

## APPENDIX

## A EQUIVARIANCE AND INVARIANCE IN GEOMETRIC OPTIMAL TRANSPORT

**Equivariance.** Molecules, typically existing within a three-dimensional physical space, are subject to geometric symmetries, including translations, rotations, and potential reflections. These are collectively referred to as the Euclidean group in 3 dimensions, denoted as  $E(3)$  (Celeghini et al., 1991).

A function  $F$  is said to be equivariant to the action of a group  $G$  if  $T_g \circ F(\mathbf{x}) = F \circ S_g(\mathbf{x})$  for all  $g \in G$ , where  $S_g, T_g$  are linear representations related to the group element  $g$  (Serre et al., 1977).

**Invariance.** A function  $F$  is said to be invariant to the action of a group  $G$  if  $F \circ \pi(\mathbf{x}) = F(\mathbf{x})$  for all  $g \in G$  and every permutation  $\pi \in S_n$ .

**Equivariance and Invariance in Molecules.** For geometric graph generation, we consider the special Euclidean group  $SE(3)$ , involving translations and rotations. Moreover, the transformations  $S_g$  or  $T_g$  can be represented by a translation  $\mathbf{t}$  and an orthogonal matrix rotation  $\mathbf{R}$ .

For a molecule  $\mathbf{g} = \langle \mathbf{x}, \mathbf{h} \rangle$ , the node features  $\mathbf{h}$  are  $SE(3)$ -invariant while the coordinates  $\mathbf{x}$  are  $SE(3)$ -equivariant, which can be expressed as  $\mathbf{R}\mathbf{x} + \mathbf{t} = (\mathbf{R}\mathbf{x}_1 + \mathbf{t}, \dots, \mathbf{R}\mathbf{x}_N + \mathbf{t})$ .

**Equivariance and Invariance in Geometric Optimal Transport.** For non-topological data, such as images, the transport cost between two given data points is fixed. However, this does not apply to topological graphs. For instance, when a topological graph (molecule) undergoes rotation or translation, the inherent properties of the molecule remain unchanged, but the cost of transporting coordinates may vary. Similarly, if the atom order in one of the molecules changes in silico, the molecule remains constant, but the transport cost of coordinates and features may alter. Therefore, the proposed optimal molecule transport problem aims to find an optimal rotation, translation, and permutation transformation for one molecule to minimize the distance, considering both coordinates and features, from another molecule.

## B PROOF FOR THEOREM 3.1

The theorem 3.1 is reproduced here for convenience:

**Theorem 3.1** *The coupling  $\hat{\Gamma}$  incurs no larger geometric transport costs than the arbitrary coupling  $\Gamma(p_0, p_1)$  in that  $\mathbb{E}[\hat{c}_g(\mathbf{z}_0, \mathbf{z}'_1)] \leq \mathbb{E}[\hat{c}_g(\mathbf{z}_0, \mathbf{z}_1)]$  where  $(\mathbf{z}_0, \mathbf{z}'_1) \in \hat{\Gamma}(p_0, p'_1)$ ,  $(\mathbf{z}_0, \mathbf{z}_1) \in \Gamma(p_0, p_1)$ , and  $\hat{c}_g(\mathbf{z}_0, \mathbf{z}_1) = \min \|\pi(\mathbf{R}\mathbf{z}_{\mathbf{x},1}^1 + \mathbf{t}, \mathbf{R}\mathbf{z}_{\mathbf{x},1}^2 + \mathbf{t}, \dots, \mathbf{R}\mathbf{z}_{\mathbf{x},1}^N + \mathbf{t}) - (\mathbf{z}_{\mathbf{x},0}^1, \mathbf{z}_{\mathbf{x},0}^2, \dots, \mathbf{z}_{\mathbf{x},0}^N)\|_2 + \min \|\pi(\mathbf{z}_{\mathbf{h},1}^1, \mathbf{z}_{\mathbf{h},1}^2, \dots, \mathbf{z}_{\mathbf{h},1}^N) - (\mathbf{z}_{\mathbf{h},0}^1, \mathbf{z}_{\mathbf{h},0}^2, \dots, \mathbf{z}_{\mathbf{h},0}^N)\|_2, \forall \pi, \mathbf{R}$ , and  $\mathbf{t}$ .*

$\mathbf{z}$  is geometry  $\mathbf{g}$  in the latent space, which is composed of  $\mathbf{z}_{\mathbf{x}} \in \mathbb{R}^{N \times 3}$  and  $\mathbf{z}_{\mathbf{h}} \in \mathbb{R}^{N \times k}$ , where  $k$  is the latent dimension characterized by  $\mathcal{E}_\phi$ .

