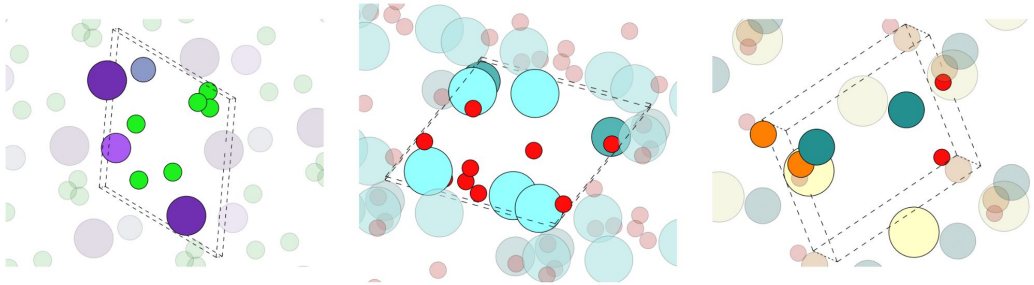


Conditional Crystal Structure Generation Using FlowMM

Oliver Tang†, Ray Zhu†,

Faradawn Yang, Hanlei Wen, Haroan Wu

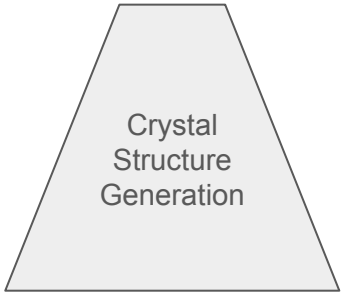
https://github.com/oytang/Agent47-Material

Miller, Benjamin Kurt, et al. "FlowMM: Generating Materials with Riemannian Flow Matching." arXiv preprint arXiv:2406.04713 (2024).

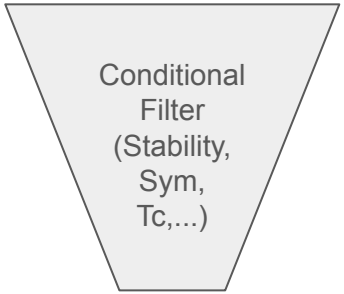
Crystal Design

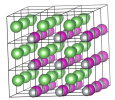
Problem Statement

Limited constraints



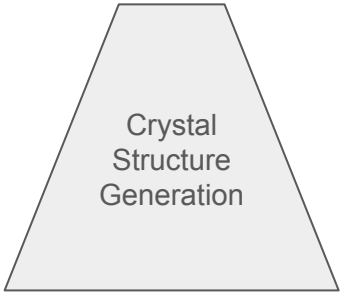
Generated structures

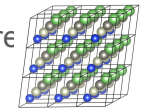


Designed structure 

- Success rate
- Stability
- Physical realisticity and symmetry
- Multi-objective generation (e.g. functionality and stability)

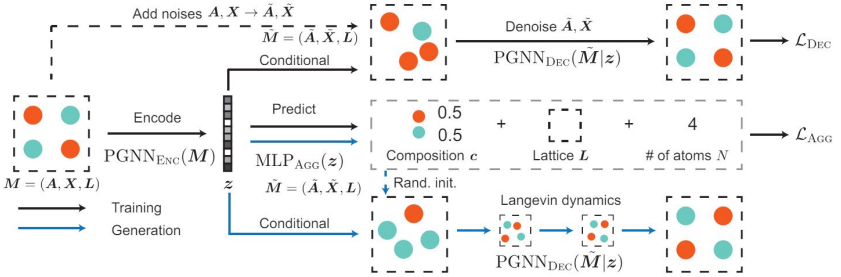
All necessary constraints (Stability, Sym, Tc,...)



