Equivariant Transformer Forcefields for Molecular Conformer Generation

May 26, 2024

Abstract

Molecular conformer generation is vital to computational chemistry and drug discovery, but it remains challenging due to the extensive range of possible conformations. In this paper, we propose a novel approach for molecular conformer generation that utilizes an Equivariant Transformwer Forcefield (ETF) pre-trained on large-scale molecular datasets to refine the quality of the conformers. This strategy begins with an initial set of conformers, which are subsequently refined through structural optimization. We demonstrate that our ETF-based optimization significantly improves the quality of the conformers generated by state-of-the-art methods, achieving a reduction in 45% the distance to the reference conformers. Furthermore, our methodology outperforms the classical forcefields by improving precision without sacrificing recall. Lastly, it can deliver competitive performance even when beginning with a simple initialization of conformers by RDKit, demonstrating its robustness and potential for extensive applications in computational chemistry and drug discovery.

1. Introduction

The generation of molecular conformers, *i.e.*, spatial arrangements of atoms in the low-energy states of a molecule, plays a critical role in computational chemistry and drug discovery. Understanding the range of conformers that a molecule can adopt is essential, as these conformers largely dictate the molecule's biological activity and physical properties (Hawkins, 2017). However, the generation of conformers is a challenging task, especially for large and flexible molecules, given the enormous conformation space.

Traditional conformer generation approaches typically rely on rules and knowledge extracted from known conformers to search through a molecule's conformational space (Hawkins, 2017). Two well-known tools in this domain are OMEGA (Hawkins et al., 2010) and ETKDG (Riniker & Landrum, 2015). OMEGA is a commercial software that applies systematic search with a set of rules and heuristics to narrow down the search space, whereas ETKDG, an open source conformer generation tool, uses a distance geometry-based stochastic approach to generate conformers with certain constraints. These rules and knowledge trade off the quality of generated conformers for efficient computation.

To obtain higher quality samples more efficiently, researchers have recently turned to deep generative models for conformer generation (Ganea et al., 2021; Xu et al., 2022; Jing et al., 2022). In particular, diffusion models, which learn to reverse the diffusion process of conformers within either Euclidean or Torsional space, have been proven to be effective in producing high-quality molecular conformers (Xu et al., 2022; Jing et al., 2022).

In this work, we reframe conformer generation as a task of conformer refinement, and introduce a machine learning forcefield-based approach to improve sample quality. Instead of directly generating conformers for a specific molecule, our approach begins with an initial set of conformers, then refines the distribution by structural optimization. We optimize the 3D structure of each individual conformer using an Equivariant Transformer Forcefield (ETF) (Feng et al., 2023). Pre-trained on the off-equilibrium conformations of small molecules and polymers, the ETF is capable of predicting the force acting on each atom in a conformation, which is then useful in pushing the conformations towards their lower energy states. It's worth noting that our approach, being orthogonal to existing ones, can seamlessly integrate with any other conformer generation methods. Furthermore, unlike deep generative models and cheminformatics methods, our approach does not require access to exemplar conformers.

We demonstrate that our approach significantly improves the quality of the conformer distribution generated by state-of-the-art methods by optimizing the structure of each individual conformer. This has resulted in a 45% reduction in the distance between the generated and reference conformers, measured in terms of Average Minimum RMSD (AMR) on the GEOM-QM9 dataset (Ax-

elrod & Gomez-Bombarelli, 2022). Unlike classical hand-engineered forcefields such as Universal Force Field (UFF) (Rappé et al., 1992) and Merck Molecular Force Field (MMFF94) (Halgren, 1996), our method does not compromise recall for precision. Moreover, we have illustrated that our approach can maintain performance comparable to that of state-of-the-art methods, even when using straightforward initializations of conformers.

The contributions of this paper are summarized as follows:

- We propose a novel method that employs the Equivariant Transformer Forcefield (ETF) for molecular conformer generation via structural optimization.
- Our technique substantially refines the quality of conformers generated by state-of-the-art deep generative models, resulting in a 45% reduction in Average Minimum RMSD (AMR), a metric that measures the distance between the generated conformers and their reference counterparts.
- We introduce a straightforward diversifying sampling strategy for initialization, using RDKit. With this simple initialization, our approach outperforms nearly all baselines on both the Coverage and AMR metrics.

2. Background

2.1. Traditional Methods for Conformer Generation

Traditional methods for conformer generation typically consist of two steps: 1) exploring the vast conformational space through conformation search methods, and 2) refining the conformers generated in the previous step by optimizing their 3D structure using forcefields.

