A Comparative Study of Flow-Based and Diffusion-Based 3D Molecule Generation Methods

Enyi Jiang (enyij2), Lily Xie (hexuanx2), Huyen Nguyen (huyentn2)

Final Presentation

OUTLINE

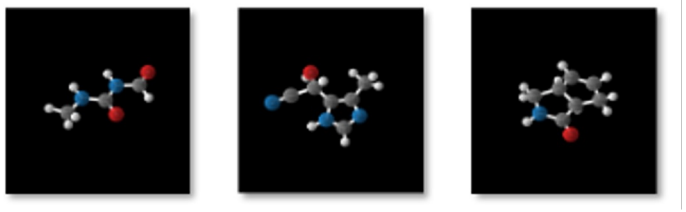
- Background and Motivation:
  - Small molecule generation
- Method
  - Geo-LDM
  - FlowMol
- Evaluation metrics
- Results:
  - Show results and discussion
- Conclusion

Small Molecule Generation

- Small molecules are low molecular weight (typically <900 Da) organic compounds that can enter cells easily and modulate biological processes by interacting with specific molecules targets. (Lipinski et al., 2001)
- Generative models enables de novo molecule generation, learn complex chemical distribution from large datasets, and design molecules
- Commonly used for drug discovery, material design, and catalyst development.

QM9

- Quantum chemistry dataset
- Detailed molecular properties for 134,000 small organic molecules



Motivation

- 3D molecule generation is key in drug discovery, material science, and protein design.
- Goal: evaluate and benchmark two state of the art generative model for 3D molecule generation
- Two major paradigms:
  - Flow-based methods (e.g. FlowMol)
  - Diffusion-based methods (e.g. GeoLDM)
- Understanding their strengths and weaknesses

Overview of GeoLDM

- Geometric Latent Diffusion Model
- First latent DM for the molecular geometry domain, composed of autoencoders encoding structures into continuous latent codes and DMs operating in the latent space
- Generate 3D molecule with high validity
- Controllable generation
  - target polarizability, dipole moment
- Scalable to large biomolecule
- Use VAE backbone

GeoLDM Training Pipeline

- Input: 3D molecular structures
- Encoding: SE(3)-equivariant autoencoder
- Latent diffusion process
  - forward (noise adding): Gaussian
  - reverse (denosing): Transformer based
- Decode: reconstruct 3D molecule
  - SE(3)-equivariant Decoder

SE(3)-Equivariant

- Special Euclidean group in 3D, represent all possible rotations and translation in 3D space
- If a molecule is rotated or translated, its latent representation and output remain consistent
- Good for 3D structure generation

Overview of FlowMol

- Type: Flow-based generative model
- Uses both continuous-time and discrete flow matching to model joint distribution of atom types and 3D coordinates
- Key Features:

❑ Model joint distribution of atom types + 3D positions
❑ Flow Matching to learn the distributions
❑ Used SimplexFlow to model discrete flow of atom coordinates and compared with naive Flow Matching.

Motivation

Problem: Generating valid, stable 3D molecules from scratch

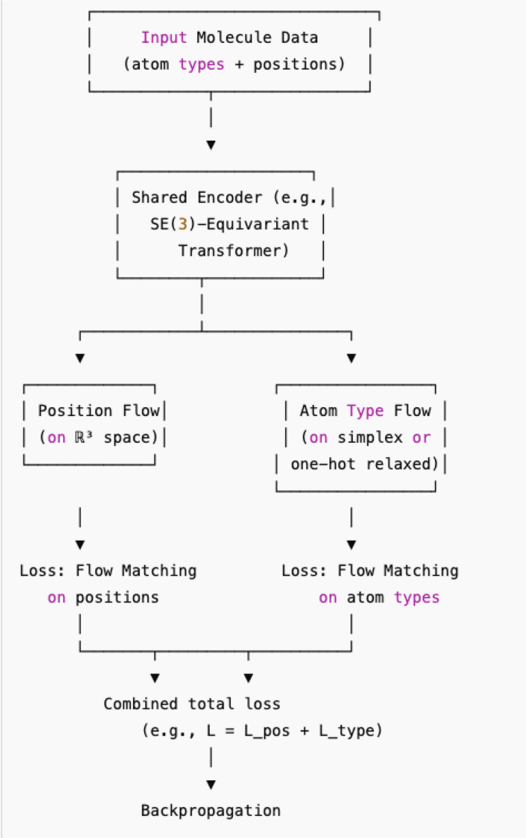
Challenges:

- Continuous (3D positions) **and** categorical (atom types) data
- Need for efficient, accurate generative model

Why Flow Matching?