Designed structure 

Crystal Design

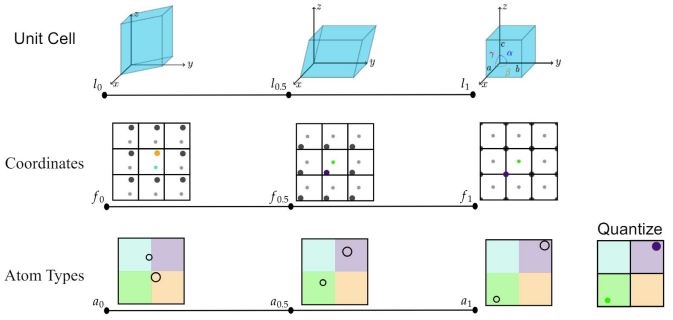
Relevant work and practical insight



CDVAE (https://arxiv.org/abs/2110.06197)

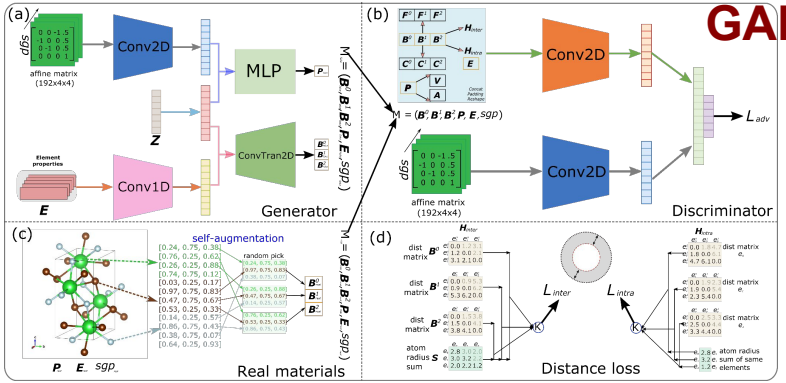
VAE

Flow



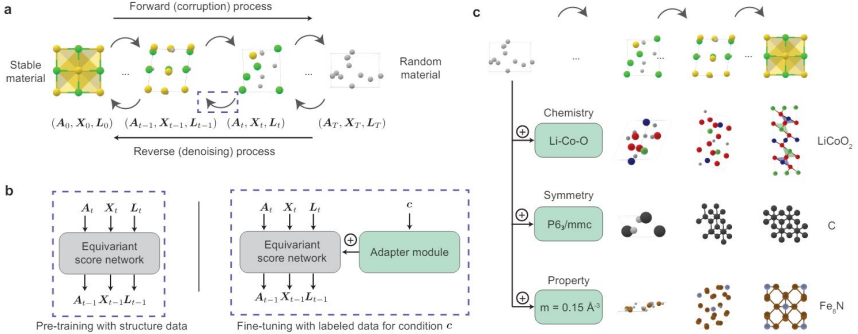
FlowMM (https://arxiv.org/abs/2406.04713)

GAN



PGCGM (https://arxiv.org/abs/2203.14352)

Diffusion



MatterGen (https://arxiv.org/abs/2312.03687)

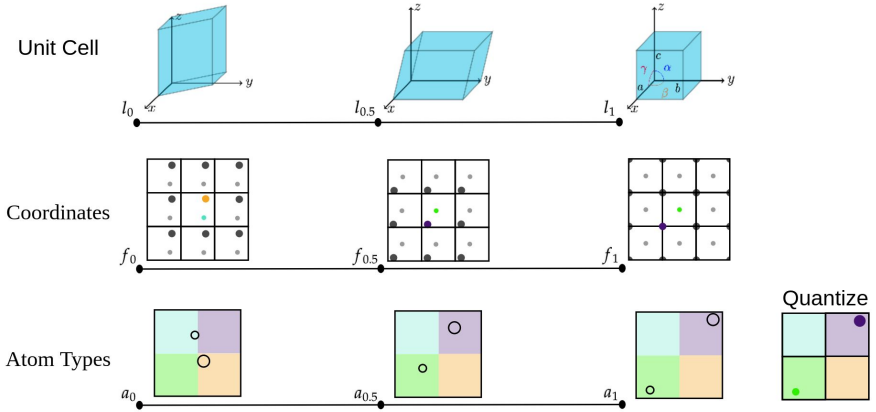
FlowMM

Riemannian Flow Matching

Fit probability distribution over crystals

Procedure

1. Represent crystal symmetries and leveraging Niggli reduction for unique unit cell representation.
2. Use permutation, translation, and rotation invariant density functions.
3. Employ a binary representation for atomic types to reduce dimensionality.
4. Apply a new objective to estimate density while maintaining symmetry invariances.
5. Conduct extensive quantum chemistry calculations to validate generated structures.

