DiSCO: Diffusion Schrödinger Bridge for Molecular Conformer Optimization

Danyeong Lee[1], **Dohoon Lee**[2, 3], **Dongmin Bang**[1, 4], **Sun Kim**[1, 4, 5, 6]
[1]Interdisciplinary Program in Bioinformatics, Seoul National University
[2]Bioinformatics Institute, Seoul National University
[3]BK21 FOUR Intelligence Computing, Seoul National University
[4]AIGENDRUG Co., Ltd.
[5]Department of Computer Science and Engineering, Seoul National University
[6]Interdisciplinary Program in Artificial Intelligence, Seoul National University
{ldy9381, apap7, eugenomics, sunkim.bioinfo}@snu.ac.kr

Abstract

The generation of energetically optimal 3D molecular conformers is crucial in cheminformatics and drug discovery. While deep generative models have been utilized for direct generation in Euclidean space, this approach encounters challenges, including the complexity of navigating a vast search space. Recent generative models that implement simplifications to circumvent these challenges have achieved state-of-the-art results, but this simplified approach unavoidably creates a gap between the generated conformers and the ground-truth conformational landscape. To bridge this gap, we introduce **DiSCO**: **Di**ffusion **S**chrödinger Bridge for Molecular **C**onformer **O**ptimization, a novel diffusion framework that enables direct learning of nonlinear diffusion processes in prior-constrained Euclidean space for the optimization of 3D molecular conformers. Through the incorporation of an SE(3)-equivariant Schrödinger bridge, we establish the roto-translational equivariance of the generated conformers. Our framework is model-agnostic and offers an easily implementable solution for the post hoc optimization of conformers produced by any generation method. Through comprehensive evaluations and analyses, we establish the strengths of our framework, substantiating the application of the Schrödinger bridge for molecular conformer optimization. First, our approach consistently outperforms four baseline approaches, producing conformers with higher diversity and improved quality. Then, we show that the intermediate conformers generated during our diffusion process exhibit valid and chemically meaningful characteristics. We also demonstrate the robustness of our method when starting from conformers of diverse quality, including those unseen during training. Lastly, we show that the precise generation of low-energy conformers via our framework helps in enhancing the downstream prediction of molecular properties. The code is available at https://github.com/Danyeong-Lee/DiSCO.

Introduction

The three-dimensional structure of molecules is a key factor in determining their molecular properties. Therefore, it is essential to accurately predict the energetically optimal 3D molecular conformations for a variety of applications, from materials science to pharmaceutical design. The goal of molecular conformer generation is to sample conformers from each local minimum of the conformational energy landscape formulated as the Boltzmann distribution (Noé et al. 2019). Traditional approaches based on molecular dynamics (MD) simulations (Pracht, Bohle, and Grimme 2020) stand as the gold standard, yielding precise results by exhaustively exploring the conformational energy landscape. However, the high level of accuracy they achieve comes with substantial computational costs, rendering MD simulations unsuitable for high-throughput applications (Hawkins 2017). In contrast, cheminformatics-based approximation tools such as OMEGA (Hawkins et al. 2010) and ETKDG (Riniker and Landrum 2015) offer more computational efficiency but often at the expense of accuracy. Additionally, force field-based methods like MMFF (Halgren 1996), although employed for efficient optimization of conformers, still fall short in achieving the desired levels of accuracy (Kanal, Keith, and Hutchison 2018). Recent studies have focused on developing a fast and accurate conformer generation approach that leverages the capabilities of data-driven machine learning techniques.

Building on this framework, the evolution of deep generative models has made considerable progress in the field of molecular conformation generation. Previous work (Mansimov et al. 2019) attempted to generate molecular conformers directly in Euclidean space. This direct approach, however, has encountered several challenges, including the necessity for SE(3)-equivariance and the complexity of navigating an expansive search space, heightened by the high degree of freedom inherent in the task. To bypass the need to address SE(3)-equivariance, alternative approaches (Simm and Hernández-Lobato 2020; Xu et al. 2021a; Shi et al. 2021; Luo et al. 2021; Ganea et al. 2021) have utilized internal molecular geometries such as interatomic distances and torsion angles. However, the accumulation of prediction errors during the reassembly of entire conformer remains as a major drawback. Also, these approaches are susceptible to challenges such as modeling invalid distance matrices (e.g., negative values) or violations of the triangle inequality. A subset of recent methodologies has effectively managed SE(3)-equivariance without utilizing intermediate variables, through their specifically designed diffusion process (Xu et al. 2022) or loss function (Zhu et al. 2022), leading to improved performance. Yet, the challenge of searching a vast Euclidean space due to the high degree of freedom remains. To effectively narrow down the search space, Jing et