Conformation Search. Classical methods for searching through conformational space for low-energy samples can be broadly divided into two categories based on their sampling methods: stochastic and systematic (Hawkins, 2017). Stochastic sampling methods predominantly rely on molecular dynamics (MD), Monte Carlo-simulated annealing (MC), distance geometry (DG), and genetic algorithms (GAs). While most MD (Tsujishita & Hirono, 1997) and MC (Wilson et al., 1991; Sperandio et al., 2009; Chang et al., 1989) based methods do not scale well, DG-based methods (Vainio & Johnson, 2007; Lagorce et al., 2009), especially ETKDG (Riniker & Landrum, 2015), strike a good balance between computational efficiency and accuracy. Generally, stochastic methods produce non-deterministic output, and the computational effort needed to find diverse and high-quality conformers is unpredictable (Hawkins, 2017). On the other hand, systematic methods comprehensively search the conformational space. Consequently, rules and knowledge bases such as allowed torsion angles, permissible paths, and libraries of 3D fragment conformations are necessary to reduce the search space (Beusen et al., 1996; Sauton et al., 2008; Smellie et al., 2003). Among them, OMEGA (Hawkins et al., 2010) represents state-of-the-art software for systematic search. Overall, systematic methods tend to be inflexible as they rely heavily on rules and existing knowledge of local structures.

Structure Optimization / Energy Minimization. Once the conformer candidates have been generated, forcefields are typically employed to optimize these conformers through energy minimization. Forcefields calculate the potential energy and forces of atoms, guiding energy minimization using optimization algorithms. *ab-initio* forcefields, such as Density Functional Theory (DFT), are highly accurate but costly, limiting their applications for large-scale applications for large molecules. In contrast, empirical or classical forcefields like the Universal Force Field (UFF) (Rappé et al., 1992) and the Merck Molecular Force Field (MMFF94) (Halgren, 1996), while less accurate, offer much faster computations. These forcefields apply simplified assumptions, fit empirical potential energy function parameters from experimental data, and typically model interactions between atom pairs and triplets, making them more suitable for large-scale simulations, albeit mainly for specific molecule types.

Machine learning forcefields attempt to bridge the gap between the accuracy of *ab-initio* forcefields and the efficiency of classical forcefields. Typically, they approximate the *ab-initio* forcefield predictions with machine learning models designed to encode invariant or equivariant features for molecules. Traditional Graph Neural Network or Message Passing based methods, such as SchNet (Schütt et al., 2018), DIME-Net (Yeh et al., 2023; Gasteiger et al., 2020), and GemNet (Gasteiger et al., 2021), build on invariant interatomic features such as bond lengths and angles. These methods can learn invariant molecular representations for potential energies. More advanced Equivariant Graph Neural Network based methods, such as NequIP (Batzner et al., 2022), EGNN (Satorras et al., 2021), and Equivariant Transformer (ET) (Thölke & De Fabritiis, 2022), directly learns equivariant features that represent a larger physically valid function space. Because of the improved representation ability, they usually achieve higher accuracy than invariant methods, at the cost of increased computation overhead to constrain the features in the equivariant space. The exception is ET (Thölke & De Fabritiis, 2022), which achieves equivariance with vector embedding and equivariant attention mechanism, with significantly higher efficiency.

2.2. Deep Learning Methods for Conformer Generation

Recently, there has been a surge of interest in employing deep learning methods for conformer generation. Earlier methods either directly predicted atomic coordinates (Mansimov et al., 2019; Zhu et al., 2022), estimated interatomic distances (Simm & Hernández-Lobato, 2019; Xu et al., 2021a;b), or predicted the gradient of coordinates (Shi et al., 2021; Luo et al., 2021). The most cutting-edge methods include GeoDiff (Xu et al., 2022) and Torsional Diffusion (Jing et al., 2022), which learn to reverse the diffusion process in Euclidean and Torsional space, respectively. GeoMol (Ganea et al., 2021), another recent method, focuses solely on predicting local structures and torsional angles.

3. Methodology

In this section, we first propose using machine learning forcefields to refine conformer samples through structural optimization. We show that the proposed approach can be unified with recent diffusion-based conformer generation methods under the same sample refinement scheme. Following this, we introduce an Equivariant Transformer Forcefield (ETF) for structural optimization and detail its model architecture, training objective, and training dataset.

3.1. A Unified Sample Refinement Scheme for Conformer Generation

Diffusion-based Probabilistic models, such as GeoDiff (Xu et al., 2022) and Torsional Diffusion (Jing et al., 2022), have demonstrated empirically their capability to generate superior conformer distributions. They outperform other machine learning and cheminformatics methods in both RMSD and chemical property predictions. These models define a Markov chain of the diffusion process by injecting random noise into the data iteratively. They then learn to reverse this diffusion process to recover data samples from the injected noise. To sample a low-energy conformer, diffusion methods initially draw a conformation from the uniform distribution, then iteratively apply the learned diffusion model to refine the conformation. However, such diffusion models are limited by the quality and quantity of available training data, potentially impeding further improvement in the quality of generated samples.

In this study, we propose utilizing machine learning forcefields to further refine the conformers generated by diffusion models via structural optimization. A forcefield is a model that uses the atomic coordinates of a molecule to predict the force acting on each atom. This forcefield can be applied to optimize a molecular conformation by minimizing the energy of the system. As the structural optimization occurs locally, with a good initial coverage of the conformational space, the forcefield-based approach can be used to further improve the quality of each individual sample, without sacrificing the diversity of the generated distribution.

