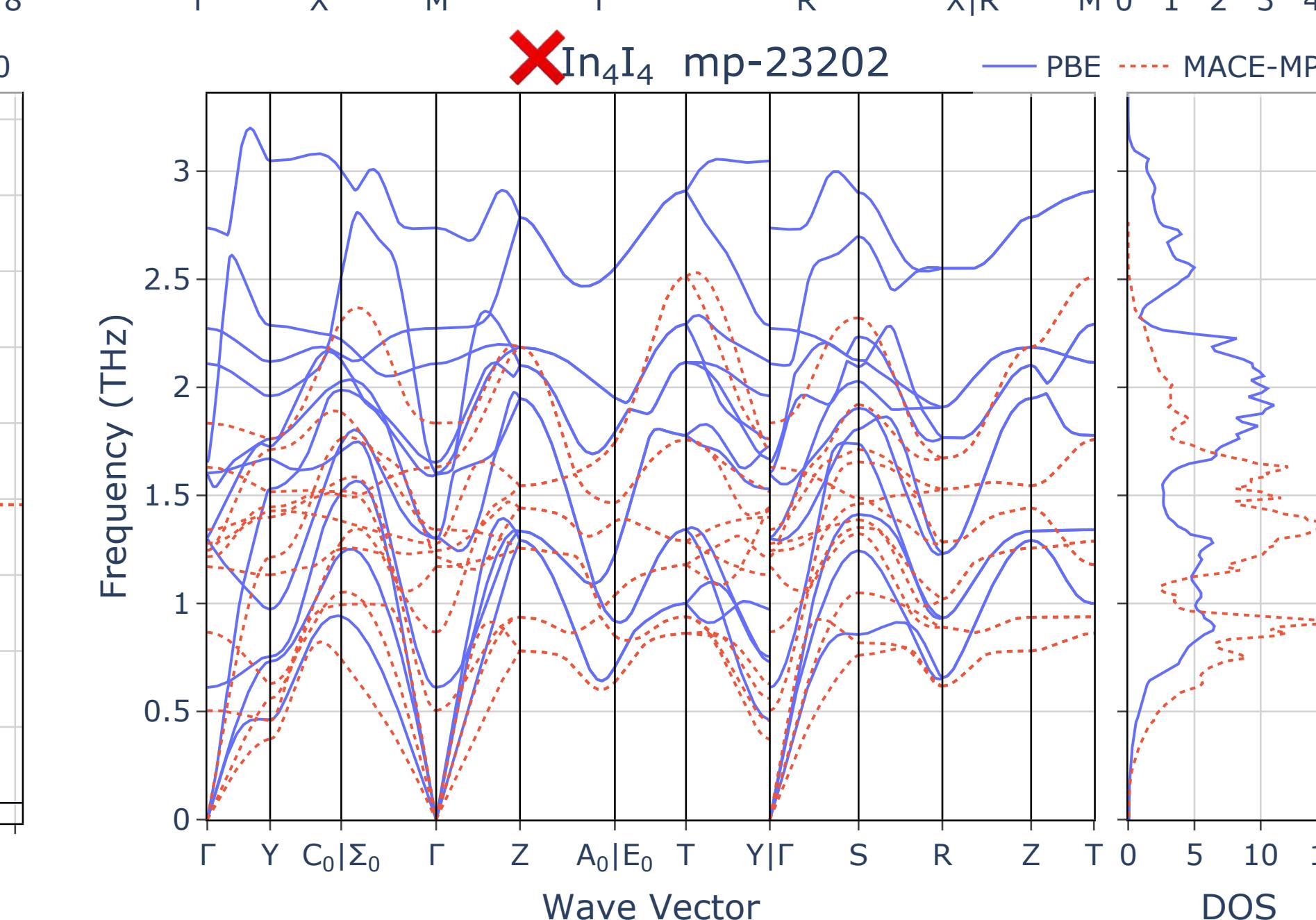
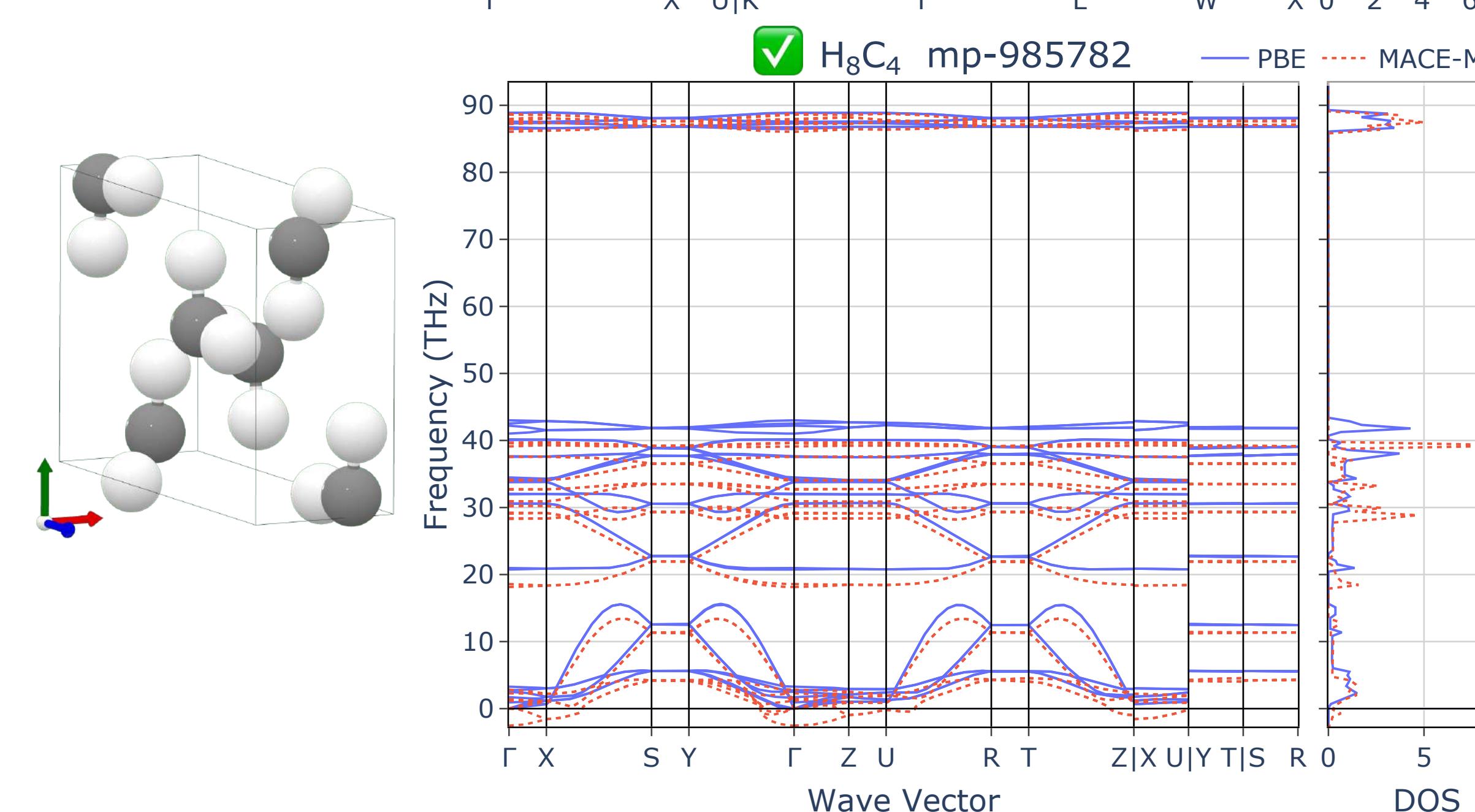
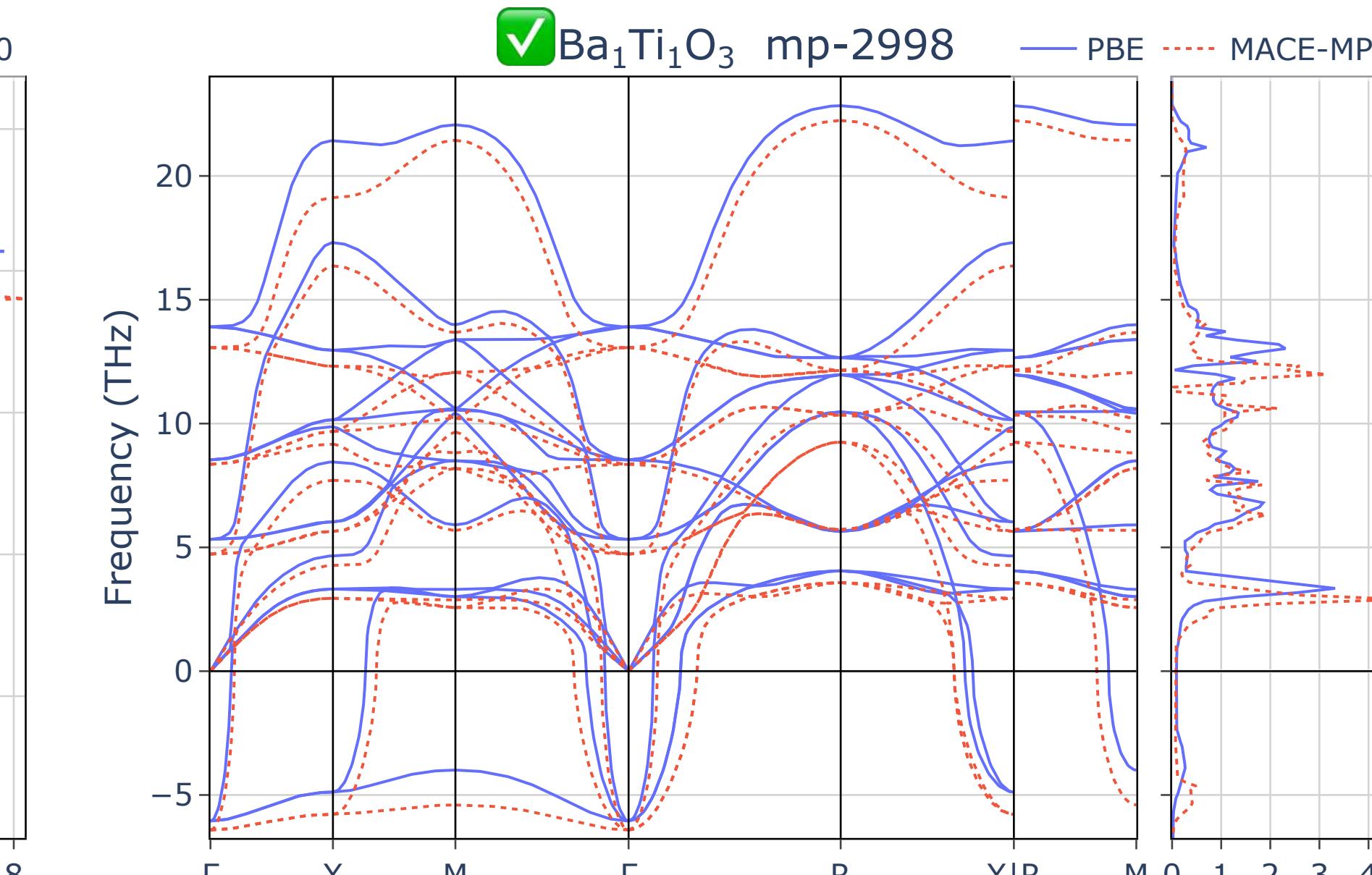
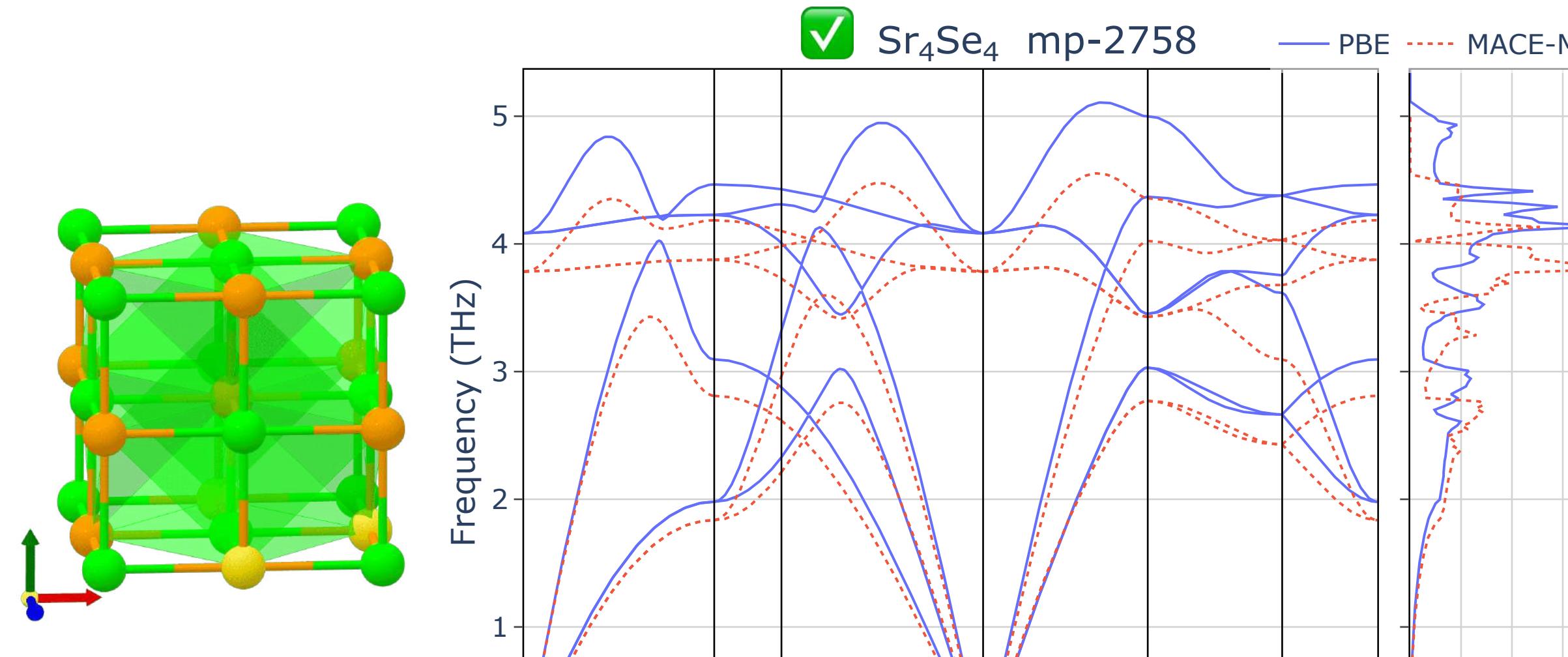
- Mark sch... - c9e1376 · 3 minutes ago · 2,504 Commits
- .github - Update GitHub Actions v... · last month
- docs - Enable all ruff rules by d... · 3 days ago
- src/atomate2 - Mark schema fields with ... · 3 minutes ago
