Energy-Inspired Molecular Conformation Optimization

ICLR 2022

Jiaqi Guan[1], Wesley Wei Qian[1], Qiang Liu[2], Wei-Ying Ma[3], Jianzhu Ma[4], Jian Peng[1,3,5]

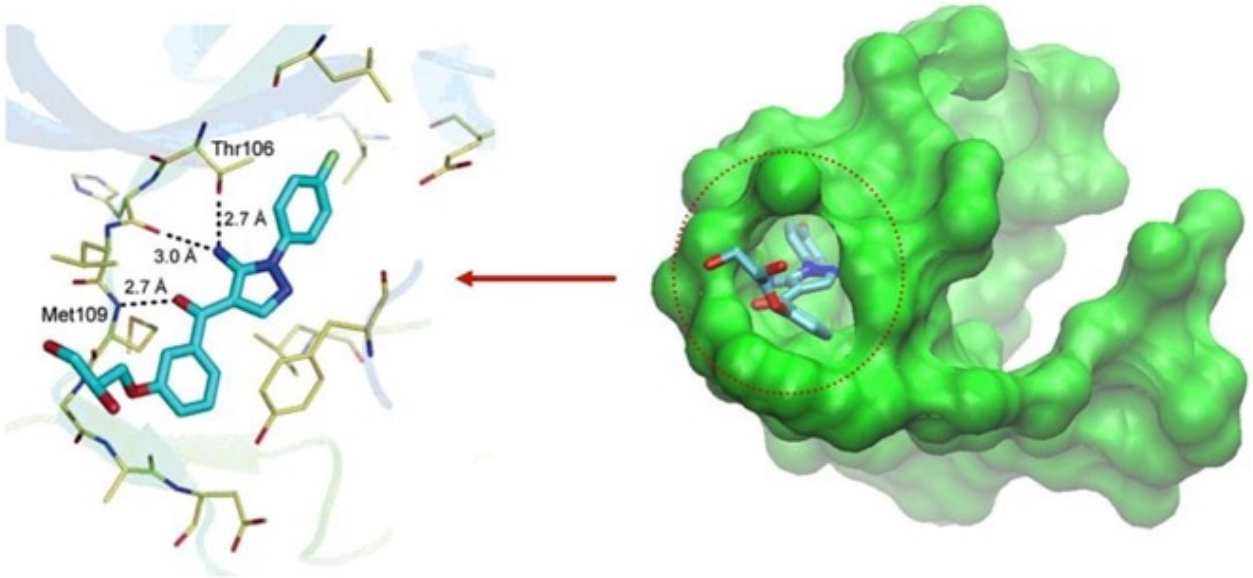
1. University of Illinois Urbana Champaign 2. University of Texas at Austin

3. AIR, Tsinghua University 4. Peking University 5. HeliXon Limited

Code: https://github.com/guanjq/confopt_official

Background

- Molecular conformer determines the steric and electronic properties of the molecules
- Efficiently and accurately generate molecular conformation can unlock lots of new opportunities including screening for protein targets



Previous Work

- Molecular Dynamics Simulation
  - Accuracy and efficiency trade-off
  - DFT based / FF based methods

- Two-Stage Deep Generative Models (First generate pairwise distances)
  - GraphDG (VAE based) [1]
  - CGCF (Continuous Normalizing Flow based) [2]