With node-granular optimal transport  $\hat{\mathbf{R}}, \hat{\mathbf{t}}$  and  $\hat{\pi}$  we have:

$$\begin{aligned} \mathbb{E}[\hat{c}_g(\mathbf{z}_0, \mathbf{z}'_1)] &= \mathbb{E}[\min \|\pi(\mathbf{R}\mathbf{z}_{\mathbf{x},1}^1 + \mathbf{t}, \mathbf{R}\mathbf{z}_{\mathbf{x},1}^2 + \mathbf{t}, \dots, \mathbf{R}\mathbf{z}_{\mathbf{x},1}^N + \mathbf{t}) - (\mathbf{z}_{\mathbf{x},0}^1, \mathbf{z}_{\mathbf{x},0}^2, \dots, \mathbf{z}_{\mathbf{x},0}^N)\|_2 \\ &\quad + \min \|\pi(\mathbf{z}_{\mathbf{h},1}^1, \mathbf{z}_{\mathbf{h},1}^2, \dots, \mathbf{z}_{\mathbf{h},1}^N) - (\mathbf{z}_{\mathbf{h},0}^1, \mathbf{z}_{\mathbf{h},0}^2, \dots, \mathbf{z}_{\mathbf{h},0}^N)\|_2, \forall \pi, \mathbf{R}, \text{ and } \mathbf{t}] \\ &= \mathbb{E}[\|\hat{\pi}(\hat{\mathbf{R}}\mathbf{z}_{\mathbf{x},1}^1 + \hat{\mathbf{t}}, \hat{\mathbf{R}}\mathbf{z}_{\mathbf{x},1}^2 + \hat{\mathbf{t}}, \dots, \hat{\mathbf{R}}\mathbf{z}_{\mathbf{x},1}^N + \hat{\mathbf{t}}) - (\mathbf{z}_{\mathbf{x},0}^1, \mathbf{z}_{\mathbf{x},0}^2, \dots, \mathbf{z}_{\mathbf{x},0}^N)\|_2 \\ &\quad + \|\hat{\pi}(\mathbf{z}_{\mathbf{h},1}^1, \mathbf{z}_{\mathbf{h},1}^2, \dots, \mathbf{z}_{\mathbf{h},1}^N) - (\mathbf{z}_{\mathbf{h},0}^1, \mathbf{z}_{\mathbf{h},0}^2, \dots, \mathbf{z}_{\mathbf{h},0}^N)\|_2] \end{aligned}$$

Let  $\hat{\mathbf{z}}_{\mathbf{x}} = \hat{\pi}(\hat{\mathbf{R}}\mathbf{z}_{\mathbf{x}}^1 + \hat{\mathbf{t}}, \hat{\mathbf{R}}\mathbf{z}_{\mathbf{x}}^2 + \hat{\mathbf{t}}, \dots, \hat{\mathbf{R}}\mathbf{z}_{\mathbf{x}}^N + \hat{\mathbf{t}})$ ,  $\hat{\mathbf{z}}_{\mathbf{h}} = \hat{\pi}(\mathbf{z}_{\mathbf{h}}^1, \mathbf{z}_{\mathbf{h}}^2, \dots, \mathbf{z}_{\mathbf{h}}^N)$ , and  $\hat{\mathbf{z}} = [\hat{\mathbf{z}}_{\mathbf{x}}, \hat{\mathbf{z}}_{\mathbf{h}}] \in \mathbb{R}^{N \times (3+k)}$ , then we have:

$$\begin{aligned} \mathbb{E}[\hat{c}_g(\mathbf{z}_0, \mathbf{z}'_1)] &= \mathbb{E}[\|(\hat{\mathbf{z}}_{\mathbf{x},1}^1 + \hat{\mathbf{z}}_{\mathbf{x},1}^2 + \dots, \hat{\mathbf{z}}_{\mathbf{x},1}^N) - (\mathbf{z}_{\mathbf{x},0}^1, \mathbf{z}_{\mathbf{x},0}^2, \dots, \mathbf{z}_{\mathbf{x},0}^N)\|_2 \\ &\quad + \|(\hat{\mathbf{z}}_{\mathbf{h},1}^1, \hat{\mathbf{z}}_{\mathbf{h},1}^2, \dots, \hat{\mathbf{z}}_{\mathbf{h},1}^N) - (\mathbf{z}_{\mathbf{h},0}^1, \mathbf{z}_{\mathbf{h},0}^2, \dots, \mathbf{z}_{\mathbf{h},0}^N)\|_2] \\ &= \mathbb{E}[\|(\hat{\mathbf{z}}_1^1 + \hat{\mathbf{z}}_1^2 + \dots, \hat{\mathbf{z}}_1^N) - (\mathbf{z}_0^1, \mathbf{z}_0^2, \dots, \mathbf{z}_0^N)\|_2] \\ &= \mathbb{E}[\|\hat{\mathbf{z}}_1 - \mathbf{z}_0\|_2]. \end{aligned}$$

Likewise, we have:

$$\mathbb{E}[\hat{c}_g(\mathbf{z}_0, \mathbf{z}_1)] = \mathbb{E}[\|\hat{\mathbf{z}}_1 - \mathbf{z}_0\|_2]. \quad (9)$$

At this point, what we aim to prove is simplified to:

$$\mathbb{E}[\|\hat{\mathbf{z}}'_1 - \mathbf{z}_0\|_2] \leq \mathbb{E}[\|\hat{\mathbf{z}}_1 - \mathbf{z}_0\|_2] \quad (10)$$

*Proof.* Given that  $\mathbf{z}'_1 = \text{ODE}_{\hat{\theta}}(\mathbf{z}_0)$ ,  $d\mathbf{z}_t = v_{\hat{\theta}}(\mathbf{z}_t, t)dt$ , we have:

$$\mathbb{E}[\hat{c}_g(\mathbf{z}_0, \mathbf{z}'_1)] = \mathbb{E} \left[ \left\| \int_0^1 v_{\hat{\theta}}(\mathbf{z}_t, t) dt \right\|_2 \right] \quad (11)$$

$\|\cdot\|_2 : \mathbb{R}^{N \times (3+k)} \rightarrow \mathbb{R}_+$  is the Euclidean norm of  $\cdot$  and it is convex, therefore, with  $\|\int_{\Omega} v dt\| \leq \int_{\Omega} \|v\| dt$  induced by Jensen's inequality we have:

$$\mathbb{E}[\hat{c}_g(\mathbf{z}_0, \mathbf{z}'_1)] \leq \mathbb{E} \left[ \int_0^1 \|v_{\hat{\theta}}(\mathbf{z}_t, t)\|_2 dt \right]. \quad (12)$$

With defined  $v_{\hat{\theta}}(\mathbf{z}_t, t) = \mathbb{E}[\mathbf{z}_1 - \mathbf{z}_0 | \mathbf{z}_t]$ , we then have:

$$\mathbb{E}[\hat{c}_g(\mathbf{z}_0, \mathbf{z}'_1)] = \mathbb{E} \left[ \int_0^1 \|\mathbb{E}[\mathbf{z}_1 - \mathbf{z}_0 | \mathbf{z}_t]\|_2 dt \right]. \quad (13)$$

Again, with the finite form of Jensen's inequality, we have:

$$\begin{aligned} \mathbb{E}[\hat{c}_g(\mathbf{z}_0, \mathbf{z}'_1)] &\leq \mathbb{E} \left[ \int_0^1 \mathbb{E}[\|\mathbf{z}_1 - \mathbf{z}_0\|_2 | \mathbf{z}_t] dt \right] \quad // \text{Jensen's inequality} \\ &= \int_0^1 \mathbb{E}[\mathbb{E}[\|\mathbf{z}_1 - \mathbf{z}_0\|_2 | \mathbf{z}_t]] dt \\ &= \int_0^1 \mathbb{E}[\|\mathbf{z}_1 - \mathbf{z}_0\|_2] dt \quad // \mathbb{E}[\|\mathbf{z}_1 - \mathbf{z}_0\|_2 | \mathbf{z}_t] = \|\mathbf{z}_1 - \mathbf{z}_0\|_2 \\ &= \mathbb{E}[\|\hat{\mathbf{z}}_1 - \hat{\mathbf{z}}_0\|_2] \\ &= \mathbb{E}[\hat{c}_g(\mathbf{z}_0, \mathbf{z}_1)] \quad // \text{By Eq. 9} \end{aligned} \quad (14)$$

Combining equations 11 to 14, Eq. 10 is proved.  $\square$

It is important to note that solving the geometric optimal transport problem in the latent space does not necessarily ensure that the molecule itself or its distribution also satisfies the optimal transport in the original space. However, given that the proposed flow model is trained in the latent space, it is sufficient to ensure that latent molecules and distributions are transported with optimal cost, thereby accelerating the flow model in the generation of molecules.