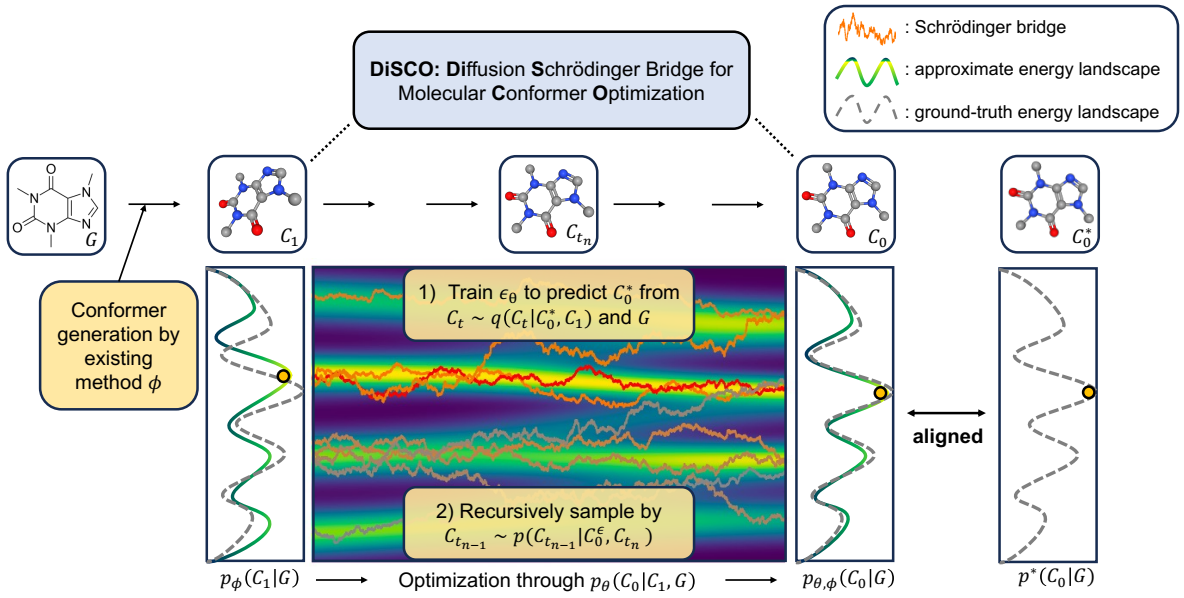

Figure 1: The overview of DiSCO. DiSCO builds a diffusion Schrödinger bridge that aligns the approximate conformational distribution of any existing method p_ϕ (viridis) to the ground-truth distribution p^* (gray).

al. (2022) proposed a potent approach that reduces the degree of freedom by fixing the local structure predicted by RDKit ETKDG (Riniker and Landrum 2015) and only explores the space of torsion angles for generation. Nevertheless, the assumption of their rigidity may result in oversimplified approximation of energy landscapes. As discussed, each existing approach has its own limitation, resulting in an unavoidable gap between the ground-truth and approximate conformational energy landscapes they offer. Our approach builds on the idea that the relaxation of local structures has the potential to produce conformers with more favorable energy profiles.

In this work, we propose a novel diffusion framework designed to refine the approximate conformational landscape inferred by existing methods, through direct modification on the 3D atomic coordinates of conformers. Our framework, named **DiSCO** (**Di**ffusion **S**chrödinger Bridge for Molecular **C**onformer **O**ptimization), builds a Euclidean Schrödinger bridge that connects the approximate distribution of any existing method and the ground-truth conformational distribution. Our approach is inspired by the work of Liu et al. (2023), which proposed a tractable Schrödinger bridge for image-to-image translation tasks. Extending this concept, we develop a conformer-to-conformer Schrödinger bridge dedicated to refining molecular conformers. By constructing an SE(3)-equivariant Schrödinger bridge, DiSCO ensures the roto-translational equivariance of the generated conformers. Since DiSCO is designed to be model-agnostic, it can be integrated as a post hoc refinement stage following any conformer generation method. Its training leverages independent and identically distributed (i.i.d.) samples that originate from the ground-truth conformational energy landscape. Consequently, the refinement precedure of DiSCO can be viewed as the energy optimization of conformers, leading to a distribution that closely resembles the Boltzmann distribution and yields more energetically favorable conformers. Starting with a pre-trained distribution from an established method harnesses valuable prior information about conformational structures, potentially leading to significant reduction in search space.

Through extensive evaluations and analyses, we demonstrate the strengths of our framework in optimizing the conformers generated by existing methods, providing a solid rationale for the utilization of the Schrödinger bridge in the molecular conformer generation task. We provide a graphical overview of our framework in Fig. 1, and summarize our contributions as follows:

- We introduce DiSCO, a novel framework that employs a diffusion Schrödinger bridge to modify the Euclidean coordinates of existing conformers, thus aligning their approximate distribution with the ground-truth energy landscape.
- Within DiSCO, an SE(3)-equivariant Schrödinger bridge is designed to guarantee the roto-translational equivariance in the generated conformers, a fundamental requirement for effective conformer generation.
- DiSCO consistently enhances both the diversity and quality of the conformers generated by existing methods, achieving state-of-the-art benchmark performance and improving molecular property predictions.
- DiSCO exhibits a range of advantageous characteristics including resilience to shifts in quality, an interpretable generative path, and relaxation of local structures, underscoring its broad practical applicability.

Related Work

Molecular Conformer Generation

Computational prediction of low-energy 3D conformation of molecules have been widely studied in both materials science and structure-based drug discovery community. Recently, there has been interest in developing deep learning-based methods, especially generative approaches, for molecular conformation prediction (Xu et al. 2023).

CVGAE (Mansimov et al. 2019) utilized VAE to generate molecular conformers directly within Euclidean space. Since then, subsequent efforts sought to utilize intermediate structures to ensure SE(3)-equivariance (Simm and Hernández-Lobato 2020; Xu et al. 2021a,b; Ganea et al. 2021). ConfGF (Shi et al. 2021) and DGSM (Luo et al. 2021) diverge from these approaches by learning the gradient of the log-density function pertaining to interatomic distances. After an initial estimate of the gradient concerning inter-atomic distances, they deduce the gradient related to atomic coordinates using the chain rule and employs Langevin dynamics to produce Euclidean conformations. GeoDiff (Xu et al. 2022) employs a Euclidean diffusion model that constructs conformers along an SE(3)-equivariant diffusion path. DMCG (Zhu et al. 2022), on the other hand, persistently refines coordinates through tailored graph neural networks and introduces a loss function invariant to both roto-translation of atomic coordinates and permutation of symmetric atoms. Torsional Diffusion (Jing et al. 2022) offers an alternative approach as a diffusion model within torsion space, subsequently reducing the degree of freedom. The model initially employs RDKit ETKDG (Riniker and Landrum 2015) algorithm to produce fixed local structures and then exclusively generates torsion angles via diffusion process. A unique approach of RDKit + Clustering (Zhou et al. 2023) first performs diverse sampling strategies, followed by K-means clustering to represent conformers using cluster centers. Notably, it achieves high diversity comparable to deep learning techniques, without the requirement for any learning procedure. Lastly, the work by Guan et al. (2022) applies SE(3)-equivariant networks to comprehend the gradient fields of the implicit conformational energy landscape. This approach facilitates the optimization of the conformer incrementally, employing the predicted gradient fields to refine the structure systematically.

