ICML
International Conference
On Machine Learning

NUS National University of Singapore | Institute for Functional Intelligent Materials

Wyckoff Transformer: Generation of Symmetric Crystals

Nikita Kazeev, Wei Nong, Ignat Romanov, Ruiming Zhu, Andrey Ustyuzhanin, Shuya Yamazaki, Kedar Hippalgaonkar

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

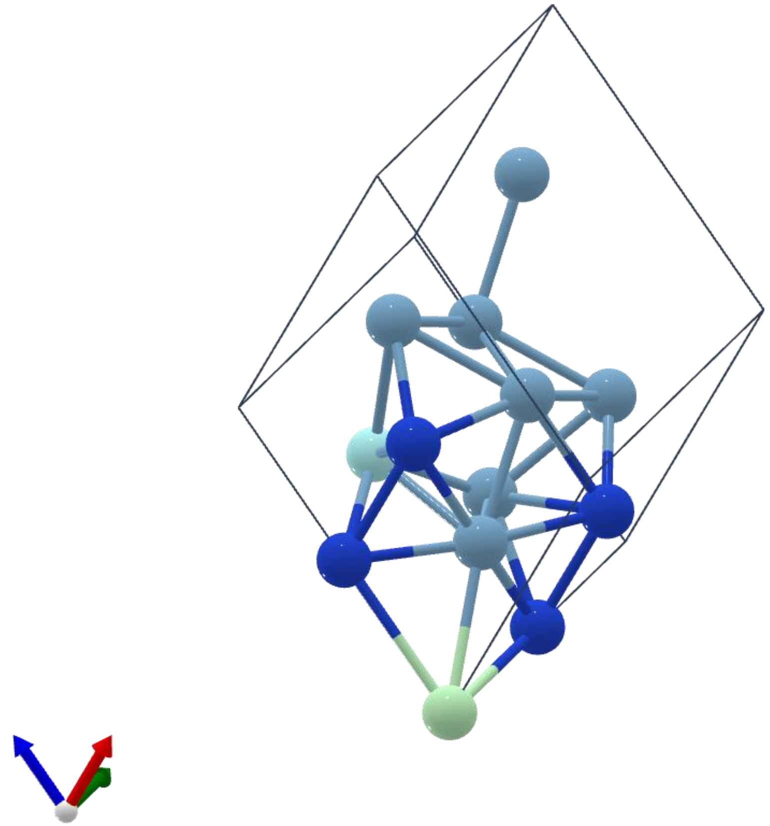
The space of possible materials is intractably vast

A rough estimate: 10^{160}

Even if quantum mechanical computation (DFT) is replaced with a 1000x faster algorithm we can't screen all

Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025

Material Generation Problem Statement



- A crystal consists of:
  - A 3D point cloud of atoms, including coordinates and chemical elements
  - Lattice parameters
- The goal is to generate crystals that are
  - Stable under normal conditions
  - Novel: not present in the training data