Interestingly, we can unify the sampling schemes of diffusion methods and local structural optimization with forcefields in Algorithm 1. In this algorithm, the sample-refining operator $\mathcal{I}_t(c_{t+1}|c_t)$ takes a conformation c_t as input and outputs a new conformation c_{t+1} with improved sample quality. The operator $\mathcal{I}_t$ can represent either a step in the denoising diffusion process or in the local optimization process. This operator is applied iteratively to the initial conformation c_0 to create a sequence of conformations $c_0, c_1, \ldots, c_T$. The final conformation c_T is then used as the output of the sample refinement process.

Note that, unlike the reverse diffusion process, the sample improvement process doesn't necessarily require random uniform initialization. In fact, the initial conformation c_0 can be any conformation, including those generated by other conformer generation methods. In our experiments, we demonstrate that even with a simple initialization scheme (using ETKDG (Riniker & Landrum, 2015) and clustering, as illustrated in Algorithm 2), our approach can still deliver state-of-the-art performance.

3.2. Equivariant Transformer Forcefields for Local Structural Optimization

In this study, we propose to use forcefields to optimize molecular conformations to their low-energy states. Consider a molecule x comprised of n_x atoms. This molecule x can exist in various 3D conformations $c_x \in \mathbb{R}^{n_x \times 3}$, each with a corresponding potential energy denoted as $E = E(x, c_x) \in \mathbb{R}$. The forcefield model $\mathcal{F}$ is a function that takes a molecule x and a conformation c_x as input and outputs the force acting on each atom: $\mathcal{F}(x, c_x) = \frac{\partial E(x,c_x)}{\partial c_x} \in \mathbb{R}^{n_x \times 3}$. Given an initial conformer c_0, the local structural optimization can then be formulated as a local optimization problem:

$$c^* = \underset{c_x}{\operatorname{argmin}} E(x, c_x) \quad \text{s.t.} \quad c_x \in \mathcal{N}(c_0), \tag{1}$$

where $\mathcal{N}(c_0)$ represent the neighborhood of c_0. We consider classic second-order optimization algorithms such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) (Broyden, 1970), Limited-memory BFGS (LBFGS) (Nocedal, 1980) and conjugate gradient (Hestenes et al., 1952) due to the small problem scale and their faster convergence rate. These methods require the access to the gradient function $\frac{\partial E(x,c_x)}{\partial c_x}$, which can be computed by the forcefield model $\mathcal{F}$. However, obtaining highly accurate forces require the use of *ab-initio* methods such as the Density Functional Theory (DFT). These methods are computationally expensive and are not suitable for force computation at inference

Algorithm 1: Diffusion + FF Optimization

Data: Molecule m

Sample $c_0 \sim$ uniform(m);
for t *in* $\{1, 2, \cdots, N_D + N_{FF}\}$ **do**
  if $t \leq N_D$ **then**
    $c_t \leftarrow \mathcal{I}_D(c_t|c_{t-1})$
  else
    $c_t \leftarrow \mathcal{I}_{FF}(c_t|c_{t-1})$
  end
end
return $c_{N_D+N_{FF}}$

Algorithm 2: ETKDG + FF Optimization

Data: Molecule m

candidates $\leftarrow$ ETKDG(m);
cluster centers $\leftarrow$ K-means(candidates);
Sample $c_0 \sim$ cluster centers.;
for t *in* $\{1, 2, \cdots, N_{FF}\}$ **do**
  $c_t \leftarrow \mathcal{I}_{FF}(c_t|c_{t-1})$
end
return $c_{N_{FF}}$

Figure 1. Algorithm 1 unifies the diffusion model and forcefield optimization within the sample refinement scheme. After sampling a conformation from a uniform distribution, it is refined for N_D steps using the learned diffusion denoising model. Subsequently, the forcefield refines it further for N_{FF} steps. The final conformation serves as one conformer sample from the process. Algorithm 2 utilizes ETKDG to generate an initial set of conformations. Once a candidate conformation set has been generated, it applies K-means clustering to extract the cluster centers for further refinement. This straightforward initialization strategy ensures an adequate level of diversity among the initially generated samples.

time. Classic forcefields such as MMFF94 and UFF, on the other hand, are fast but inaccurate.

We propose to use the ET-OREO (Feng et al., 2023) forcefield model to strike the balance between accuracy and efficiency. ET-OREO is an Equivariant Transformer-based model, pre-trained on four public datasets comprising over 15 million in equilibrium and off-equilibrium molecules. The model achieves state-of-the-art force prediction accuracy and molecular dynamics simulations. At the same time, ET-OREO achieves 3 times faster inference than the similarly performing model, NequIP, and achieves high fidelity and robustness in MD simulations.

Model Architecture. We use an Equivariant Transformer(Thölke & De Fabritiis, 2022) (ET) for learning molecular embeddings. ET leverages the equivariant attention mechanism to capture the quantum mechanical interactions between atoms. We represent a ET model parameterized with θ as Φ_θ, which maps (x, c_x) to $\Phi_\theta(x, c_x) \in \mathbb{R}$ and $\nabla_{c_X}\Phi_\theta(x, c_x) \in \mathbb{R}_{x\times 3}$, approximating the potential energy and forces, respectively.