- tests - Mark schema fields with ... · 3 minutes ago
- .codecov.yml - Update .codecov.yml · 2 years ago
- .gitattributes - Update .gitattributes · 2 years ago
- .gitignore - Add doctrees to gitignore · 8 months ago
- .pre-commit-config.yaml - Auto-update pre-commit ... · 3 days ago
- ADMIN.md - bump min emmet-core to... · 3 months ago
- CHANGELOG.md - Add optional 3rd static fo... · last month
- CODE_OF_CONDUCT.... - Refactor docs · 2 years ago

```
@dataclass
class PhononMaker(Maker):
    """
    Maker to calculate harmonic phonons with a force field.

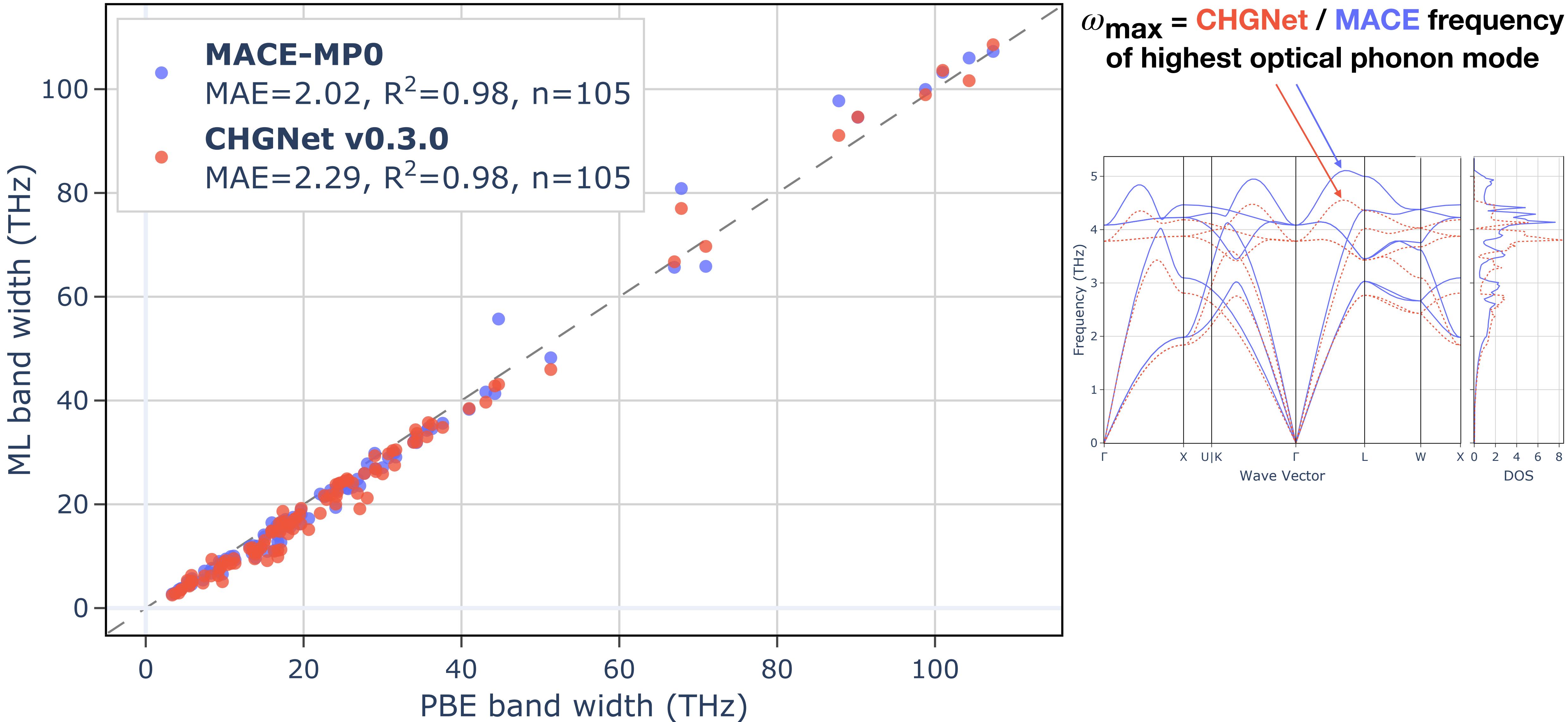
    Calculate the harmonic phonons of a material. Initially, a tight structural
    relaxation is performed to obtain a structure without forces on the atoms.
    Subsequently, supercells with one displaced atom are generated and accurate
    forces are computed for these structures. With the help of phonopy, these
    forces are then converted into a dynamical matrix. To correct for polarization
    effects, a correction of the dynamical matrix based on BORN charges can
    be performed. The BORN charges can be supplied manually.
    Finally, phonon densities of states, phonon band structures
    and thermodynamic properties are computed.

    .. Note::
        It is heavily recommended to symmetrize the structure before passing it to
        this flow. Otherwise, a different space group might be detected and too
        many displacement calculations will be generated.
        It is recommended to check the convergence parameters here and
        adjust them if necessary. The default might not be strict enough
        for your specific case.
    
```