## C ALGORITHMS

This section contains the main algorithms of the proposed GOAT. First, we present the algorithm for solving optimal molecule transport and unified flow in Algorithm 1 and Algorithm 2, respectively. Algorithm 3 presents the pseudo-code for training the GOAT. Algorithm 4 presents the process of fast molecule generation with GOAT.

**Algorithm 1** Optimal Molecule Transport

---

```

1: Input:  $\mathbf{z}_1$  and  $\mathbf{z}_0$ .
2: Output:  $\hat{\mathbf{z}}_1$  and  $\mathbf{z}_0$ .
3: Optimal Molecule Transport:
4:  $M_{c_g}[i, j] \leftarrow \|\mathbf{z}_1^i - \mathbf{z}_0^j\|^2 \leftarrow \|\mathbf{z}_{\mathbf{x},1}^i - \mathbf{z}_{\mathbf{x},0}^j\|^2 + \|\mathbf{z}_{\mathbf{h},1}^i - \mathbf{z}_{\mathbf{h},0}^j\|^2$  // Construct Atom-level Transport
   Cost Matrix
5:  $\hat{\pi} \leftarrow$  Hungarian algorithm (Kuhn, 1955) // Optimal Permutation
6:  $\hat{\mathbf{R}} \leftarrow$  Kabsch algorithm (Kabsch, 1976) // Optimal Rotation
7:  $\hat{\mathbf{z}}_1 = \pi(\hat{\mathbf{R}}\mathbf{z}_1)$  // Optimal Molecule Transport
8: return  $\hat{\mathbf{z}}_1, \mathbf{z}_0$ 

```

---

**Algorithm 2** Equivariant Autoencoder

---

```

1: Input: geometric data point  $\mathbf{g} = \langle \mathbf{x}, \mathbf{h} \rangle$ , equivariant encoder  $\mathcal{E}_\phi$ 
2: Output: encoded data point  $\mathbf{z}$ 
3: Unified Flow:
4:  $\mathbf{x} \leftarrow \mathbf{x} - \mathbf{G}(\mathbf{x})$  // Translate to CoM Space
5:  $\mu_{\mathbf{x}}, \mu_{\mathbf{h}} \leftarrow \mathcal{E}_\phi(\mathbf{x}, \mathbf{h})$  // Encode
6:  $\langle \epsilon_{\mathbf{x}}, \epsilon_{\mathbf{h}} \rangle \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  // Sample noise for Equivariant Autoencoder
7:  $\epsilon_{\mathbf{x}} \leftarrow \epsilon_{\mathbf{x}} - \mathbf{G}(\epsilon_{\mathbf{x}})$  // Translate to CoM Space
8:  $\mathbf{z}_{\mathbf{x}}, \mathbf{z}_{\mathbf{h}} \leftarrow \mu + \langle \epsilon_{\mathbf{x}}, \epsilon_{\mathbf{h}} \rangle \odot \sigma_0$  // Obtain Latent Representation
9:  $\mathbf{z} \leftarrow [\mathbf{z}_{\mathbf{x}}, \mathbf{z}_{\mathbf{h}}]$ 
10: return  $\mathbf{z}$ 