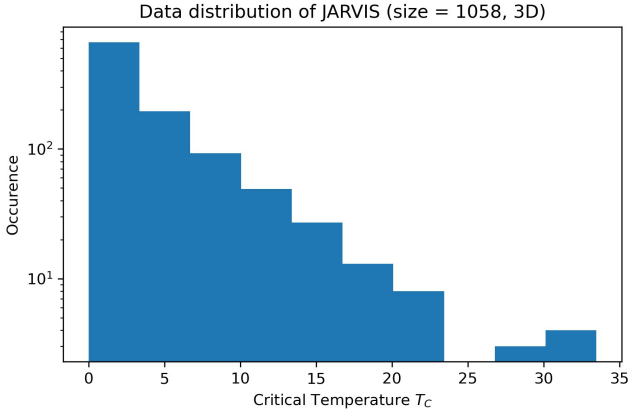


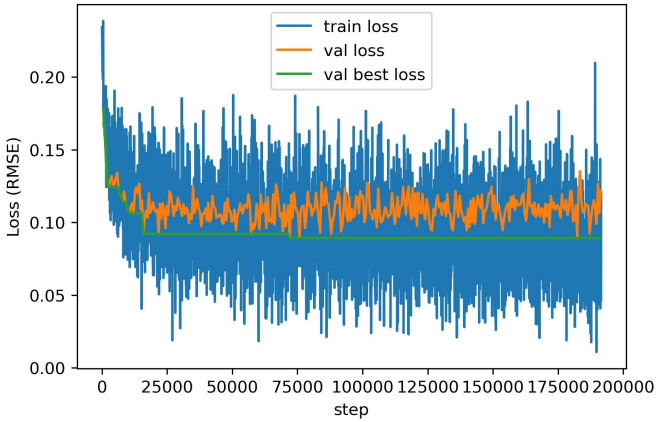
Miller, Benjamin Kurt, et al. "FlowMM: Generating Materials with Riemannian Flow Matching." arXiv preprint arXiv:2406.04713 (2024).

Model Training

Data

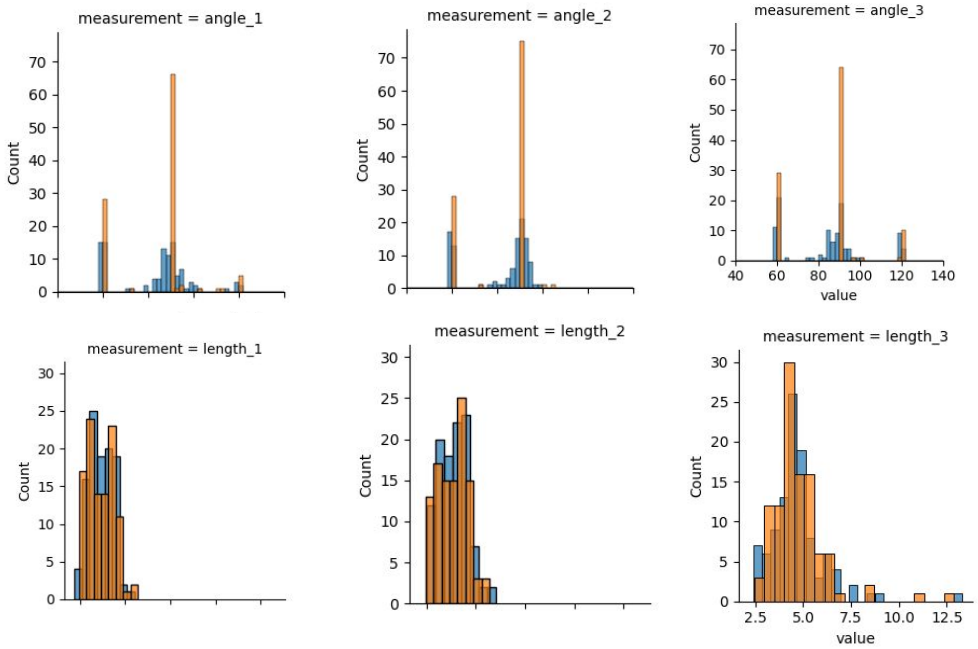
- Superconductor database - JARVIS 3D
  - 1058 structures with Tc
- Crystal structure data conversion
  - JARVIS atoms object -> CIF
- Train/val/test split = 8:1:1
  - 846:106:106 structures
- Best model at 679 epochs
  - ~24 hours training on NVIDIA A100
  - <12 hours with distributed training