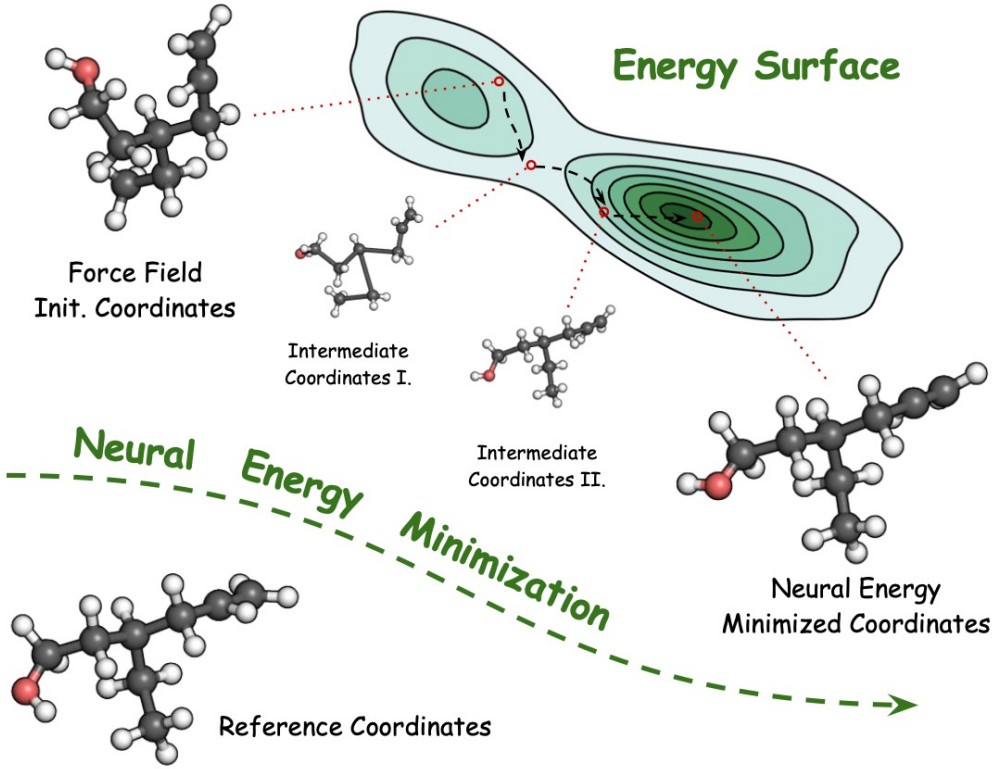
- One-Stage Deep Generative Models
  - ConfGF (Score-Matching based) [3]

1. Mansimov, E., Mahmood, O., Kang, S., & Cho, K. (2019). Molecular geometry prediction using a deep generative graph neural network. *Scientific reports*.
2. Xu, M., Luo, S., Bengio, Y., Peng, J., & Tang, J. (2021). Learning neural generative dynamics for molecular conformation generation. *ICLR 2021.*
3. Shi, C., Luo, S., Xu, M., & Tang, J. (2021). Learning gradient fields for molecular conformation generation. *ICML 2021.*

Motivation

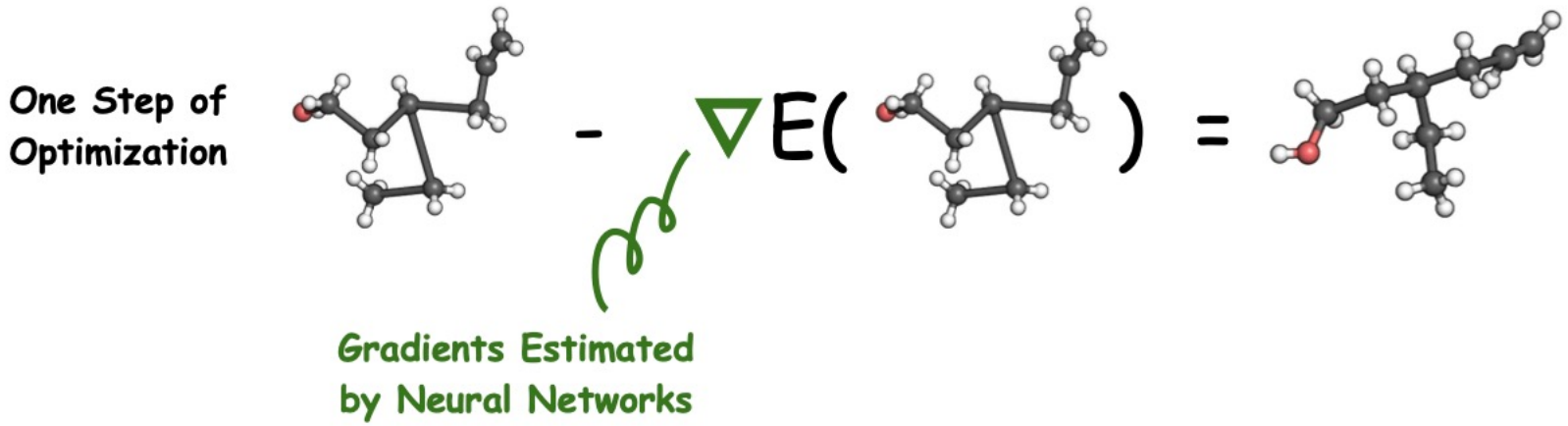
- Nature has an energy function for conformer
- The energy function governs an energy surface
- Stable equilibrium conformer corresponds a local minimum on the potential energy surface