```

---

**Algorithm 3** Geometric Optimal Transport

---

```

1: Input: data distribution  $p_1$ , equivariant encoder  $\mathcal{E}_\phi$ , decoder  $\mathcal{D}_\epsilon$ , flow network  $v_\theta$ 
2: Output: GOAT:  $(\hat{v}_\theta)$ 
3: for  $\mathbf{g}_1 = \langle \mathbf{x}, \mathbf{h} \rangle \sim p_1$  do
4:    $\mathbf{z}_1 \leftarrow$  Equivariant Autoencoder( $\mathbf{g}_1$ ) // Algorithm 2
5:    $\mathbf{z}_0 \leftarrow \langle \mathbf{z}_{\mathbf{x},0}, \mathbf{z}_{\mathbf{h},0} \rangle \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  // Sample noise from base distribution  $p_0$ 
6:    $\hat{\mathbf{z}}_1, \mathbf{z}_0 =$  Optimal Molecule Transport ( $\mathbf{z}_1, \mathbf{z}_0$ ) // Algorithm 1
7:    $\mathcal{L}_{F1}(\theta) = \mathbb{E}_{t,p_0,p_1} \|v_\theta(\hat{\mathbf{z}}_t, t) - (\hat{\mathbf{z}}_1 - \mathbf{z}_0)\|^2$  // Loss for the flow
8:    $\hat{\theta} \leftarrow$  optimizer( $\mathcal{L}_F, \theta$ ) // Optimize
9: end for
10: for  $\mathbf{g}_1 = \langle \mathbf{x}, \mathbf{h} \rangle \sim p_1$  do
11:    $\mathbf{z}_0, \mathbf{z}'_1, \mathbf{g}'_1 \leftarrow$  Sampling( $\mathcal{D}_\epsilon, \hat{\theta}$ ) // Algorithm 4
12:   if  $\mathbf{g}'_1$  meets quality (measure by RdKit (Landrum et al., 2016)) then
13:      $\hat{\mathbf{z}}'_1, \mathbf{z}_0 =$  Optimal Molecule Transport ( $\mathbf{z}'_1, \mathbf{z}_0$ ) // Algorithm 1
14:      $\mathcal{L}_{F1}(\theta) = \mathbb{E}_{t,p_0,p_1} \|v_\theta(\hat{\mathbf{z}}'_t, t) - (\hat{\mathbf{z}}'_1 - \mathbf{z}_0)\|^2$  // Loss for the flow
15:      $\hat{\theta} \leftarrow$  optimizer( $\mathcal{L}_F, \theta$ ) // Optimize
16:   end if
17: end for
18: return  $\hat{\theta}$ 

```

---

**Algorithm 4** Sampling

---

```

1: Input: equivariant decoder  $\mathcal{D}_\epsilon$ , flow network  $\theta$ .
2: Output: noise:  $\mathbf{z}_0$ , generated latent sample:  $\mathbf{z}'_1$ , generated molecule:  $\mathbf{g}'_1$ .
3:  $\mathbf{z}_0 \leftarrow \langle \mathbf{z}_{\mathbf{x},0}, \mathbf{z}_{\mathbf{h},0} \rangle \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  // Sample noise from base distribution  $p_0$ 
4:  $\mathbf{z}'_1 \leftarrow \text{ODE}_{v_\theta}(\mathbf{z}_0)$ 
5:  $\mathbf{g}'_1 \leftarrow \mathcal{D}_\epsilon(\mathbf{z}'_1)$  // Solve ODE
6: return  $\mathbf{z}_0, \mathbf{z}'_1, \mathbf{g}'_1$ 

```

---

## D RELATED WORKS

**Molecule Generation Models.** Initial research in molecule generation primarily concentrated on the creation of molecules as 2D graphs (Jin et al., 2018; Liu et al., 2018; Shi et al., 2019). However, the field has seen a shift in interest towards 3D molecule generation. Techniques such as G-SchNet (Gebauer et al., 2019) and G-SphereNet (Luo & Ji, 2022) employ autoregressive methods to incrementally construct molecules by progressively linking atoms or molecular fragments. These approaches necessitate either a detailed formulation of a complex action space or an ordering of actions.

Motivated by the success of Diffusion Models (DMs) in image generation, the focus has now turned to their application in 3D molecule generation from noise (Hoogeboom et al., 2022; Xu et al., 2023; Wu et al., 2022; Han et al., 2023). To address the inconsistency of unified Gaussian diffusion across various modalities, a latent space was introduced by (Xu et al., 2023). To resolve the atom-bond inconsistency issue, (Peng et al., 2023) proposed different noise schedulers for different modalities to accommodate noise sensitivity. However, diffusion-based models consistently face the challenge of slow sampling speed, resulting in a significant computational burden for generation. To enhance the speed, recent proposals have introduced flow matching-based (Song et al., 2023a) and Bayesian flow network-based (Song et al., 2023b) models. Despite these advancements, there remains substantial potential for improvement in these frameworks regarding speed, novelty, and ultimate significance.