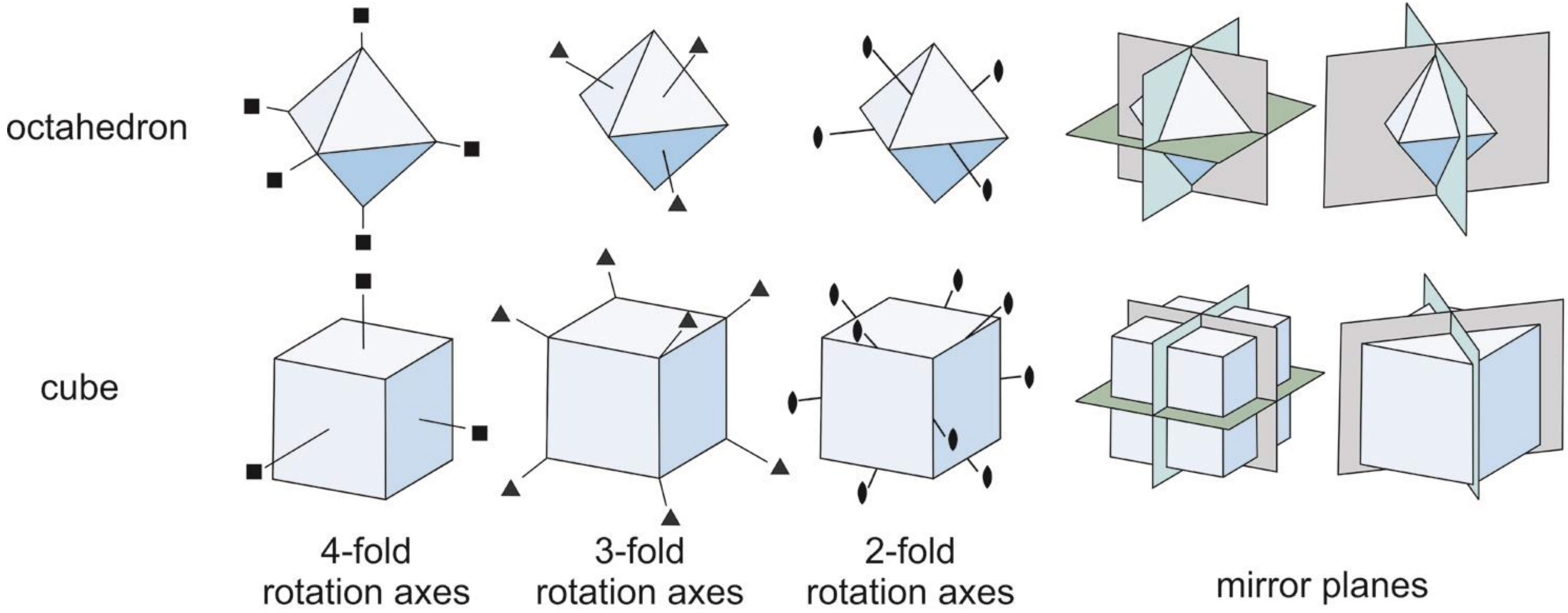
Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025

ML Challenges

1. Atom types are discrete, coordinates are continuous
2. The number of atoms varies
3. Structures are permutation, rotation and translation invariant
4. **Lots of local energy minima**

Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025

Symmetries in Crystals

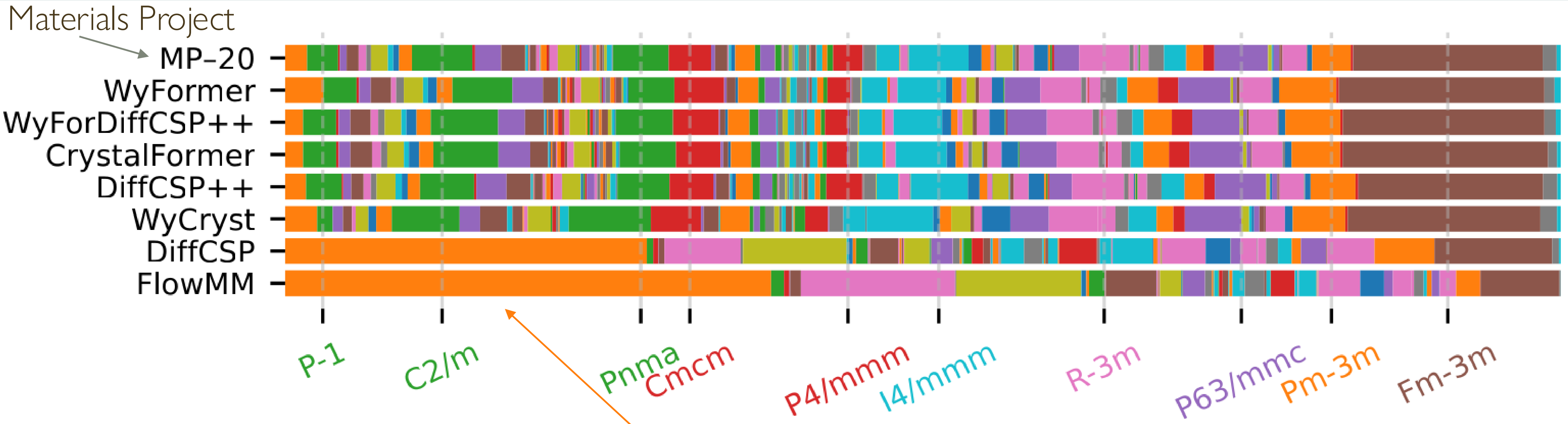


Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025

Not all of the symmetry elements are shown in these drawings

https://opengeology.org/Mineralogy/10-crystal-morphology-and-symmetry/

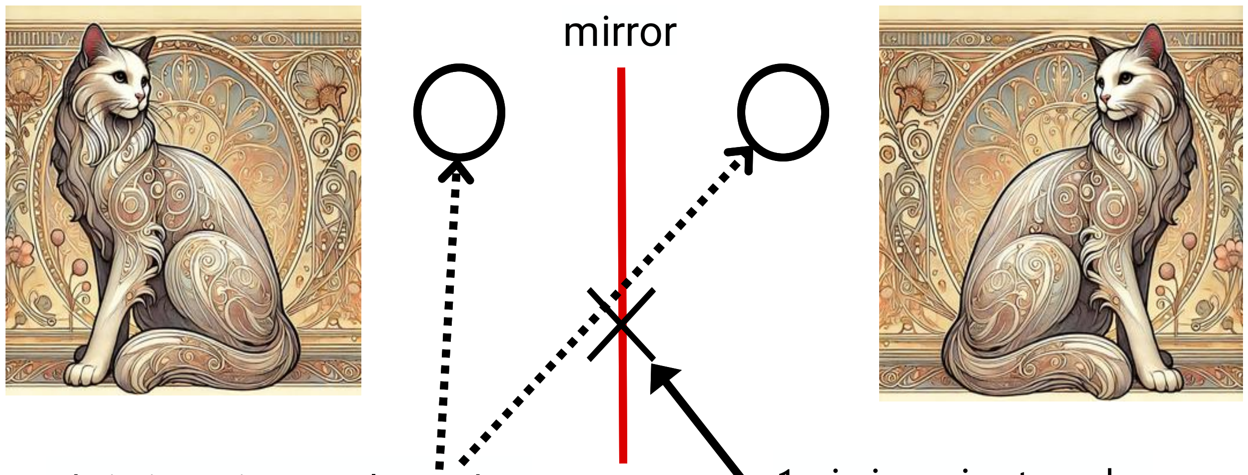
Most Real Crystals Have Internal Symmetry



Space group other than P1 indicates presence of symmetry beyond the lattice translation

Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025

Wyckoff Positions



2b is invariant only under identity transformation. The two o's are just examples; almost whole plane is 2b

1a is invariant under mirror transformation. The two x are just examples; the whole line is 1a

Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025

Wyckoff Positions for a Space Group

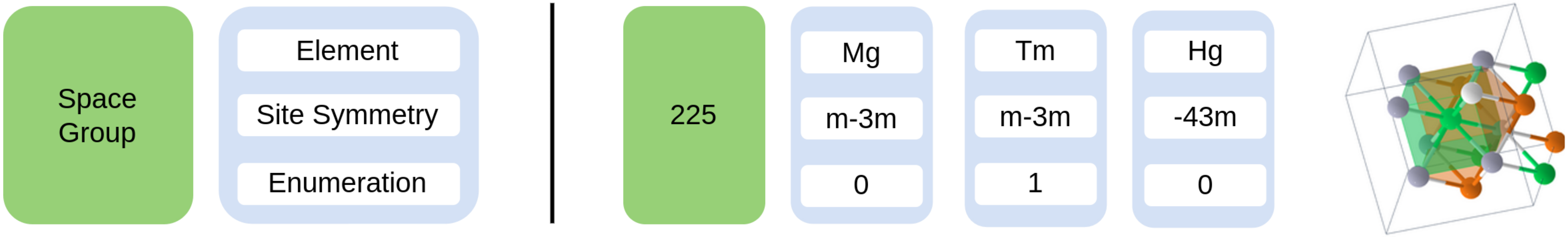
Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025