In order to obtain the potential energy and forces predictions, the model consists of three parts: 1) the embedding layer, 2) the update layers, and 3) the output layer. The embedding layer transforms each atom into vector representations that encapsulate their quantum mechanical information and their interactions within the vicinity. This layer generates both scalar and vector embeddings, with the former derived by integrating an intrinsic vector (containing atom-specific data) and a neighborhood vector (accounting for atomic neighborhood interactions). The embedding layer converts each atom into vector representations that incorporate their quantum mechanical properties and nearby interactions. Update layers handle the sequential transformation of these embeddings within each network layer with *equivariant attention mechanisms*. Equivariant attention mechanisms encode interatomic interactions by considering the relative distance between atoms. The output network generates scalar predictions $\Phi_\theta(x, c_x)$ for each atom through Gated Equivariant Blocks (Schütt et al., 2021). The scalar prediction $\Phi_\theta(x, c_x)$ serves as the approximation for the potential energy, and its gradient w.r.t coordinates $\nabla_{c_X}\Phi_\theta(x, c_x)$ as forces prediction.

For the results to be physically well-defined, the model must be equivariant w.r.t. the SE(3) group on $\mathbb{R}^3$ to input coordinates. The SE(3) group contains rotation and translation operations on 3D coordinates. Intuitively, equivariance dictates that the model predicts potential energy that remains unchanged when input coordinates are translated or rotated. Furthermore, the model should predict forces that transform according to the input coordinates. Formally,

$$\Phi_\theta(x, g(c_x)) = \Phi_\theta(x, c_x),$$

$$\nabla_{c_x}\Phi_\theta(x, g(c_x)) = g(\nabla_{c_x}\Phi_\theta(x, c_x)),$$

where g is taken from the SE(3) group. Our Equivariant Transformer achieves the desired equivariance by leveraging both scalar and vector embeddings for each atom: the scalar embedding depends on the interatomic distances $\|c_x^i - c_x^j\|_2^2 \in \mathbb{R}$, where i, j index atoms in the molecule c. The scalar embedding is invariant to the inputs as the interatomic distances remain unchanged with rotation and translation operations. The vector embedding depends on the interatomic distance vectors $c_x^i - c_x^j \in \mathbb{R}^3$, which change according to the rotation operations applied to inputs. Thus, the Equivariant Transformer is able to parameterize a large

range of physically valid functions from input coordinates to potential energy and forces.

Compared to Equivariant Graph Neural Networks (Satorras et al., 2021), Equivariant Transformer enjoys improved efficiency with the equivariant attention mechanism, which naturally models the interactions between atoms. Furthermore, the scalar embedding in higher dimensional latent space in Equivariant Transformer endows it with higher expressivity of equivariant features. Compared to e3nn, Equivariant Transformer does not require the complex and inefficient calculation of SE(3) representations in high dimensions to achieve similar levels of expressivity in equivariant features. Hence, we chose Equivariant Transformer as our backbone model to achieve both high expressivity and computational efficiency.

Force-Centric Training. We use a variant of ET-OREO which only require off-equilibrium samples to train the model. The model parameter is optimized by minimizing the following objective function:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \|\nabla_{c_x} \Phi_\theta(x, c_x) - \mathcal{F}(x, c_x)\|. \qquad (2)$$

The objective function (2) is solely focused on forces optimization. We implement this design for a number of reasons. 1) Forces are a locally well-defined quantity that is more generalizable across different molecules; 2) By learning forces as a gradient, the model naturally learns the potential energy surface up to linear transforms. From an optimization perspective, the model learns the same locally optimal conformations. While involving potential energy in the objective function is tempting, ET-OREO (Feng et al., 2023) shows that the joint optimization of potential energy and forces is difficult in a multi-molecule setting.

Training Data. The training data consists of small off-equilibrium molecules from several different domains from three different sources: *poly24*, MD17, and ANI1-x. *Poly24* consists of polymer structures, while MD17 and ANI1-x consists of small organic molecules. The details of the data composition are given below.

Poly24: MD Simulations for Polymers. This dataset is proposed by (Feng et al., 2023) and composes of DFT-based molecular dynamics simulations for 24 polymer types. For each type, a polymer is constructed with a specific monomer repeated L times in a loop structure. In total, we run generally 10 DFT simulations for different initialization of each L-loop ($L = 1, 3, 4, 5, 6$) polymer across the 24 types of polymers. In this paper, we only use polymers with less than or equal to 64 atoms for training ETF, totaling 3,851,540 conformations.

MD17 and ANI1-x: small organic molecules. In addition to *poly24*, we have utilized three existing public datasets, namely MD17(Chmiela et al., 2017; 2018) and ANI1-x(Smith et al., 2020) for training our forcefield. These datasets contain small organic molecules in a vacuum, and property prediction for such molecules is an area of great interest to the cheminformatics community. Machine learning for the molecules community has also extensively studied and benchmarked these datasets.