Criterion	Flow Matching	Diffusion Models
Sampling Speed	**Faster** (one-shot generation)	**Slower** (iterative process)
Training Complexity	**Moderate** (direct flow learning)	**Higher** (requires denoising steps and noise schedules)
Equivariance for 3D Geometries	**Built-in SE(3)-equivariance**	**Requires extra effort**
Handling Categorical + Continuous	**Naturally suited**	**Not ideal** without modifications
Scalability	**Highly scalable**	**Less scalable** due to iterative process

FlowMol Architecture



$$X = \{x_i\}_{i=1}^{N} \in \mathbb{R}^{N\times 3}$$

$$E = \{e_{ij}\ \forall i,j \in [N] \mid i \neq j\} \in \mathbb{R}^{(N^2-N)\times n}$$

$$C = \{c_i\}_{i=1}^{N} \in \mathbb{R}^{N\times n_c} \qquad A = \{a_i\}_{i=1}^{N} \in \mathbb{R}^{N\times n_a}$$

The molecule g = (X, A, C, E)

is represented with a graph, and a shared encoder is used to encode all embeddings of atom types, coordinates, charge type and bond order types.

FlowMol Architecture

The total loss is weight across different modalities:

$$L = \eta_X L_X + \eta_A L_A + \eta_C L_C + \eta_E L_E$$

L_X: Loss on the **positions** $\mathcal{L}_{EP} = \mathbb{E}_{t,g_t}\left[\frac{\alpha'_t}{1-\alpha_t}||\hat{g}_1(g_t) - g_1||\right]$

L_A, L_C, L_E: Loss on the **atom types, charge type and bond order types;**

categorical distribution, cross-entropy loss

Contributions of FlowMol

- FlowMol **does train multiple flows**, one for each type of variable — **position and atom type**.
- Each flow is adapted to the nature of the variable: continuous for geometry, categorical or relaxed-continuous for atom types.
- Despite theoretical elegance, the naive treatment of categorical variables **outperformed** the more sophisticated SimplexFlow.
- FlowMol improves upon the existing state of the art flow matching method for molecule generation; however, it still does not outperform diffusion models trained for the same task

Evaluation Metrics

- Validity: % of chemically valid molecules (A generated molecule must be **chemically valid**, meaning it adheres to known rules of chemical bonding (e.g., no disallowed valences, proper atom types, and bond orders)).
- Uniqueness: % of outputs that are unique

$$\text{Uniqueness} = \frac{\text{Number of unique molecules}}{\text{Total number of generated molecules}}$$

- Diversity: **average pairwise dissimilarity, measure** how **different** they are in terms of structure, function, or properties.

$$\text{Diversity} = 1 - \text{Average Tanimoto Similarity}$$

Evaluation Metrics

- MMFF Energy: estimates the **total potential energy** of a molecule using classical physics-based approximation; it shows physical plausibility (molecule with **high MMFF energy** and are likely **invalid or non-synthesizable)**.

- Equivariance Error (SE(3) robustness): measures how much the model's output **deviates** when the **input molecule is transformed** under SE(3).

$$\text{Equivariance Error} = \|f(RX + t) - Rf(X) - t\|$$

By physical realism, molecular properties and relationships should be invariant or equivariant to spatial changes, so the smaller Equivariance Error the better.