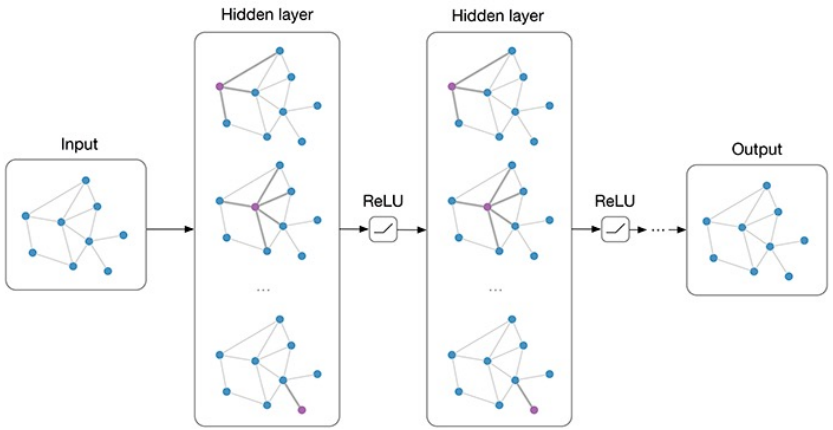
To get a conformer with lowest energy, we can navigate the energy surface (perform gradient descent) by updating the 3D coordinates



Motivation

Use a graph neural network to learn the gradient field of conformational energy from data





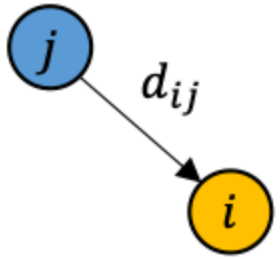
Should be 3D rotational and translational equivariant

→ Our models derived from the energy minimization perspective will be naturally SE(3)-equivariant

Recipe For Deriving Energy-Inspired Models

- Assume an underlying energy function

I. Energy Function

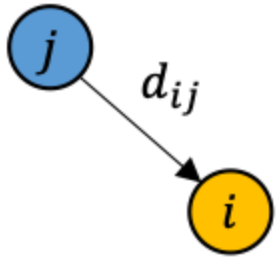


$$E(\mathbf{x}; G) = \sum_{i \in \mathcal{V}} \sum_{j \neq i} \phi(d_{ij}; G)$$

Recipe For Deriving Energy-Inspired Models

- Assume an underlying energy function
- Take gradients of the energy function w.r.t conformation $\mathbf{x}$

I. Energy Function



$$E(\mathbf{x}; G) = \sum_{i \in \mathcal{V}} \sum_{j \neq i} \phi(d_{ij}; G)$$

II. Energy Gradients w.r.t x

$$f_\theta(\mathbf{x}; G)|_{x_i} = \frac{\partial E(\mathbf{x}; G)}{\partial x_i} = \sum_{j \neq i} \frac{\partial \phi(d_{ij}; G)}{\partial x_i}$$

$$= \sum_{j \neq i} \frac{\partial d_{ij}}{\partial x_i} \frac{\partial \phi(d_{ij}; G)}{\partial d_{ij}}$$

$$= \sum_{j \neq i} \frac{(x_i - x_j)}{d_{ij}} \frac{\partial \phi(d_{ij}; G)}{\partial d_{ij}}$$