In summary, we have 3,851,540 3D conformations from *poly24*, 3,611,115 from MD17, and 4,956,005 from ANI1-x, totaling 12,339,673 conformations for training ETF.

4. Experiments

In this section, we present a comprehensive empirical evaluation of our Equivariant Transformer Forcefield-based structural optimization approach, referred to as ETF optimization for the remainder of this section, for generating molecular conformers. We assess our methodology's performance using a diverse set of small organic molecules and compare our approach with a range of state-of-the-art molecular conformer generation baselines, including both cheminformatics and machine-learning-based methods. For a fair comparison, we utilize the GEOM-QM9 dataset, a widely recognized benchmark for molecular conformer generation techniques. We employ two primary metrics, Average Minimum Root Mean Square Distance (AMR) and Coverage, to compare the conformers generated by our approach with the reference conformers in the dataset. The results of our experiments underscore the effectiveness of our ETF optimization approach in generating high-quality molecular conformers.

4.1. Experimental Settings

Dataset. Our evaluation is based on the GEOM-QM9 dataset (Axelrod & Gomez-Bombarelli, 2022), consistent with prior studies (Jing et al., 2022; Ganea et al., 2021). Derived from the well-established QM9 database (Ramakrishnan et al., 2014), the GEOM-QM9 dataset serves as a common benchmark for comparing molecular conformer generation methods. It contains conformers for 133k small organic molecules. In line with previous research, we maintain the same train/validation/test splits for our experiments.

Metrics. Our comparative analysis employs two primary metrics, Average Minimum RMSD (AMR), and Coverage, to measure the difference between the generated conformers and the dataset's reference conformers. The Coverage metric measures the percentage of one conformer set present within the δ-neighborhood of the other, whereas AMR quantifies the average distance between each con-

former in one set relative to the other. Both metrics are computed under the Recall and Precision settings. The Recall setting assesses how extensively the generated conformers covers the reference conformers, while the Precision setting evaluates the accuracy of the generated conformers.

Let $\mathcal{G} = \{\hat{c}_g\}_{g=1}^{|\mathcal{G}|}$, and $\mathcal{R} = \{c_r\}_{r=1}^{|\mathcal{R}|}$ represent the sets of generated and reference conformers for the same molecule, respectively. In line with existing work, each method is permitted to generate twice as many conformers as in the reference group: $|\mathcal{G}| := 2|\mathcal{R}|$. The Coverage and AMR metrics under the Recall setting are defined as follows: The Coverage metrics under the Recall setting, Coverage-Recall$_\delta(\mathcal{G}|\mathcal{R})$ is defined as follows:

$$\frac{1}{|\mathcal{R}|}\Big|\{c_r \in \mathcal{R} : \exists\, \hat{c}_g \in \mathcal{G},\ \text{RMSD}(c_r, \hat{c}_g) < \delta\}\Big|, \quad (3)$$

while AMR-Recall$(\mathcal{G}|\mathcal{R})$ is defined as

$$\frac{1}{|\mathcal{R}|}\sum_{c_r \in \mathcal{R}} \min_{\hat{c}_g \in |\mathcal{G}|} \text{RMSD}(c_r, \hat{c}_g) \quad (4)$$

. In the equations above, RMSD denotes the Root Mean Square Distance of atom-wise distances following the alignment of the conformer pair. Similarly, the Coverage and AMR metrics under the Precision setting can be computed by swapping the roles of $\mathcal{G}$ and $\mathcal{R}$ in the above equations.

Baselines. We compare our ETF optimization approach with a range of state-of-the-art molecular conformer generation baselines, including both classical cheminformatics methods and recent deep generative model based approaches. The baselines we consider are: The following conformer generation baselines are chosen as comparison to our ETF optimization approach, including both classical methods and recent deep generative model approaches:

- **RDKit's ETKDG.** ETKDG (Riniker & Landrum, 2015) is a well-accepted method for 3D conformer generation implemented in the RDKit cheminformatics library. ETKDG uses a distance geometry-based approach with embedded torsion angle preferences from small molecule crystal data.
- **OMEGA.** OMEGA (Hawkins et al., 2010) is a widely used commercial software for generating conformers. OMEGA applies a systematic search method based on rules derived from observed geometries in databases.
- **GeoMol.** GeoMol (Ganea et al., 2021) employs a SE(3)-invariant graph neural network to predict adjacent atomic coordinates and torision angles.
- **GeoDiff.** GeoDiff (Xu et al., 2022) is a Euclidean-space diffusion model that generate models conformers as 3D point clouds.
- **Torsional Diffusion.** Torsional Diffusion (Jing et al., 2022) constructs a diffusion model in the space of torsional angles, resulting in a significant reduction in dimensionality.

4.2. Molecular Conformer Generation

Main Results. Table 1 outlines a comparison between our proposed Equivariant Transformer Forcefield (ETF)-based structural optimization approach and existing baselines. Overall, the best performance is achieved when ETF optimization is applied to the Torsional Diffusion samples. A notable reduction of 35% in mean AMR and 45% in median AMR is observed in both the Recall and Precision settings. This suggests considerable potential for improvement even for the leading existing deep generative model, and demonstrating the effectiveness of ETF optimization.