Diffusion Schrödinger Bridge

The Schrödinger bridge problem represents an entropy-regularized optimal transport problem, aiming to identify the optimal path connecting two arbitrary distributions (Schrödinger 1932; Léonard 2014). Within this context, De Bortoli et al. (2021) introduced diffusion-based techniques to approximate the Iterative Proportional Fitting (IPF) procedure. This method exhibits potential in applications related to generative modeling and data interpolation, offering insights into the solution of the Schrödinger bridge problem.

A recent development by Liu et al. (2023) unveiled a tractable class of Schrödinger bridges, formulated as a convex combination of degraded and clean images. Building on this technique, the proposed Image-to-Image Schrödinger Bridge (I2SB) was applied for solving multitude of image-to-image translation tasks. Our research draws inspiration from this particular study, expanding upon the concept to construct conformer-to-conformer Schrödinger bridges.

Proposed Framework

Notations

In this paper, a 2D molecular graph is denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ symbolizes a set of atoms and $\mathcal{E}$ symbolizes a set of bonds. Within this notation, each node $v_i \in \mathcal{V}$ encapsulates the attributes of the i-th atom, including characteristics such as atom type and chirality, while each edge e_{ij} conveys the type of bond between v_i and v_j (i.e., single, double, triple, aromatic).

A 3D conformer $\mathcal{C} \in \mathbb{R}^{|V|\times 3}$ is represented by a collection of each atom v_i's 3D coordinate $\mathbf{c}_i \in \mathbb{R}^3$. We denote a stochastic process of conformers with $\mathcal{C}_t$, which is indexed by $t \in [0, 1]$. We use $0 = t_0 < \ldots < t_n < \ldots < t_N = 1$ to denote discretized time steps, and $\mathcal{C}_n$ is a shorthand for $\mathcal{C}_{t_n}$.

Problem Definition

Molecular conformer generation can be represented as a conditional generation task with the fundamental aim of producing low-energy conformers given a molecular graph $\mathcal{G}$. Contemporary approaches approximate the ground truth energy landscape through their respective methods denoted by ϕ, and sample low-energy conformers from the resulting distribution $p_\phi(\mathcal{C}|\mathcal{G})$. In reality, however, perfect approximation of the ground truth distribution remains challenging, causing the apparent gap between the ground truth energy landscape $p^*(\mathcal{C}|\mathcal{G})$ and the approximate energy landscape $p_\phi(\mathcal{C}|\mathcal{G})$. We consider the conformers generated by ϕ as latent variable $\mathcal{C}_1$, and using this latent variable $\mathcal{C}_1$, we factorize the ground truth p^* as follows:

$$p^*(\mathcal{C}|\mathcal{G}) = \int_{\mathcal{C}_1} p(\mathcal{C}|\mathcal{C}_1, \mathcal{G}) p_\phi(\mathcal{C}_1|\mathcal{G}) d\mathcal{C}_1. \tag{1}$$

Here, we want to find the conditional distribution $p(\mathcal{C}|\mathcal{C}_1, \mathcal{G})$ that yields the ground truth distribution when marginalized with $\mathcal{C}_1$ from existing methods ϕ. Within this context, $p(\mathcal{C}|\mathcal{C}_1, \mathcal{G})$ can be interpreted as a refinement procedure of $\mathcal{C}_1$ to the ground truth conformer. By achieving a precise approximation of the $p(\mathcal{C}|\mathcal{C}_1, \mathcal{G})$, it is feasible through marginalization to calibrate the learned distribution p_ϕ to align more closely with p^*. Hence, our primary objective is to parameterize and train the conformer optimization model $p_\theta(\mathcal{C}_0|\mathcal{C}_1, \mathcal{G})$, which refines $\mathcal{C}_1$ into a more accurate conformer, given the molecular graph $\mathcal{G}$. The refinement through parameterized model is articulated as:

$$p_{\theta,\phi}(\mathcal{C}_0|\mathcal{G}) = \int_{\mathcal{C}_1} p_\theta(\mathcal{C}_0|\mathcal{C}_1, \mathcal{G}) p_\phi(\mathcal{C}_1|\mathcal{G}) d\mathcal{C}_1. \tag{2}$$

According to Eq. 2, the optimization of conformers sampled from p_ϕ with p_θ is equivalent to the sampling from the refined distribution $p_{\theta,\phi}$. To enable this transformation, we design our conformer optimization model p_θ upon on the

framework of diffusion Schrödinger bridge, which has been utilized for the connection of two arbitrary distributions (De Bortoli et al. 2021; Liu et al. 2023).