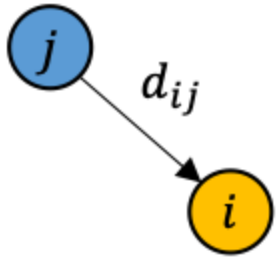
$$= \sum_{j \neq i} (x_i - x_j) g_\theta(d_{ij}; G)$$

Recipe For Deriving Energy-Inspired Models

- Assume an underlying energy function
- Take gradients of the energy function w.r.t conformation $\mathbf{x}$

→ Obtain a SE(3)-equivariant network naturally

I. Energy Function



$$E(\mathbf{x}; G) = \sum_{i \in \mathcal{V}} \sum_{j \neq i} \phi(d_{ij}; G)$$

II. Energy Gradients w.r.t x

$$f_\theta(\mathbf{x}; G)|_{x_i} = \frac{\partial E(\mathbf{x}; G)}{\partial x_i} = \sum_{j \neq i} \frac{\partial \phi(d_{ij}; G)}{\partial x_i}$$

$$= \sum_{j \neq i} \frac{\partial d_{ij}}{\partial x_i} \frac{\partial \phi(d_{ij}; G)}{\partial d_{ij}}$$

$$= \sum_{j \neq i} \frac{(x_i - x_j)}{d_{ij}} \frac{\partial \phi(d_{ij}; G)}{\partial d_{ij}}$$

$$= \sum_{j \neq i} (x_i - x_j) g_\theta(d_{ij}; G)$$

Result: a SE(3)-Equivariant model

Each update is a SE(3)-equivariant operation

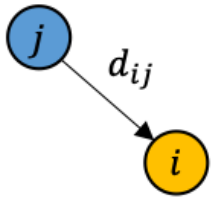
Proof:

Assume input x is $\hat{x} = Rx + t$ and translate with t :

$$\begin{aligned} \hat{x}^{t+1} &= \hat{x}^t - f_\theta(\hat{x}^t; G) \\ &= Rx^t + t - Rf_\theta(x^t; G) \\ &= R(x^t - f_\theta(x^t; G)) + t \\ &= Rx^{t+1} + t \end{aligned}$$

More Variants

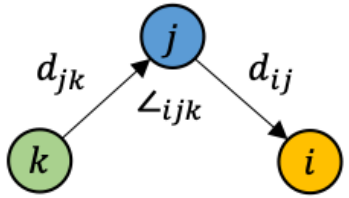
Two-Atom Model



$$E(\mathbf{x}; G) = \sum_{i \in \mathcal{V}} \sum_{j \neq i} \phi(d_{ij}; G)$$

$$x_i^{l+1} = x_i^l + \sum_{j \neq i} (x_i^l - x_j^l) f_{h \to x}(\| x_i^l - x_j^l \|, h_i^l, h_j^l; e_{ij})$$

Three-Atom Model



$$E(\mathbf{x}; G) = \sum_{i \in \mathcal{V}} \sum_{j \neq ik \neq i,j} \sum \phi(d_{ij}, d_{ik}, \angle_{jik}; G)$$

$$x_i^{l+1} = x_i^l + \sum_{jk} (x_i - x_j) g_\psi(d_{ij}, d_{ik}, r_{ij}^T r_{ik}) + (x_i - x_k) h_\omega(d_{ij}, d_{ik}, r_{ij}^T r_{ik})$$

Benefits:

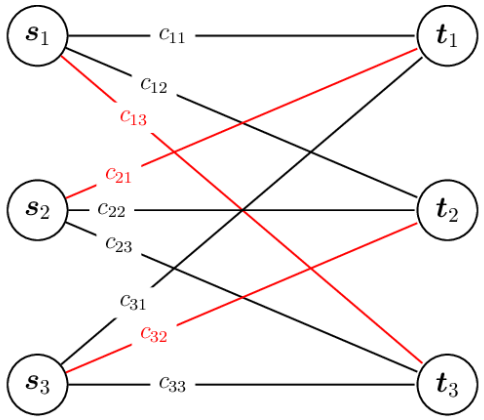
- Instead of proposing a model heuristically based on MPNN (like EGNN), the derived model is SE(3)-equivariant naturally and aligns the underlying physical assumption
- Avoid complicated spherical-harmonics-based techniques, like SE(3)-Transformer
- One-stage; Faster sampling compared to ConfGF