Conditional Generation

Crystal Structure Prediction



Task: reconstruct the test set

# of steps: 1000

Dataset	Match Rate (%)	RMSE	Size
JARVIS-3D	**47.00**	**0.21**	**1058**
MP-20	61.39	0.0566	45231
Perov-5	53.15	0.0992	18928
Carbon-24	23.47	0.4122	10153
MPTS-52	17.54	0.1726	40476

Unconditional Generation

De Novo Generation

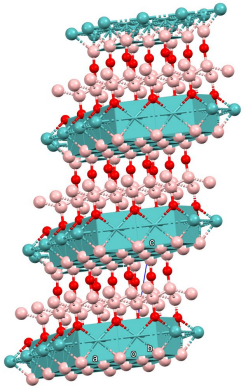
Structure generation
(10,000 structures)

↓

Structure relaxation using CHGNet

Dataset	MP-20	JARVIS-3D
Size	**45231**	**1058**
Validity % Structural	96.85	88.21
Validity % Composition	83.19	68.43
Wdist (ρ)	0.239	0.533
Wdist (N_{el})	0.083	1.122

Example

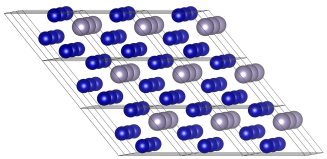


Spacegroup: Cm

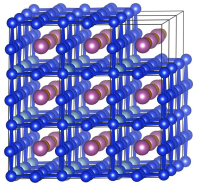
E(gen) = -33.78 eV/atom

E(relax) = -44.76 eV/atom

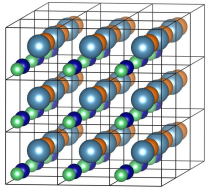
De Novo Generation



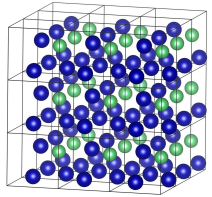
Co3Sn
I4/mmm
E = -25.11 eV/atom
Ehull = 111 meV/atom



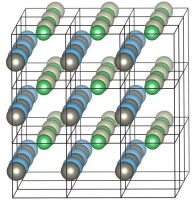
ScAlFeSi
P4/mmm
E = -46.03 eV/atom
Ehull = 321 meV/atom



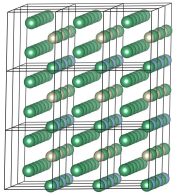
CaMgBeCr
P4mm
E = -15.13 eV/atom
Ehull = 587 meV/atom



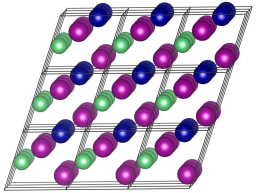
BeCo3
Pm-3m
E = -25.33 eV/atom
Ehull = 85 meV/atom



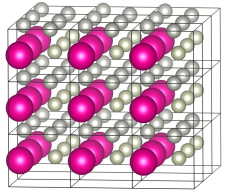
TiBeZnRh
P4mm
E = -21.98 eV/atom
Ehull = 67 meV/atom



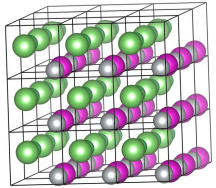
Nb6AlRh
Cmmm
E = -73.00 eVatom
Ehull = 112 meV/atom



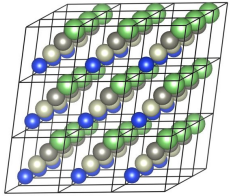
Mn2BeCr
Fm-3m
E = -31.93 eV/atom
Ehull = 40 meV/atom



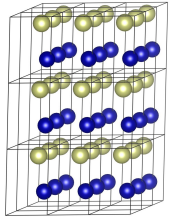
RbPd2Rh
P4/mmm
E = -16.40 eV/atom
Ehull = 575 meV/atom



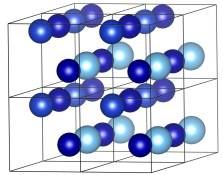
LiBeCdNi
P4mm
E = -21.91 eV/atom
Ehull = 80 meV/atom



LiZrSiRh
F-43m
E = -17.97 eV/atom
Ehull = 22 meV/atom



CoIr
P-6m2
E = -15.88 eV/atom
Ehull = 31 meV/atom



TiCo2Cu
P4mm
E = -26.55 eV/atom
Ehull = 145 meV/atom

Conclusion

FlowMM is an effective and efficient generative model for predicting and generating stable crystal structures.