Schrödinger Bridge for Molecular Conformer Optimization

Our diffusion process is constructed upon the tractable class of Schrödinger bridge, a concept brought forth by Liu et al. (2023). First, given a conformer generation method ϕ, we define a Schrödinger bridge connecting ground truth conformer distribution p^* with an approximate distribution p_ϕ through a pair of stochastic differential equations (SDEs):

$$\mathrm{d}\mathcal{C}_t = [f_t + \beta_t \nabla \log \Psi(\mathcal{C}_t, t)]\mathrm{d}t + \sqrt{\beta_t}\, \mathrm{d}W_t, \tag{3a}$$

$$\mathrm{d}\mathcal{C}_t = [f_t - \beta_t \nabla \log \widehat{\Psi}(\mathcal{C}_t, t)]\mathrm{d}t + \sqrt{\beta_t}\, \mathrm{d}\bar{W}_t, \tag{3b}$$

where $\mathcal{C}_0^* \sim p^*$ and $\mathcal{C}_1 \sim p_\phi$. The standard Brownian motion and its reversed counterpart are denoted by W_t and $\bar{W}_t$, respectively (Anderson 1982). Time-varying energy potentials are given by the functions $\Psi, \widehat{\Psi} \in C^{2,1}\left(\mathbb{R}^d, [0,1]\right)$, which are solutions to the coupled partial differential equations (PDEs) below:

$$\begin{aligned} &\begin{cases} \frac{\partial \Psi(\mathcal{C},t)}{\partial t} = -\nabla \Psi^\top f - \frac{1}{2}\beta \Delta \Psi \\ \frac{\partial \widehat{\Psi}(\mathcal{C},t)}{\partial t} = -\nabla \cdot (\widehat{\Psi} f) + \frac{1}{2}\beta \Delta \widehat{\Psi} \end{cases} \\ &\text{s.t.} \quad \Psi(\mathcal{C},0)\widehat{\Psi}(\mathcal{C},0) = p^*(\mathcal{C}), \\ &\qquad\quad \Psi(\mathcal{C},1)\widehat{\Psi}(\mathcal{C},1) = p_\phi(\mathcal{C}) \end{aligned} \tag{4}$$

It is known that the paths of the SDEs (Eq. 3) are equivalent. Our ultimate objective is to learn the path dictated by the SDE (Eq. 3) and adopt it as an optimization model, facilitating its application to the conformers derived from p_ϕ. Recently, the Schrödinger bridge problem has been approached through iterative projection methods (Chen, Georgiou, and Pavon 2021) to overcome the coupling constraints in Eq. 4. However, in our work, we leverage a recently proposed tractable Schrödinger bridge (Liu et al. 2023) that bypasses the need for any iterative fitting procedure.

Tractable Schrödinger Bridge According to the insights provided by Liu et al. (2023), the specific diffusion process shares the same marginal density with the SDEs in Eq. 3 can be constructed without the need to directly solve the SDEs, provided that the boundary pairs are accessible during training. This diffusion process is modeled as the convex combination of the associated boundary pairs. A more extensive explanation can be found in the Appendix.

In our case, the boundary pair is $(\mathcal{C}_0^*, \mathcal{C}_1)$ where $\mathcal{C}_0^* \sim p^*(\cdot|\mathcal{G})$ represents the ground truth conformer and $\mathcal{C}_1 \sim p_\phi(\cdot|\mathcal{G})$ is the conformer generated by the existing method. Given this setting, the posterior of SDE (3) associated with the boundary pair can be analytically derived as:

$$\begin{aligned} &q(\mathcal{C}_t|\mathcal{C}_0^*, \mathcal{C}_1) = \mathcal{N}(\mathcal{C}_t; \mu_t\left(\mathcal{C}_0^*, \mathcal{C}_1\right), \Sigma_t), \\ &\mu_t = \frac{\bar{\sigma}_t^2}{\bar{\sigma}_t^2 + \sigma_t^2}\mathcal{C}_0^* + \frac{\sigma_t^2}{\bar{\sigma}_t^2 + \sigma_t^2}\mathcal{C}_1, \quad \Sigma_t = \frac{\sigma_t^2 \bar{\sigma}_t^2}{\bar{\sigma}_t^2 + \sigma_t^2} \cdot I, \end{aligned} \tag{5}$$

where $\sigma_t^2 := \int_0^t \beta_\tau \mathrm{d}\tau$ and $\bar{\sigma}_t^2 := \int_t^1 \beta_\tau \mathrm{d}\tau$. Moreover, above posterior marginalizes the recursive posterior sampling in denoising diffusion probabilistic model (DDPM) (Nichol and Dhariwal 2021) as follows:

$$q(\mathcal{C}_n|\mathcal{C}_0^*, \mathcal{C}_N) = \int \prod_{k=n}^{N-1} p(\mathcal{C}_k|\mathcal{C}_0^*, \mathcal{C}_{k+1}) d\mathcal{C}_{k+1}. \tag{6}$$

Interestingly, the analytic posterior of Schrödinger bridge (Eq. 3) for the given boundary pair $(\mathcal{C}_0^*, \mathcal{C}_1)$ is equivalent to the marginal density induced by DDPM, assuming $\mathcal{C}_N \sim p_\phi$. Consequently, if we can accurately predict $\mathcal{C}_0^*$ from $\mathcal{C}_n$ at test time, the standard posterior sampling of DDPM can be applied to traverse the path induced by Eq. 3, starting from $\mathcal{C}_N \sim p_\phi$. To this end, we utilize a neural network $\epsilon_\theta(\mathcal{C}_t, \mathcal{G}, t)$, parameterized with SE(3)-equivariant neural networks, and our training objective becomes:

$$\|\epsilon_\theta(\mathcal{C}_t, \mathcal{G}, t) - \frac{\mathcal{C}_t - \mathcal{C}_0^*}{\sigma_t}\| \tag{7}$$

where $\mathcal{C}_t$ is drawn from Eq. 5. This parameterization is similar to that of standard DDPM (Nichol and Dhariwal 2021), where a neural network is trained to predict the noise added to clean samples, rather than predicting clean samples directly. Practically, we stochastically optimize Eq. 7, randomly selecting $\mathcal{C}_0^*$ and $\mathcal{C}_1$ from pre-computed sets of ground-truth and generated conformers for a molecule $\mathcal{G}$.