Experiments – Setup

- Molecular Conformation Optimization
  - L2 Loss on pairwise distance matrix
  - Datasets: QM9 and lowest energy conformer in GEOM-Drug
  - Baselines: RDKit, EGNN, SE(3)-Transformer
  - Metrics: mean / median RMSD

- Molecular Conformation Generation
  - Apply the optimal transport loss

$$L_{OT} = \min_{\pi \in \Gamma} \sum_{i,j} \pi_{i,j} \mathbf{C}(\boldsymbol{X}_i^*, \boldsymbol{X}_j)$$



  - Datasets: GEOM-QM9 and GEOM-Drug (multiple conformers)
  - Baselines: RDKit, multiple deep generative models
  - Metrics: Coverage rate (COV), Mismatching rate (MIS), Matching score (MAT)

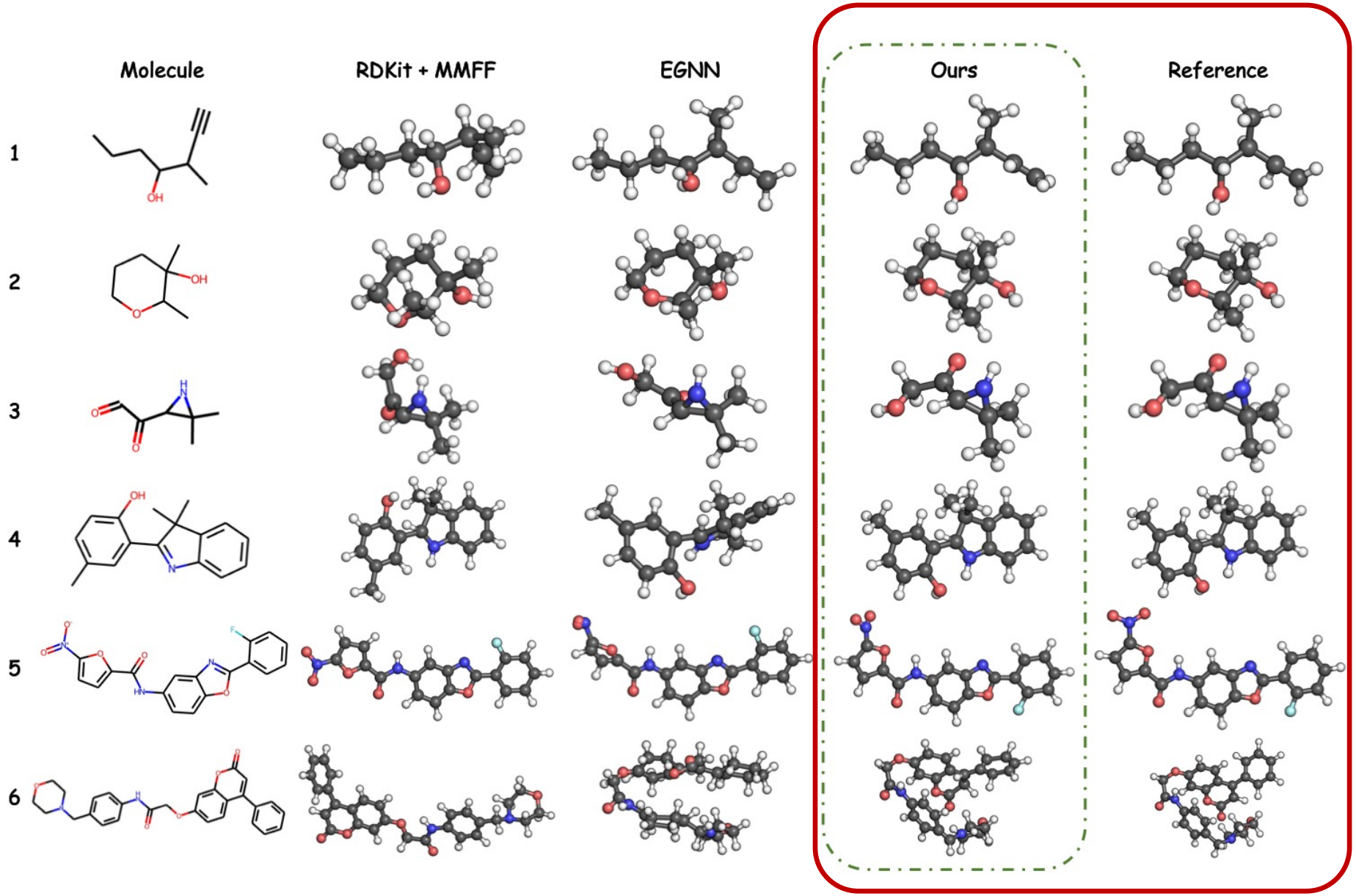
Experiments – Optimization Results

Molecular Conformation Optimization

- Better RMSD

Model \ Dataset	QM9		GEOM-Drug	
	mean RMSD (↓)	median RMSD (↓)	mean RMSD (↓)	median RMSD (↓)
RDKit+MMFF	0.3872 ± 0.0029	0.2756 ± 0.0075	1.7913 ± 0.0030	1.6433 ± 0.0097
SE(3)-Tr. [8]	0.2476 ± 0.0021	0.1657 ± 0.0022	1.0050 ± 0.0022	0.9139 ± 0.0041
EGNN [34]	0.2101 ± 0.0009	0.1356 ± 0.0013	1.0405 ± 0.0018	0.9598 ± 0.0038
Ours-TwoAtom	0.1415 ± 0.0004	0.0534 ± 0.0002	0.8839 ± 0.0014	0.7733 ± 0.0026
Ours-Ext$_v$	0.1383 ± 0.0005	**0.0505** ± 0.0001	0.8691 ± 0.0015	0.7535 ± 0.0028
Ours-ThreeAtom	**0.1374** ± 0.0004	0.0522 ± 0.0002	**0.8567** ± 0.0014	**0.7192** ± 0.0024