Phonon Bands + DOS from MACE-MP



PBE vs MACE phonon band width ω_{\max}



PBE vs ML imaginary phonon modes

any k -point

MACE-MPO

		Stable	Unstable
PBE	Stable	True Stable 37.1%	False Unstable 7.6%
	Unstable	False Stable 12.4%	True Unstable 42.9%

Acc=80%, N=105, Tol=0.05

Γ -point only

MACE-MPO

		Γ -Stable	Γ -Unstable
PBE	Γ -Stable	True Γ -Stable 44.8%	False Γ -Unstable 5.7%
	Γ -Unstable	False Γ -Stable 12.4%	True Γ -Unstable 37.1%

Acc=82%, N=105, Tol=0.05

any k -point

CHGNet v0.3.0

		Γ -Stable	Γ -Unstable
PBE	Γ -Stable	True Γ -Stable 6.7%	False Γ -Unstable 43.8%
	Γ -Unstable	False Γ -Stable 1.0%	True Γ -Unstable 48.6%

Acc=55%, N=105, Tol=0.05

Γ -point only

CHGNet v0.3.0

		Stable	Unstable
PBE	Stable	True Stable 5.7%	False Unstable 39.0%
	Unstable	False Stable 0.0%	True Unstable 55.2%

Acc=61%, N=105, Tol=0.05

57% of 105 test structures are PBE-dynamically unstable

Future Research Directions

- Design large-scale phonon benchmark since thermodynamic stability seems "solved"
- How to address PES softness?
- What are good metrics to compare ML foundation models going forward?
- How to commoditize foundation models / make them usable by non-experts to unlock their full utility?
- How to best train foundation models on multiple modalities (slabs, defects, amorphous, 2D, molecules, ...)