Once a well-trained ϵ_θ is obtained, we simulate the generative path of the Schrödinger bridge (Eq. 3) by recursively sampling $\mathcal{C}_{n-1} \sim p(\mathcal{C}_{n-1}|\mathcal{C}_0^\epsilon, \mathcal{C}_n)$ from $n = N$ to 1, given $\mathcal{C}_0^\epsilon = \mathcal{C}_n - \sigma_n \cdot \epsilon_\theta(\mathcal{C}_n, \mathcal{G}, n)$. Please refer to Alg. 1 and Alg. 2 for detailed training and generation procedures.

In summary, we present a direct approach to learning the diffusion processes that share marginal densities with (Eq. 3) as a convex combination of conformers from boundary distributions p^* and p_ϕ. Intuitively, the forward diffusion path is a convex combination between the paired data $(\mathcal{C}_0^*, \mathcal{C}_1)$, such that the intermediate conformers $\mathcal{C}_t$ constantly degrade from $\mathcal{C}_0^*$ to $\mathcal{C}_1$. On the other hand, the reverse diffusion path is a convex combination of predicted $\mathcal{C}_0^\epsilon$ and $\mathcal{C}_t$.

Notably, within the stochastic differential equation (SDE) (Eq. 3) and the analytic posterior (Eq. 5), the molecular graph $\mathcal{G}$ is not explicitly considered. Nevertheless, during training, we sample boundary pairs from distributions conditioned on a specific molecule $\mathcal{G}$ and train the neural network ϵ_θ to take into account the molecular graph $\mathcal{G}$. Consequently, our diffusion path is learned to form a direct connection between conformers pertaining to the same molecule $\mathcal{G}$.

SE(3)-Equivariance In order to maintain the roto-translational equivariance of our diffusion process, we follow the approach proposed by Xu et al. (2022), involving the alignment and zero center of mass (CoM). During training, we first translate both the ground truth conformer $\mathcal{C}_0^*$ and the starting conformer $\mathcal{C}_1$ to zero CoM. We then align $\mathcal{C}_0^*$ to $\mathcal{C}_1$ employing the Kabsch alignment algorithm (Kabsch 1976). Subsequent to calculating $\mathcal{C}_t$ using Eq. 5, $\mathcal{C}_0^*$ is once more aligned to $\mathcal{C}_t$. This ensures that the training label $\frac{\mathcal{C}_t - \mathcal{C}_0^*}{\sigma_t}$ remains equivariant to the roto-translation of $\mathcal{C}_t$. Training SE(3)-equivariant networks with equivariant labels enables the equivariant prediction of $\mathcal{C}_0^\epsilon$. Through this process, our diffusion paths, originating from any given $\mathcal{C}_1$, learn to arrive at the aligned ground-truth conformer $\mathcal{C}_0^*$. It is pertinent

Algorithm 1: Training

Input: a set of training molecules $\mathcal{G}$, ground truth p^* and existing conformer generation model p_ϕ
Output: trained ϵ_θ

repeat
  $t \sim \mathcal{U}([0,1])$, $\mathcal{C}_0^* \sim p^*(\mathcal{C}_0|\mathcal{G})$, $\mathcal{C}_1 \sim p_\phi(\mathcal{C}_1|\mathcal{G})$
  Translate $\mathcal{C}_0^*$ and $\mathcal{C}_1$ to zero CoM
  Align $\mathcal{C}_0^*$ to $\mathcal{C}_1$ using Kabsch algorithm
  $\mathcal{C}_t \sim q(\mathcal{C}_t|\mathcal{C}_0^*, \mathcal{C}_1)$ using Eq. 5
  Take gradient descent step on $\epsilon_\theta(\mathcal{C}_t, \mathcal{G}, t)$ using Eq. 7
until converges

Algorithm 2: Generation

Input: molecular graph $\mathcal{G}$, trained ϵ_θ, a predicted conformer $\mathcal{C}_1 \sim p_\phi(\mathcal{C}_1|\mathcal{G})$
Output: an optimized conformer $\mathcal{C}_0 \sim p_{\theta,\phi}(\mathcal{C}_0|\mathcal{C}_1, \mathcal{G})$

Translate $\mathcal{C}_1$ to zero CoM
for $n = N$ to 1 **do**
  Predict $\mathcal{C}_0^\epsilon$ using ϵ_θ
  $\mathcal{C}_{n-1} \sim p(\mathcal{C}_{n-1}|\mathcal{C}_0^\epsilon, \mathcal{C}_n)$ according to DDPM
end for
return $\mathcal{C}_0$

to note that during the test, the only manipulation required for $\mathcal{C}_1$ is the translation to zero CoM, which is a trivial computation.