Experiment Setting

- Dataset: QM9
- Models: FlowMol, GeoLDM
  - FlowMol: CTMC flow type
  - GeoLDM: Geometric Latent Diffusion Models
- Sampled molecule number: 10,000

Experimental Results
(Valid/Complete/Valid & Complete)

	Valid	Complete	Valid & Complete
FlowMol	**1.00**	0.998	**9677.00**
GeoLDM	**1.00**	**1.000**	9514.00

GeoLDM has better completeness, yet FlowMol is better for valid & complete

Experimental Results

(Avg molecule size/Atom Stability/Mol Stability)

	Avg molecule size	Atom Stability	Mol Stability
FlowMol	8.549	0.999	0.989
GeoLDM	**8.850**	**1.000**	**1.000**

GeoLDM has larger molecule size and stability

Experimental Results
(Uniqueness/Diversity/Novelty)

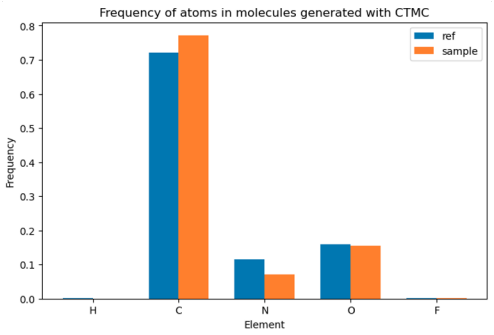
	Uniqueness	Diversity (Avg/Med)	Novelty
FlowMol	0.962	0.8718/0.8926	**0.871**
GeoLDM	**0.980**	**0.8936 / 0.9091**	0.640

GeoLDM has better uniqueness and diversity;
FlowMol has significantly better novelty.

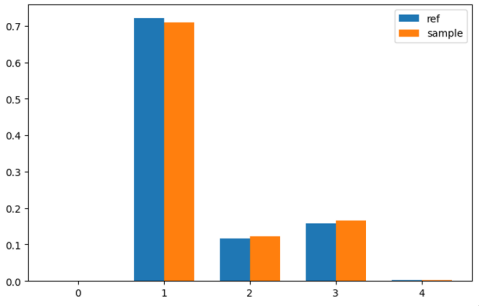
Experimental Results

(Sampling Runtime / Frequency of Atoms)

	Runtime (100 molecules)
FlowMol	**14.034 s**
GeoLDM	85.923 s



FlowMol



GeoLDM

FlowMol has a quicker runtime; GeoLDM has a closer distribution of atoms compared with reference

Strengths & Weaknesses

- FlowMol:

+ Efficient sampling
+ Validity & completeness
+ Novelty
- Molecule size
- Stability
- Uniqueness
- Diversity

- GeoLDM:

+ Molecule size
+ Stability
+ Uniqueness
+ Diversity
- Slower inference
- Validity & completeness
- Novelty

Analysis

Feature	FlowMol (Flow-based)	GeoLDM (Diffusion-based)
Sampling Efficiency	One-shot generation (fast)	Iterative denoising (slow)
Validity & Completeness	Invertibility ensures valid outputs, structural mapping encourages full molecules	Denoising errors can cause invalidity, may fail to fully reconstruct molecules
Stability	No correction mechanism	Stepwise correction improves stability
Molecule Size	Simpler structures preferred	Can handle large/complex molecules
Novelty	Continuous time space enables novelty	Training bias may reduce novelty
Uniqueness/Diversity	Lower due to mode collapse	Higher due to noise-based sampling

Conclusion

- Flow-based and diffusion-based methods both advance 3D molecule generation
- FlowMol: speed & novelty & validity/completeness
- GeoLDM: stability & diversity & uniqueness
- Future: design unified framework for combining the strengths of two lines of work

References & Links

- FlowMol: github.com/Dunni3/FlowMol
- Dunn, I., & Koes, D. R. (2024). Mixed continuous and categorical flow matching for 3d de novo molecule generation. *ArXiv*, arXiv-2404.
- GeoLDM: github.com/MinkaiXu/GeoLDM
- Xu, M., Powers, A. S., Dror, R. O., Ermon, S., & Leskovec, J. (2023, July). Geometric latent diffusion models for 3d molecule generation. In *International Conference on Machine Learning*(pp. 38592-38610). PMLR.