Acknowledgements

Rhys Goodall



Yuan Chiang



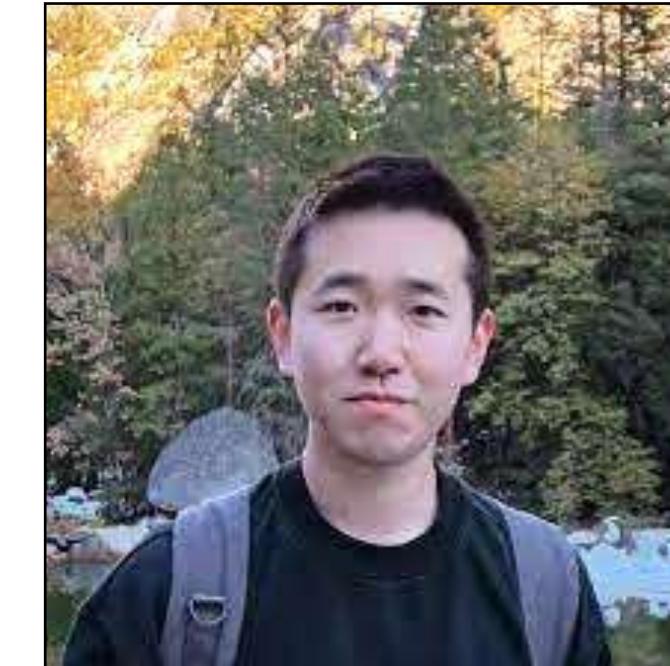
Philipp Benner



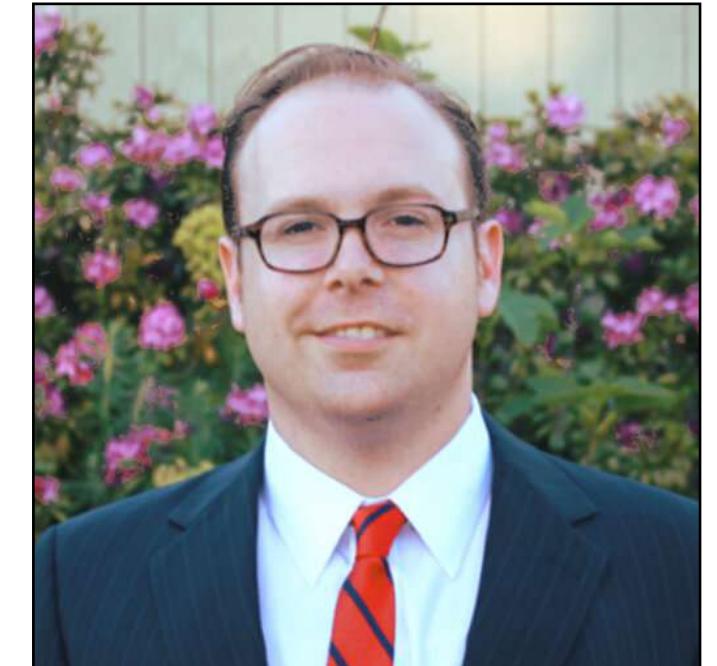
Aaron Kaplan



Bowen Deng



Wesley Surta



Michael Gaulois



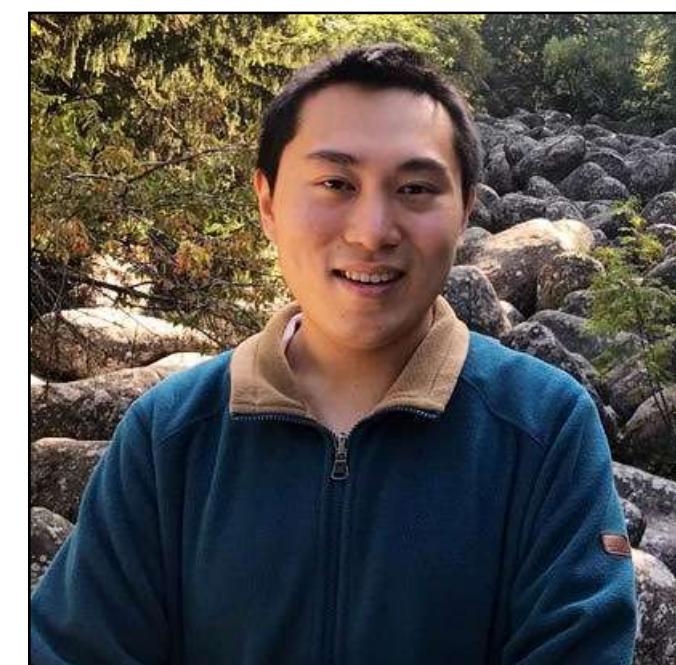
Gábor Csányi



Shyue Ping Ong



Alpha Lee



Kristin Persson



Anubhav Jain





Questions?