Contrasting our approach to I2SB We note a key difference between our approach and that of Liu et al. (2023). In the latter's case, the boundary pairs are the clean images $X_0 \sim p_{\mathcal{A}}(X_0)$ and corresponding degraded images modeled by $p_{\mathcal{B}}(X_1|X_0)$. Therefore, Schrödinger bridges were constructed to connect individual clean images X_0 to the associated degraded distributions $p_{\mathcal{B}}(X_1|X_0)$. In contrast, our scenario presents a more complex challenge, as acquiring the conditional distribution $p(\mathcal{C}_1|\mathcal{C}_0^*, \mathcal{G})$ is not straightforward. Hence, we assume that C_0^* and C_1 are independent. This allows $\mathcal{C}_1$ to be simply sampled from $p_\phi(\mathcal{C}_1|\mathcal{G})$, enabling us to construct Schrödinger bridges that connect the entire distribution of ground truth conformers to the distribution of starting conformers. This simplification is realized by assuming independence, a choice that may appear to limit the overall model. However, it provides a foundation that can be further developed. In future research, this model could be enhanced by taking into consideration the distance on the conformational energy landscape, allowing for a more nuanced understanding of the relationships between different conformers and improving the fidelity of the bridges constructed.

Neural Network Architecture

In our approach, the neural network ϵ_θ is parameterized using SE(3)-equivariant networks. This architecture consists of an embedding layer, atomic convolution layers, and fully-connected layers. The initial layers generate atom-specific embeddings, followed by atomic convolution layers that capture spatial relationships and symmetries. The final fully-connected layers convert processed embeddings into 3D coordinate vectors, ensuring direct correspondence to geometric structures. Our network implementation utilizes the e3nn library (Geiger and Smidt 2022) for SE(3)-equivariant operations. Comprehensive details about the specific configurations, hyperparameters, and design choices of our neural network architecture can be found in the Appendix.

Experiments

Dataset

We used GEOM-QM9 and GEOM-Drugs, widely used benchmark datasets in molecular conformer generation. The GEOM-QM9 dataset consists of molecules with nine or fewer atoms, while GEOM-Drugs includes larger drug-like molecules. We follow the same dataset split as Xu et al. (2021) and Shi et al. (2021). Each dataset is made up of 40,000 molecules for training, 5,000 molecules for validation, and 200 molecules for testing. Each molecule has a set of ground-truth conformers, referred to as reference conformers, which have been calculated using CREST (Pracht, Bohle, and Grimme 2020).

Baselines

We selected the following four baseline methods:

- **RDKit ETKDG**: A rapid and efficient algorithm commonly used as the default in RDKit.
- **Torsional Diffusion**: A diffusion model operating in torsion space. We retrained Torsional Diffusion on our dataset split, as it was originally trained on a different split.
- **DMCG**: A graph neural network approach that predicts 3D coordinates directly, thus bypassing the need for internal representations of molecular geometry. We utilized pre-trained checkpoint provided in its official code repository.
- **RDKit + Clustering**: An approach that involves clustering of conformers generated through RDKit.

Experimental Setup

For each molecule in the training and validation dataset, $10K$ conformers are generated by the baseline method, where K represents the number of conformers associated with each molecule. This provides a substantial sample space for model training. After the training of DiSCO, the effectiveness of optimization is assessed using two critical aspects: the ensemble root mean square deviation (RMSD) and the ensemble property.

Ensemble RMSD The ensemble RMSD is a widely recognized metric in molecular conformer generation tasks, providing insights into both the diversity and quality of the generated conformer set relative to a reference set.

For the evaluation of ensemble RMSD, $2K$ conformers are generated by the baseline method and subsequently optimized by DiSCO. Diversity is measured using recall-based metrics (COV-R and MAT-R), whereas quality is assessed through precision-based metrics (COV-P and MAT-P). A

comprehensive discussion of these RMSD metrics is available in the Appendix.

Ensemble Property Prediction We also evaluate the ensemble property prediction by contrasting the predicted property values of the reference conformer set against those of the generated set. A subset of 30 molecules from GEOM-QM9 is used, and 50 conformers are generated for each. The xTB package (Bannwarth, Ehlert, and Grimme 2019) is employed to compute properties such as energy and HOMO-LUMO gap, following the evaluation pipeline from Jing et al. (2022).

Hyperparameter Selection The selection of DiSCO's hyperparameters is guided by the harmonic mean of the ensemble RMSD metrics on the validation set. The search space includes the number of diffusion steps (options: 5, 7, 10, or 15), noise scheduling (quadratic or sigmoid) and the number of training epochs (100 or 250). A complete description of the hyperparameter settings and the specifics of noise scheduling can be found in the Appendix.

Experimental Results

DiSCO successfully optimizes the conformers generated by four baseline methods We have thoroughly examined the performance of DiSCO in conformer optimization using two datasets, GEOM-QM9 and GEOM-Drugs. Initially, we trained DiSCO with conformers generated from each of the four baseline methods, respectively, and measured the ensemble RMSD metrics for each set of conformers before and after DiSCO optimization. The results for the GEOM-Drugs dataset (Table 1) show the improvement in most metrics after optimization with DiSCO. Notably, both precision and recall (measuring quality and diversity, respectively) increased for all baseline methods, demonstrating DiSCO's capability to enhance the diversity and quality of the generated conformers. This robust optimization capability is applicable to conformers from any existing method, resulting in state-of-the-art performance across most metrics except for the mean COV-R. Similar results on the GEOM-QM9 dataset support our findings, with detailed comparisons provided in the Appendix.

DiSCO demonstrates robustness to changes in $\mathcal{C}_1$ quality We conducted experiments to assess the impact of initial conformer quality on the DiSCO's performance and its behavior in handling unexpected quality shifts during test time. To study this aspect, we designed three distinct quality alteration strategies following conformer generation by Torsional Diffusion:

1. Adding Gaussian noise with varying scales to input conformers' coordinates,
2. Modulating the number of diffusion steps in Torsional Diffusion before passing the resulting conformers to DiSCO,
3. Employing MMFF (Merck Molecular Force Field, Halgren., 1996) with varying iteration counts. Greater MMFF iterations enhance conformational stability through force field-based computations.