**Flow Models.** Introduced in (Chen et al., 2018), Continuous Normalizing Flows (CNFs) represent a continuous-time variant of Normalizing Flows (Rezende & Mohamed, 2015). Subsequently, flow matching (Lipman et al., 2022) and rectified flow (Liu et al., 2022) were proposed to circumvent the need for ODE simulations during forward and backward propagation in CNF, and they introduced optimal transport for faster generation. Leveraging these advanced flow models, (Garcia Satorras et al., 2021) pioneered the use of flow models for molecule generation, which was later followed by the proposal of (Song et al., 2023a), based on hybrid transport. Beyond the realm of 3D molecule generation, the concept of flow matching and optimal transport has also found applications in many-body systems (Garcia Satorras et al., 2021) and molecule simulations (Midgley et al., 2023). Despite these advancements, existing models primarily focus on atomic coordinates, leaving the challenge of geometric optimal transport unresolved.

## E DATASET

### E.1 QM9 DATASET

QM9 (Ramakrishnan et al., 2014) is a comprehensive dataset that provides geometric, energetic, electronic, and thermodynamic properties for a subset of the GDB-17 database (Ruddigkeit et al., 2012) comprises a total of 130,831 molecules<sup>2</sup>. We utilize the train/validation/test partitions delineated in (Anderson et al., 2019), comprising 100K, 18K, and 13K samples for each respective partition.

### E.2 GEOM-DRUG DATASET

GEOM-DRUG (Geometric Ensemble Of Molecules) dataset (Axelrod & Gómez-Bombarelli, 2022) encompasses around 450,000 molecules, each with an average of 44.2 atoms and a maximum of 181 atoms<sup>3</sup>. We build the GEOM-DRUG dataset following (Hoogeboom et al., 2022) with the provided code.

## F IMPLEMENTATION DETAILS

In this study, all the neural networks utilized for the encoder, flow network, and decoder are implemented using EGNNs (Satorras et al., 2021). The dimension of latent invariant features, denoted as  $k$ , is set to 2 for QM9 and 1 for GEOM-DRUG, to map the molecule for a unified flow matching.

<sup>2</sup><https://springernature.figshare.com/ndownloader/files/3195389>

<sup>3</sup><https://dataverse.harvard.edu/file.xhtml?fileId=4360331&version=2.0>

For the training of the flow neural network, we employ EGNNs with 9 layers and 256 hidden features on QM9, and 4 layers and 256 hidden features on GEOM-DRUG, with a batch size of 64 and 16, respectively.

In the case of equivariant autoencoders, the decoder is parameterized in the same manner as the encoder, but the encoder is implemented with a 1-layer EGNN. This shallow encoder effectively constrains the encoding capacity and aids in regularizing the latent space (Xu et al., 2023).

All models utilize SiLU activations and are trained until convergence. Across all experiments, the Adam optimizer (Kingma & Ba, 2015) with a constant learning rate of  $10^{-4}$  is chosen as our default training configuration. The training process for QM9 takes approximately 3000 epochs, while for GEOM-DRUG, it takes about 20 epochs.

With the flow model trained on QM9 or GEOM-DRUG, we then generate and purify the coupling to obtain a total of 100K molecular pairs, which form the estimated couplings.

### Hardware Configuration

1. GPU: NVIDIA GeForce RTX 3090
2. CPU: Intel(R) Xeon(R) Platinum 8338C CPU
3. Memory: 512 GB
4. Time: Around 7 days for QM9 and 20 days for GEOM-DRUG.

## G MORE EXPERIMENTAL RESULTS

We present the full results in Tables 5 and 6. In our detailed experimental results on QM9, we reproduced EDM, GeoLDM, and EquiFM on the QM9 dataset to obtain the actual generation time consumption with the same compute configuration. As a result, the proposed method achieves the fastest sampling speed, which is consistent with the measurement of sampling steps. We also witness a huge generation speed improvement by the proposed GOAT for GEOM-DRUG.

In addition to supplementing the actual time used for generation, we also added the metrics of molecule stability, and it is obvious that all methods achieve nearly 0% molecule stability in GEOM-DRUG. This is because metrics, atom and molecule stability, create errors during bond type prediction based on pair-wise atom types and distances. Therefore, we concentrate on metrics measured by RdKit.