Limited Improvement on Recall Coverage. While our method significantly improves most metrics when initialized with Torsional Diffusion samples, the improvement in the Coverage rate under the Recall setting is relatively modest. This observation is consistent with our understanding of local optimization – since it does not leap over the energy barrier, it does not aid in broadening coverage to include more low-energy modes.

RDKit Initializations. Even with basic RDKit initialization (ETKDG + ETF), the conformers obtained post-ETF optimization still outperform the baselines significantly on AMR metrics. However, due to the aforementioned limitations, this combination tends to have low coverage of reference conformers. To improve the coverage, we preprocess the ETKDG conformers with K-means clustering, i.e., we generate 10 times more conformers using ETKDG and extract the K-means cluster centers for initialization. This initialization strategy (ETKDG + K-means + ETF) notably increases Recall Coverage by 5% and reduces Recall AMR by 20%, further demonstrating the effectiveness of ETF optimization, even with simple initialization.

Ablation Studies. We conducted several ablation studies to examine the effect of the ETF model size, the optimization algorithm used, and the choice of forcefields. The results are shown in Table 2. Unless otherwise stated, all methods default to using the 12-layer ETF model in conjunction with the BFGS optimization algorithm, with initial conformers generated by Torsional Diffusion. Our findings indicate that increasing the number of layers from 8 to 12 significantly reduces the AMR, suggesting that a larger model is capable of learning a more accurate

Method	Recall				Precision			
	Coverage ↑		AMR ↓		Coverage ↑		AMR ↓	
	Mean	Med	Mean	Med	Mean	Med	Mean	Med
ETKDG (Riniker & Landrum, 2015)	82.3	100.0	0.234	0.198	84.3	100.0	0.236	0.206
OMEGA (Hawkins et al., 2010)	85.5	100.0	0.177	0.126	82.9	100.0	0.224	0.186
GeoMol (Ganea et al., 2021)	**91.5**	100.0	0.225	0.193	86.7	100.0	0.270	0.241
GeoDiff (Xu et al., 2022)	76.5	100.0	0.297	0.229	50.0	33.5	0.524	0.510
Torsional Diffusion (Jing et al., 2022)	88.4	100.0	0.178	0.147	84.5	100.0	0.221	0.195
ETKDG + ETF	83.5	100.0	0.169	0.108	86.5	100.0	0.155	0.108
ETKDG + K-means + ETF	87.5	100.0	0.135	0.087	86.2	100.0	0.158	0.110
Torsional Diffusion + ETF	88.8	**100.0**	**0.116**	**0.078**	**87.0**	**100.0**	**0.144**	**0.106**

Table 1. Main results for molecular conformer generation. ETKDG and OMEGA are cheminformatics software tools, while GeoMol, GeoDiff, and Torsional Diffusion represent state-of-the-art deep generative models. The Coverage and AMR metrics are defined in Eq (3) and Eq (4), respectively. The term 'Recall' refers to the computation of metrics with respect to each reference conformer, whereas 'Precision' refers to the computation of metrics for each generated conformer. The application of ETF optimization results in a significant improvement in all metrics under both the Recall and Precision settings. ETF optimization, even with simple RDKit and K-means initialization, outperforms all baseline methods except for GeoMol in terms of the Coverage metrics.

Method	Recall				Precision			
	Coverage ↑		AMR ↓		Coverage ↑		AMR ↓	
	Mean	Med	Mean	Med	Mean	Med	Mean	Med
L8	88.7	100.0	0.151	0.123	85.9	100.0	0.189	0.160
L12	**88.8**	**100.0**	**0.116**	**0.078**	**87.0**	**100.0**	**0.144**	**0.106**
CG (Hestenes et al., 1952)	88.6	100.0	0.130	0.093	86.5	100.0	0.161	0.127
LBFGS (Nocedal, 1980)	**89.0**	100.0	0.118	0.082	86.9	100.0	0.150	0.114
BFGS (Broyden, 1970)	88.8	**100.0**	**0.116**	**0.078**	**87.0**	**100.0**	**0.144**	**0.106**
UFF (Rappé et al., 1992)	84.5	100.0	0.202	0.173	85.7	100.0	0.185	0.142
MMFF94 (Halgren, 1996)	86.0	100.0	0.182	0.148	**87.2**	100.0	0.167	0.129
ETF	**88.8**	**100.0**	**0.116**	**0.078**	87.0	**100.0**	**0.144**	**0.106**

Table 2. Ablation studies on the ETF model size, optimization algorithm used, and the choice of forcefields. L8 and L12 denote the 8-layer and 12-layer ETF models, respectively. CG refers to the conjugate gradient optimization algorithm, while UFF and MMFF94 represent classical forcefield models. Unless specified otherwise, all methods default to using the L12 ETF model combined with the BFGS optimization algorithm, with initial conformers generated by Torsional Diffusion.

forcefield. Additionally, we observed that the BFGS optimization algorithm (Broyden, 1970) is preferred over LBFGS (Nocedal, 1980) and the conjugate gradient (CG) method (Hestenes et al., 1952).