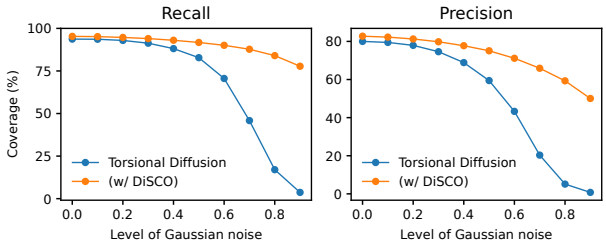

Figure 2: Mean coverage for recall (*left*) and precision (*right*) with the addition of varying levels (in standard deviation) of Gaussian noise to the output of Torsional Diffusion.

These alterations were exclusively introduced during the test phase, allowing us to simulate scenarios where unexpected quality shifts might arise in practical usage. We note that model training utilized standard-quality conformers generated by Torsional Diffusion.

Fig. 2 illustrates DiSCO's optimization performance on conformers perturbed with different levels of Gaussian noise, consistently enhancing both the diversity and quality across all noise levels. This improvement becomes more pronounced with increasing noise, showcasing the robustness of our approach. These outcomes indicate that our optimization model, denote as p_θ, possesses the ability to offset errors in the baseline model p_ϕ, even under practical conditions. It can adapt and improve despite unexpected alterations in conformer quality. Further details on the other two strategies (diffusion step modulation and MMFF iterations) are available in the Appendix. Together, these results reinforce that DiSCO is not only robust to variations in input quality but can also adapt to unanticipated challenges in the quality of conformers, making it a resilient and practical framework for real-world applications.

Intermediate conformers within our diffusion path are also valid We conducted an in-depth analysis to ascertain the effectiveness of the Schrödinger bridge within our DiSCO framework, particularly focusing on the quality of intermediate conformers produced along the generative path. The assessment was initiated by evaluating the conformers generated during the progressive timesteps of DiSCO, commencing with Torsional Diffusion. A set of relevant metrics, including ensemble RMSD and others, was utilized to characterize these intermediate conformers.

The evolution of the average COV-R and COV-P is illustrated in Fig. 3a and b, respectively. A consistent trend of improvement across all metrics (trends of other metrics are provided in the Appendix) was evident throughout the diffusion process. This pattern signifies that DiSCO's intermediate conformers consistently outperformed the initial conformers in terms of RMSD. Moreover, the mean absolute error of energy in comparison to the reference conformer set (calculated by xTB) exhibited a continual decline (Fig. 3c). This evidence further underlines that the intermediate conformers are not only numerically accurate but also possess chemical significance. From a broader perspective, the

Models	COV-R (%) ↑		MAT-R (Å) ↓		COV-P (%) ↑		MAT-P (Å) ↓	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
CONFVAE	55.20	59.43	1.2380	1.1417	22.96	14.05	1.8287	1.8159
CONFGF	62.15	70.93	1.1629	1.1596	23.42	15.52	1.7219	1.6863
GEODIFF	89.13	97.88	0.8629	0.8529	61.47	64.55	1.1712	1.1232
RDKit ETKDG	59.45	60.88	1.2242	1.1653	72.06	89.94	1.0966	0.9528
(w/ DiSCO)	77.22	86.73	0.9638	0.9100	77.17	**94.32**	1.0046	0.9045
Torsional Diffusion	93.71	**100.00**	0.7008	0.6667	80.00	90.98	0.9221	0.8746
(w/ DiSCO)	95.31	**100.00**	**0.6496**	**0.6163**	82.73	93.33	**0.8688**	**0.8097**
DMCG	**96.76**	**100.00**	0.7183	0.7169	79.37	85.13	0.9422	0.9027
(w/ DiSCO)	96.10	**100.00**	0.7119	0.7123	**83.31**	91.13	0.8853	0.8346
RDKit + Clustering	88.13	**100.00**	0.8082	0.7813	69.46	77.08	1.0764	0.9881
(w/ DiSCO)	94.89	**100.00**	0.7135	0.6905	76.56	84.44	0.9722	0.9118

Table 1: Results on the GEOM-Drugs dataset.

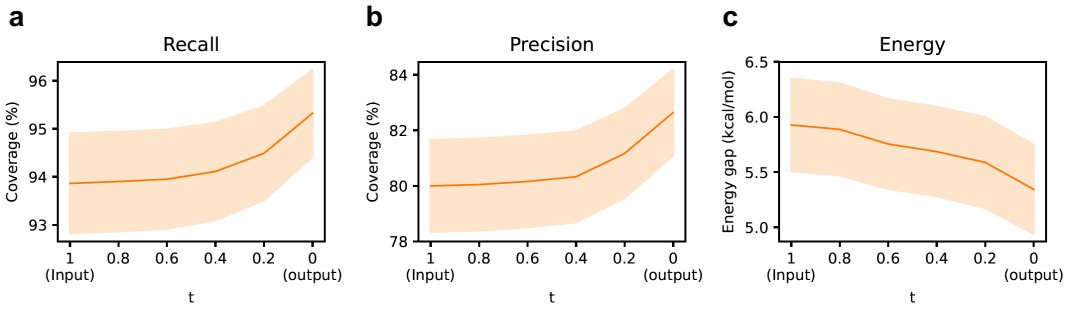

Figure 3: Mean coverage for recall (a), precision (b), and mean energy gap (c) of intermediate conformers in DiSCO. t denotes the timesteps in diffusion path. Shades denote the standard errors.