Lastly, we produced the full results of GeoBFN using sampling steps from 50 to 1,000. It is worth noting that the novelty and significance continue to decrease on QM9 datasets as sampling steps increase, which aligns with our conjecture in the experiments. Besides, we also observed that its performance on GEOM-DRUG also decreased in terms of validity. Combined with its efficiency and quality, we believe that our method GOAT has competitive performance compared with GeoBFN.

We present the visualization of generated molecules on QM9 and GEOM-DRUG in Figures 5 and 6.

Table 5: Comparisons of generation quality (larger is better) in terms of Atom Stability, Molecule Stability, Validity, Uniqueness, Novelty, and Significance. And comparisons of generation efficiency regarding generation time and sampling steps for one molecule (less is better). The **best** results are highlighted in bold.

QM9								
# Metrics	Efficiency		Atom Sta	Mol Sta	Quality (%)			
	S-Time	Steps			Valid	Uniqueness	Novelty	Significance
Data	-	-	99.0	95.2	97.7	100.0	-	-
ENF	-	-	85.0	4.9	40.2	98.0	-	-
G-Schnet	-	-	95.7	68.1	85.5	93.9	-	-
GDM-aug	1.50	1000	97.6	71.6	90.4	<b>99.0</b>	66.8	73.9
EDM	1.68	1000	98.7	82.0	91.9	98.7	65.7	64.8
EDM-Bridge	-	1000	98.8	84.6	92.0	98.6	-	-
GeoLDM	1.86	1000	98.9	89.4	93.8	98.8	58.1	53.9
GeoBFN	-	50	98.3	85.1	92.3	98.3	72.9	66.1
	0.16	100	98.6	87.2	93.0	98.4	70.3	64.4
	-	500	98.8	88.4	93.4	98.3	67.7	62.1
	-	1000	<b>99.1</b>	<b>90.9</b>	<b>95.3</b>	97.6	66.4	61.8
EquiFM	0.37	200	98.9	88.3	94.7	98.7	57.4	53.7
GOAT	<b>0.12</b>	<b>90</b>	98.4	84.1	90.0	<b>99.0</b>	<b>78.6</b>	<b>72.3</b>

Table 6: Comparisons of generation quality (larger is better) in terms of Atom Stability, Molecule Stability, Validity, Uniqueness, Novelty, and Significance. And comparisons of generation efficiency regarding generation time and sampling steps per molecule (less is better). The **best** results are highlighted in bold.

GEOM-DRUG						
# Metrics	Efficiency		Atom Sta	Quality (%)		
	S-Time	Steps		Mol Sta	Valid	Uniqueness
Data	-	-	86.5	0.0	99.9	100.0
ENF	-	-	-	-	-	-
G-Schnet	-	-	-	-	-	-
GDM-aug	-	1000	77.7	-	91.8	-
EDM	14.88	1000	81.3	0.0	92.6	99.9
EDM-Bridge	-	1000	82.4	-	92.8	-
GeoBFN	-	50	78.9	-	93.1	-
	-	100	81.4	-	93.5	-
	-	500	85.6	-	92.1	-
	-	1000	<b>86.2</b>	-	91.7	-
GeoLDM	12.84	1000	84.4	0.0	<b>99.3</b>	99.9
EquiFM	-	200	84.1	-	98.9	-
GOAT	<b>0.94</b>	<b>90</b>	84.8	0.0	96.2	<b>99.9</b>