ETF versus Classical Forcefields. When compared with classical forcefields such as Universal Force Field (UFF) (Rappé et al., 1992) and Merck Molecular Force Field (MMFF94) (Halgren, 1996), local optimization using our ETF yields significantly improved Coverage and AMR scores. Unlike UFF and MMFF94, ETF optimization does not compromise recall scores for the sake of precision scores, when compared to the pre-optimized Torsional Diffusion samples. This result suggests that the ETF closely approximates the exact physical force, allowing ETF optimization to recover more modes without mode collapsing.

5. Conclusion

In this work, we introduced an innovative approach for molecular conformers generations based on an Equivariant Transformer Forcefield model. Extensive experiments on the GEOM-QM9 dataset confirmed the significant improvements our method offers in terms of the generated conformers' quality, with a notable reduction in Average Minimum RMSD and improved performance on the Coverage metric. Our method was effective even with simple initialization of conformers, highlighting its robustness.

The results of our study underscore the vast potential of pre-trained machine learning forcefields in the realm of computational chemistry and drug discovery. Future studies could look into non-local structural optimization methods, which may help further improve the Recall Coverage. Also, the scalability of our method to larger and more complex molecules remains an exciting area to explore.

References

Simon Axelrod and Rafael Gomez-Bombarelli. Geom, energy-annotated molecular conformations for property prediction and molecular generation. *Scientific Data*, 9 (1):185, 2022.

Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature communications*, 13(1):2453, 2022.

Denise D Beusen, EF Berkley Shands, SF Karasek, Garland R Marshall, and Richard A Dammkoehler. Systematic search in conformational analysis. *Journal of Molecular Structure: THEOCHEM*, 370(2-3):157–171, 1996.

Charles George Broyden. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA Journal of Applied Mathematics*, 6(1):76–90, 1970.

George Chang, Wayne C Guida, and W Clark Still. An internal-coordinate monte carlo method for searching conformational space. *Journal of the American Chemical Society*, 111(12):4379–4386, 1989.

Stefan Chmiela, Alexandre Tkatchenko, Huziel E Sauceda, Igor Poltavsky, Kristof T Schütt, and Klaus-Robert Müller. Machine learning of accurate energy-conserving molecular force fields. *Science advances*, 3(5): e1603015, 2017.

Stefan Chmiela, Huziel E Sauceda, Klaus-Robert Müller, and Alexandre Tkatchenko. Towards exact molecular dynamics simulations with machine-learned force fields. *Nature communications*, 9(1):3887, 2018.

Rui Feng, Qi Zhu, Huan Tran, Binghong Chen, Aubrey Toland, Rampi Ramprasad, and Chao Zhang. May the force be with you: Unified force-centric pre-training for 3d molecular conformations. 2023.

Octavian Ganea, Lagnajit Pattanaik, Connor Coley, Regina Barzilay, Klavs Jensen, William Green, and Tommi Jaakkola. Geomol: Torsional geometric generation of molecular 3d conformer ensembles. *Advances in Neural Information Processing Systems*, 34:13757–13769, 2021.

Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional message passing for molecular graphs. *arXiv preprint arXiv:2003.03123*, 2020.

Johannes Gasteiger, Florian Becker, and Stephan Günnemann. Gemnet: Universal directional graph neural networks for molecules. *Advances in Neural Information Processing Systems*, 34:6790–6802, 2021.

Thomas A Halgren. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *Journal of computational chemistry*, 17(5-6): 490–519, 1996.

Paul CD Hawkins. Conformation generation: the state of the art. *Journal of chemical information and modeling*, 57(8):1747–1756, 2017.

Paul CD Hawkins, A Geoffrey Skillman, Gregory L Warren, Benjamin A Ellingson, and Matthew T Stahl. Conformer generation with omega: algorithm and validation using high quality structures from the protein databank and cambridge structural database. *Journal of chemical information and modeling*, 50(4):572–584, 2010.

Magnus R Hestenes, Eduard Stiefel, et al. Methods of conjugate gradients for solving linear systems. *Journal of research of the National Bureau of Standards*, 49(6):409–436, 1952.

Bowen Jing, Gabriele Corso, Jeffrey Chang, Regina Barzilay, and Tommi Jaakkola. Torsional diffusion for molecular conformer generation. *Advances in Neural Information Processing Systems*, 35:24240–24253, 2022.

David Lagorce, Tania Pencheva, Bruno O Villoutreix, and Maria A Miteva. Dg-ammos: A new tool to generate 3d conformation of small molecules using d istance g eometry and a utomated m olecular m echanics o ptimization for in silico s creening. *BMC Chemical Biology*, 9(1): 1–10, 2009.

Shitong Luo, Chence Shi, Minkai Xu, and Jian Tang. Predicting molecular conformation via dynamic graph score matching. *Advances in Neural Information Processing Systems*, 34:19784–19795, 2021.