Experiments – Visualization



Experiments – Generation Results

Molecular Conformation Generation

- Evaluate with different initialization
- Better precision and recall

Dataset	GEOM-QM9						GEOM-Drugs					
Metric	COV(%) (↑)		MIS(%) (↓)		MAT(Å) (↓)		COV(%) (↑)		MIS(%) (↓)		MAT(Å) (↓)	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
RDKit	83.26	90.78	8.13	1.00	0.3447	0.2935	60.91	65.70	27.95	12.07	1.2026	1.1252
CVGAE	0.09	0.00	-	-	1.6713	1.6088	0.00	0.00	-	-	3.0702	2.9937
GraphDG	73.33	84.21	56.09	64.66	0.4245	0.3973	8.27	0.00	97.92	100.00	1.9722	1.9845
CGCF	78.05	82.48	63.51	64.66	0.4219	0.3900	53.96	57.06	78.32	86.28	1.2487	1.2247
ConfGF	88.49	**94.13**	53.56	56.59	0.2673	0.2685	62.15	70.93	76.58	84.48	1.1629	1.1596
Ours-Random	**88.83**	93.18	30.21	30.74	0.3778	0.3736	**76.50**	**83.78**	31.40	23.03	**1.0694**	1.0583
Ours-RDKit	86.68	91.34	**5.46**	**0.00**	**0.2667**	**0.2125**	67.72	75.30	**21.52**	**2.66**	1.0739	**1.0372**

Thanks for watching!

Energy-Inspired Molecular Conformation Optimization

ICLR 2022

- Jiaqi Guan[1], Wesley Wei Qian[1], Qiang Liu[2], Wei-Ying Ma[3], Jianzhu Ma[4], Jian Peng[1,3,5]

1. University of Illinois Urbana Champaign 2. University of Texas at Austin
- 3. AIR, Tsinghua University 4. Peking University 5. HeliXon Limited