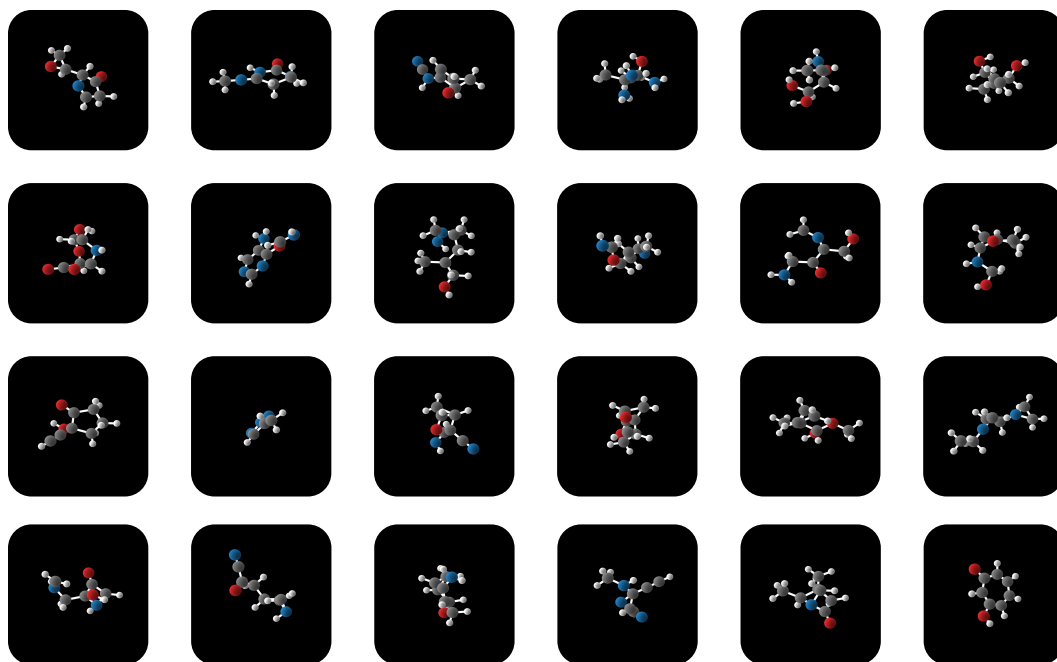


Figure 5: Molecules Generated by GOAT trained on QM9.

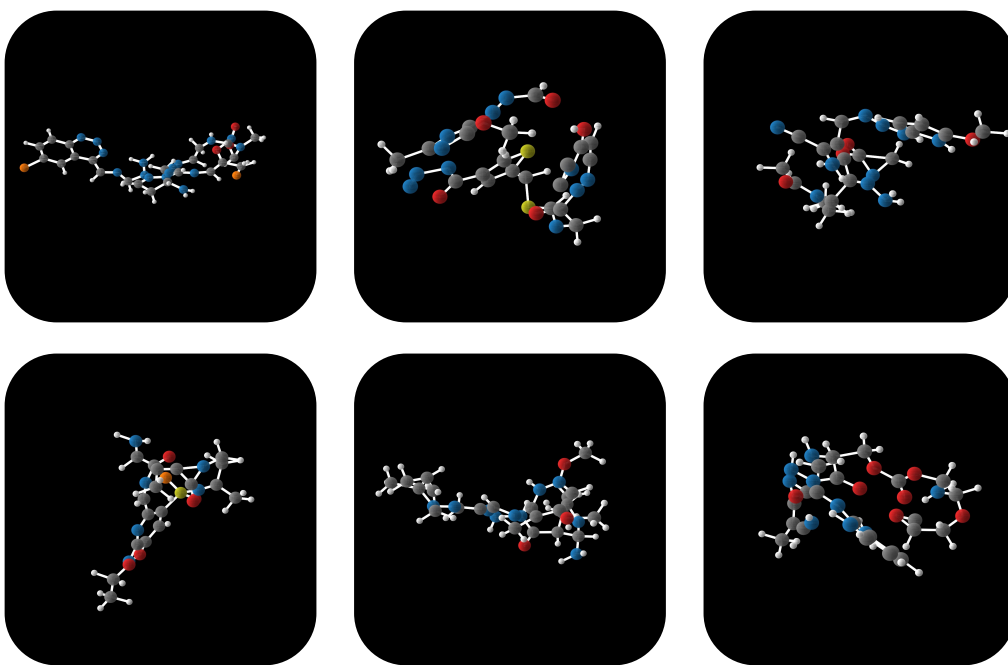


Figure 6: Molecules Generated by GOAT trained on GEOM-DRUG.

## H IMPACT STATEMENTS

This paper contributes to the advancement of generative Artificial Intelligence (AI) in scientific domains, including material science, chemistry, and biology. The insights gained will significantly enhance generative AI technologies, thereby streamlining the process of scientific knowledge discovery.

The application of machine learning to molecule generation expands the possibilities for molecule design beyond therapeutic purposes, potentially leading to the creation of illicit drugs or hazardous substances. This potential for misuse and unforeseen consequences underscores the need for stringent ethical guidelines, robust regulation, and responsible use of these technologies to safeguard individuals and society.