Wyckoff Positions of Group *I4/mmm* (No. 139)

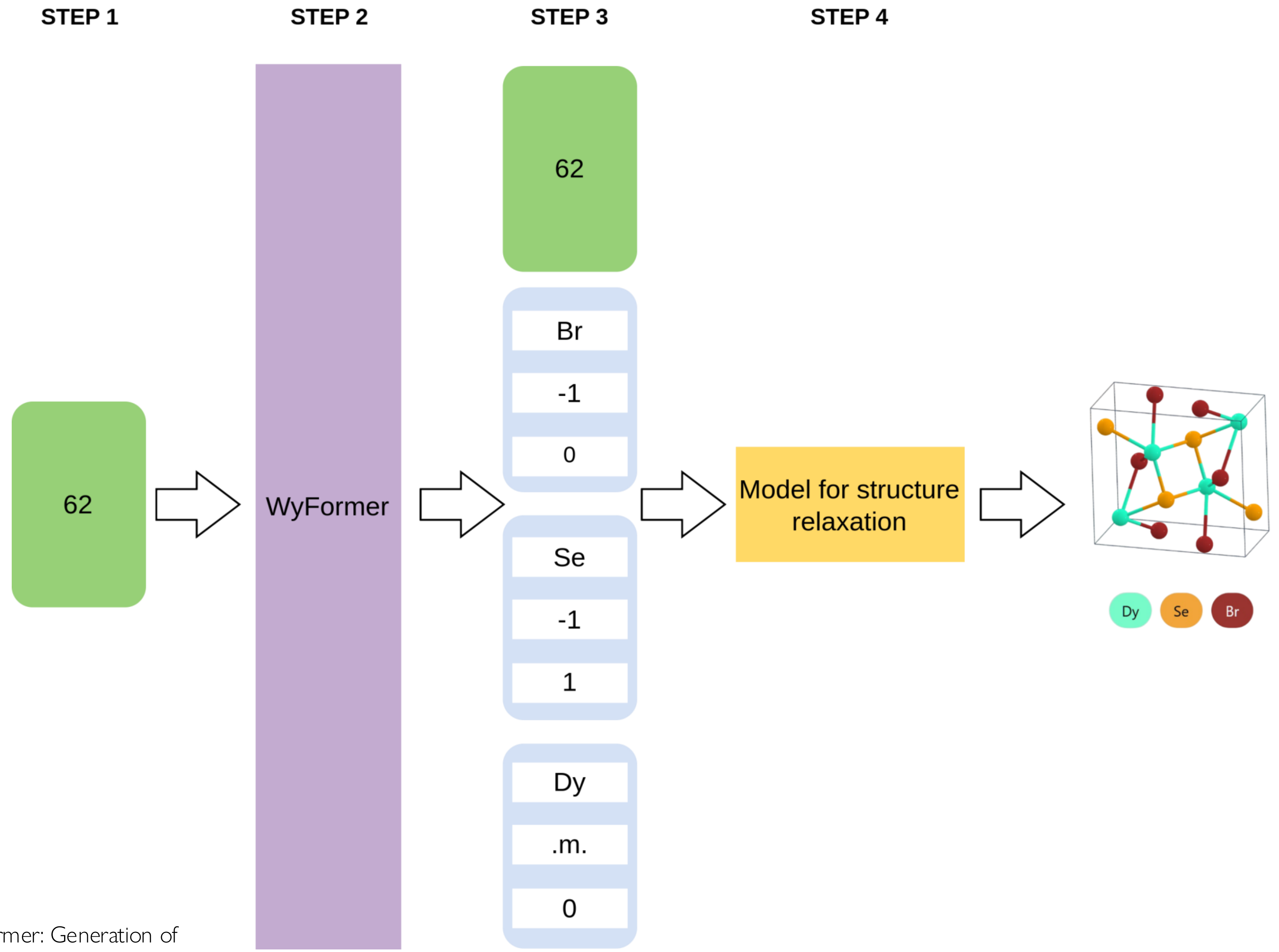
Multiplicity	Wyckoff letter	Site symmetry	Coordinates (0,0,0) + (1/2,1/2,1/2) +
32	o	1	(x,y,z) (-x,-y,z) (-y,x,z) (y,-x,z) (-x,y,-z) (x,-y,-z) (y,x,-z) (-y,-x,-z) (-x,-y,-z) (x,y,-z) (y,-x,-z) (-y,x,-z) (x,-y,z) (-x,y,z) (-y,-x,z) (y,x,z)
16	n	.m.	(0,y,z) (0,-y,z) (-y,0,z) (y,0,z) (0,y,-z) (0,-y,-z) (y,0,-z) (-y,0,-z)
16	m	..m	(x,x,z) (-x,-x,z) (-x,x,z) (x,-x,z) (-x,x,-z) (x,-x,-z) (x,x,-z) (-x,-x,-z)
16	l	m..	(x,y,0) (-x,-y,0) (-y,x,0) (y,-x,0) (-x,y,0) (x,-y,0) (y,x,0) (-y,-x,0)
16	k	..2	(x,x+1/2,1/4) (-x,-x+1/2,1/4) (-x+1/2,x,1/4) (x+1/2,-x,1/4) (-x,-x+1/2,3/4) (x,x+1/2,3/4) (x+1/2,-x,3/4) (-x+1/2,x,3/4)
8	j	m2m .	(x,1/2,0) (-x,1/2,0) (1/2,x,0) (1/2,-x,0)
8	i	m2m .	(x,0,0) (-x,0,0) (0,x,0) (0,-x,0)
8	h	m.2 m	(x,x,0) (-x,-x,0) (-x,x,0) (x,-x,0)
8	g	2mm .	(0,1/2,z) (1/2,0,z) (0,1/2,-z) (1/2,0,-z)
8	f	..2/m	(1/4,1/4,1/4) (3/4,3/4,1/4) (3/4,1/4,1/4) (1/4,3/4,1/4)
4	e	4mm	(0,0,z) (0,0,-z)
4	d	-4m2	(0,1/2,1/4) (1/2,0,1/4)
4	c	mmm .	(0,1/2,0) (1/2,0,0)
2	b	4/mmm	(0,0,1/2)
2	a	4/mmm	(0,0,0)