Elman Mansimov, Omar Mahmood, Seokho Kang, and Kyunghyun Cho. Molecular geometry prediction using a deep generative graph neural network. *Scientific reports*, 9(1):20381, 2019.

Jorge Nocedal. Updating quasi-newton matrices with limited storage. *Mathematics of computation*, 35(151):773–782, 1980.

Raghunathan Ramakrishnan, Pavlo O Dral, Matthias Rupp, and O Anatole Von Lilienfeld. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific data*, 1(1):1–7, 2014.

Anthony K Rappé, Carla J Casewit, KS Colwell, William A Goddard III, and W Mason Skiff. Uff, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American chemical society*, 114(25):10024–10035, 1992.

Sereina Riniker and Gregory A Landrum. Better informed distance geometry: using what we know to improve conformation generation. *Journal of chemical information and modeling*, 55(12):2562–2574, 2015.

Vıctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E (n) equivariant graph neural networks. In *International conference on machine learning*, pp. 9323–9332. PMLR, 2021.

Nicolas Sauton, David Lagorce, Bruno O Villoutreix, and Maria A Miteva. Ms-dock: accurate multiple conformation generator and rigid docking protocol for multi-step virtual ligand screening. *BMC bioinformatics*, 9(1):1–12, 2008.

Kristof Schütt, Oliver Unke, and Michael Gastegger. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International Conference on Machine Learning*, pp. 9377–9388. PMLR, 2021.

Kristof T Schütt, Huziel E Sauceda, P-J Kindermans, Alexandre Tkatchenko, and K-R Müller. Schnet–a deep learning architecture for molecules and materials. *The Journal of Chemical Physics*, 148(24), 2018.

Chence Shi, Shitong Luo, Minkai Xu, and Jian Tang. Learning gradient fields for molecular conformation generation. In *International conference on machine learning*, pp. 9558–9568. PMLR, 2021.

Gregor NC Simm and José Miguel Hernández-Lobato. A generative model for molecular distance geometry. *arXiv preprint arXiv:1909.11459*, 2019.

Andrew Smellie, Robert Stanton, Randy Henne, and Steve Teig. Conformational analysis by intersection: Conan. *Journal of computational chemistry*, 24(1):10–20, 2003.

Justin S Smith, Roman Zubatyuk, Benjamin Nebgen, Nicholas Lubbers, Kipton Barros, Adrian E Roitberg, Olexandr Isayev, and Sergei Tretiak. The ani-1ccx and ani-1x data sets, coupled-cluster and density functional theory properties for molecules. *Scientific data*, 7(1): 134, 2020.

Olivier Sperandio, Marc Souaille, François Delfaud, Maria A Miteva, and Bruno O Villoutreix. Med-3dmc: A new tool to generate 3d conformation ensembles of small molecules with a monte carlo sampling of the conformational space. *European journal of medicinal chemistry*, 44(4):1405–1409, 2009.

Philipp Thölke and Gianni De Fabritiis. Torchmd-net: equivariant transformers for neural network based molecular potentials. *arXiv preprint arXiv:2202.02541*, 2022.

Hideki Tsujishita and Shuichi Hirono. Camdas: An automated conformational analysis system using molecular dynamics. *Journal of computer-aided molecular design*, 11:305–315, 1997.

Mikko J Vainio and Mark S Johnson. Generating conformer ensembles using a multiobjective genetic algorithm. *Journal of chemical information and modeling*, 47(6):2462–2474, 2007.

Stephen R Wilson, Weili Cui, Jules W Moskowitz, and Kevin E Schmidt. Applications of simulated annealing to the conformational analysis of flexible molecules. *Journal of computational chemistry*, 12(3):342–349, 1991.

Minkai Xu, Shitong Luo, Yoshua Bengio, Jian Peng, and Jian Tang. Learning neural generative dynamics for molecular conformation generation. *arXiv preprint arXiv:2102.10240*, 2021a.

Minkai Xu, Wujie Wang, Shitong Luo, Chence Shi, Yoshua Bengio, Rafael Gomez-Bombarelli, and Jian Tang. An end-to-end framework for molecular conformation generation via bilevel programming. In *International Conference on Machine Learning*, pp. 11537–11547. PMLR, 2021b.

Minkai Xu, Lantao Yu, Yang Song, Chence Shi, Stefano Ermon, and Jian Tang. Geodiff: A geometric diffusion model for molecular conformation generation. *arXiv preprint arXiv:2203.02923*, 2022.

Shu-Hao Yeh, Shuangyu Xie, Di Wang, Wei Yan, and Dezhen Song. Dime-net: Neural network-based dynamic intrinsic parameter rectification for cameras with optical image stabilization system. *arXiv preprint arXiv:2303.11307*, 2023.

Jinhua Zhu, Yingce Xia, Chang Liu, Lijun Wu, Shufang Xie, Yusong Wang, Tong Wang, Tao Qin, Wengang Zhou, Houqiang Li, et al. Direct molecular conformation generation. *arXiv preprint arXiv:2202.01356*, 2022.