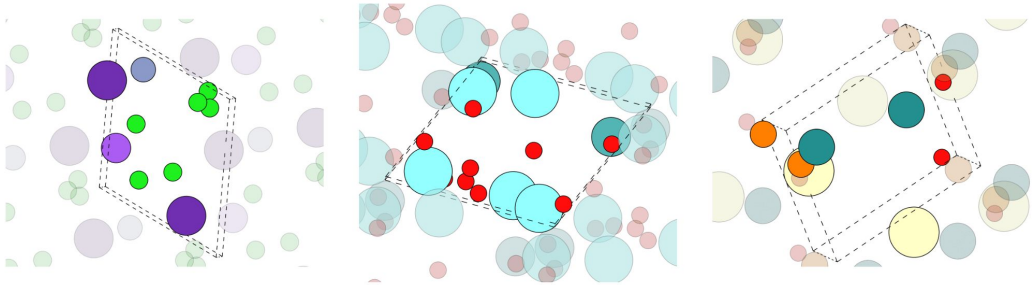
By leveraging Riemannian Flow Matching, FlowMM achieves superior performance and efficiency in model training and evaluation.

Outlook

1. Reduce the number of integration steps required for FlowMM to generate stable materials
2. Incorporate DFT calculation to further validate the generated structure
3. Expand its application to larger datasets and different types of materials

Project Github:

https://github.com/oytang/Agent47-Material



Conditional Crystal Structure Generation Using FlowMM

Oliver Tang†, Ray Zhu†,

Faradawn Yang, Hanlei Wen, Haroan Wu

https://github.com/oytang/Agent47-Material

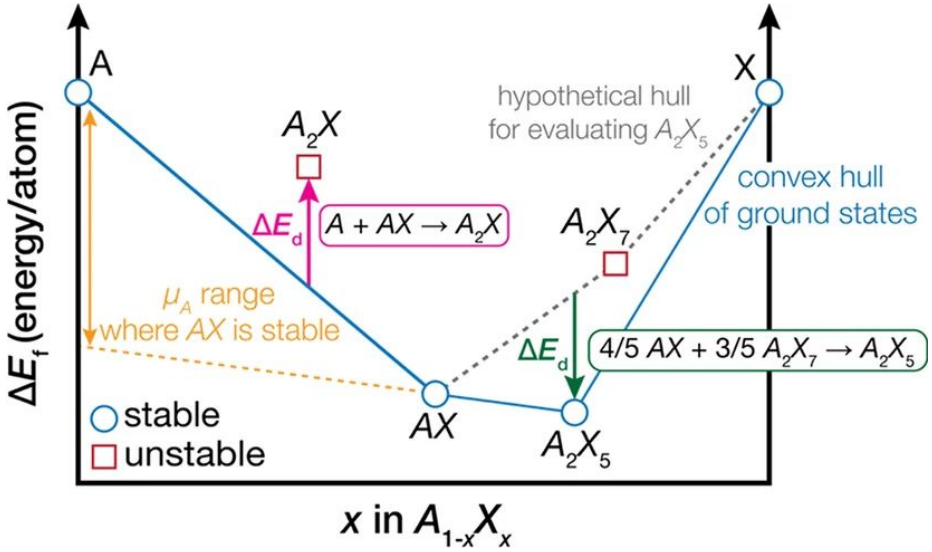
Miller, Benjamin Kurt, et al. "FlowMM: Generating Materials with Riemannian Flow Matching." arXiv preprint arXiv:2406.04713 (2024).

Backup Slides

Convex Hull Energy Calculation

Tool: *pymatgen.analysis.phase_diagram.PhaseDiagram*

Database: Materials Project



https://docs.materialsproject.org/methodology/materials-methodology/thermodynamic-stability/phase-diagrams-pds