Aroyo, et. al. Zeitschrift fuer Kristallographie (2006), 221, 1, 15-27.

Wyckoff Positions as Tokens

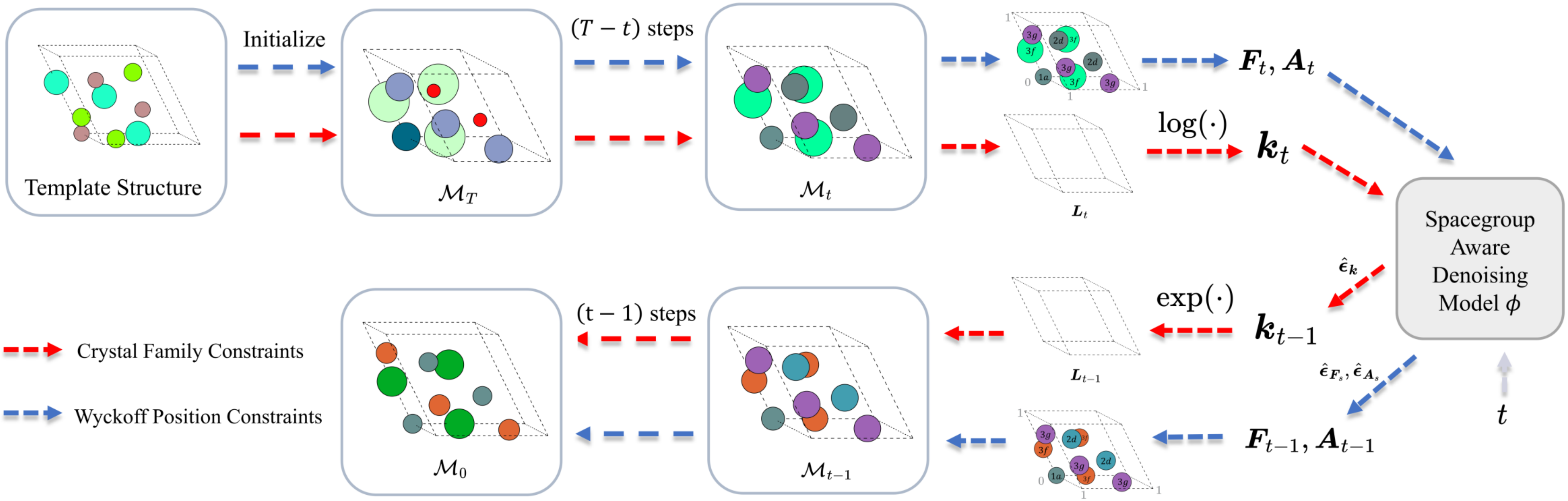


Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025



Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025

Relaxation: DiffCSP++



Jiao, Rui, et al. "Crystal structure prediction by joint equivariant diffusion." Advances in Neural Information Processing Systems 36 (2023): 17464-17497.

Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025

Wyckoff Transformer

An autoregressive generative Transformer:

- Encoder-only
- Permutation-invariant: no positional encoding, shuffle token order
- Predict the probability for the next part of token

Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025

Results

Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025

Table 1. Evaluation. Symmetry metrics are computed only using novel structurally valid examples. Note that the 1000 and 105-example metrics are computed using MP-20 train and validation as reference datasets for novelty, while the 10 000-example S.U.N. only uses MP-20 train to remain compatible with the reported values. **Bold** indicates the values within $p = 0.1$ statistical significance threshold from the best. Values marked by * were computed by Miller et al. (2024), the rest by us; see note H.1 for an important caveat; in short, the values in (brackets) are less accurate, but are compatible with each other.

Method/Metric	**Novel Unique** Templates (#) ↑	**P1 (%)** ref = 1.7	**Space Group** χ^2 ↓	**S.U.N. % ↑** $E_{\text{hull}} < 80$ meV	**S.S.U.N. % ↑**	**S.U.N. % ↑** $E_{\text{hull}} < 0$ meV
Sample size	1000	1000	1000	105	105	10 000
Relaxation	CHGNet	CHGNet	CHGNet	DFT	DFT	DFT
WyFormerCHGNet	**180**	3.24	0.223	**23.1**	**22.3**	–
WyFormerDiffCSP++	**186**	**1.46**	**0.212**	**22.2**	**21.1**	**3.83 (4.14)**
DiffCSP++	*10*	2.57	0.255	14.4	**14.4**	–
CrystalFormer	74	**0.91**	0.276	**20.1**	**20.1**	–
SymmCD	101	**2.35**	0.24	**20.7**	**20.7**	–
WyCryst	**165**	4.79	0.710	5.5	5.5	–
DiffCSP	76	36.57	7.989	**22.2**	**20.6**	– (3.34*)
FlowMM	51	44.27	12.423	**17.8**	**16.9**	– (2.34*)
WyFormer *MPTS–52*	386	0	0.225	–	–	–

The nature is elegant, so is WyFormer

WyFormer

- Best-in-class symmetry-conditioned generation
- Relative 20% S.U.N. improvement
- Physics-motivated inductive bias
- Unparalleled inference speed

Kazeev et al., Wyckoff transformer: Generation of Symmetric Crystals, ICML 2025