Method	$\overline{E}$	$E_{\min}$	$\overline{\Delta\epsilon}$	$\Delta\epsilon_{\min}$
RDKit ETKDG	1.27	0.85	1.37	2.62
(w/ DiSCO)	0.79	0.34	1.38	2.39
Torsional Diffusion	1.09	1.40	1.34	4.07
(w/ DiSCO)	0.85	0.33	1.32	3.70

Table 2: MAE of predicted ensemble properties with xTB.

generative path followed by DiSCO can be seen as a systematic and guided exploration for energetically favorable conformers. This exploration, while maintaining proximity to the initial conformers on the conformational energy landscape, ensures that the solutions are both chemically meaningful and accurate. The transparent and interpretable optimization path offered by DiSCO holds the potential to yield significant advantages when extended to other contexts, such as the modeling of molecular docking.

DiSCO improves ensemble property prediction performances We embarked on an ensemble property prediction task to investigate the ability of DiSCO to identify low-energy conformers and improve the prediction of specific molecular properties. We assessed five key molecular properties, quantifying the mean absolute error before and after DiSCO optimization (Table 2). We could only employ two baselines that are capable of generating hydrogen atom coordinates, which is a requirement for exact property computation using xTB. Our findings illustrate a noticeable improvement with DiSCO, particularly in the context of energy predictions. These findings emphasize DiSCO's capability to identify low-energy conformers, which enhances the accuracy of molecular property prediction, and which reinforces its value in tasks demanding precise conformational details and energy estimation.

Conclusion

Through DiSCO, we effectively bridge the gap between generated conformers and the ground-truth conformational landscape. Throughout comprehensive experiments, DiSCO has consistently produced diverse and high-quality conformers across various datasets and baseline methods. Its ability to handle conformers of different qualities and its resilience to unexpected quality shifts further demonstrate its practicality. Additionally, the chemically meaningful intermediate conformers and the improved prediction of molecular properties underscore its value in cheminformatics and drug discovery. The accomplishment of DiSCO unveils new possibilities of diffusion bridges for future endeavors seeking to achieve more precise, interpretable, and effective molecular conformer optimization.

Acknowledgments

This research was supported by the Bio Medical Technology Development Program of the National Research Foundation (NRF), funded by the Ministry of Science ICT (NRF-2022M3E5F3085677, NRF-2022M3E5F3085681, RS-2023-00257479), and Institute of Information communications Technology Planning Evaluation (IITP) grant funded by the Korea government (MSIT) [NO.2021-0-01343, Artificial Intelligence Graduate School Program (Seoul National University)]. The ICT at Seoul National University provided research facilities for this study.

References

Anderson, B. D. 1982. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3): 313–326.

Bannwarth, C.; Ehlert, S.; and Grimme, S. 2019. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *Journal of Chemical Theory and Computation*, 15(3): 1652–1671.

Chen, Y.; Georgiou, T. T.; and Pavon, M. 2021. Stochastic Control Liaisons: Richard Sinkhorn Meets Gaspard Monge on a Schrödinger Bridge. *SIAM Review*, 63(2): 249–313.

De Bortoli, V.; Thornton, J.; Heng, J.; and Doucet, A. 2021. Diffusion Schrödinger Bridge with Applications to Score-Based Generative Modeling. In Beygelzimer, A.; Dauphin, Y.; Liang, P.; and Vaughan, J. W., eds., *Advances in Neural Information Processing Systems*.

Ganea, O.; Pattanaik, L.; Coley, C.; Barzilay, R.; Jensen, K.; Green, W.; and Jaakkola, T. 2021. GeoMol: Torsional Geometric Generation of Molecular 3D Conformer Ensembles. In Ranzato, M.; Beygelzimer, A.; Dauphin, Y.; Liang, P.; and Vaughan, J. W., eds., *Advances in Neural Information Processing Systems*, volume 34, 13757–13769. Curran Associates, Inc.

Geiger, M.; and Smidt, T. 2022. e3nn: Euclidean Neural Networks. arXiv:2207.09453.

Guan, J.; Qian, W. W.; qiang liu; Ma, W.-Y.; Ma, J.; and Peng, J. 2022. Energy-Inspired Molecular Conformation Optimization. In *International Conference on Learning Representations*.

Halgren, T. A. 1996. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 17(5-6): 490–519.

Hawkins, P. C. D. 2017. Conformation Generation: The State of the Art. *Journal of Chemical Information and Modeling*, 57(8): 1747–1756.

Hawkins, P. C. D.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; and Stahl, M. T. 2010. 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.

Jing, B.; Corso, G.; Chang, J.; Barzilay, R.; and Jaakkola, T. 2022. Torsional Diffusion for Molecular Conformer Generation. In Koyejo, S.; Mohamed, S.; Agarwal, A.; Belgrave, D.; Cho, K.; and Oh, A., eds., *Advances in Neural Information Processing Systems*, volume 35, 24240–24253. Curran Associates, Inc.

Kabsch, W. 1976. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, 32(5): 922–923.

Kanal, I. Y.; Keith, J. A.; and Hutchison, G. R. 2018. A sobering assessment of small-molecule force field methods for low energy conformer predictions. *International Journal of Quantum Chemistry*, 118(5): e25512.

Liu, G.-H.; Vahdat, A.; Huang, D.-A.; Theodorou, E. A.; Nie, W.; and Anandkumar, A. 2023. I^2SB: Image-to-Image Schrödinger Bridge. arXiv:2302.05872.

Luo, S.; Shi, C.; Xu, M.; and Tang, J. 2021. Predicting Molecular Conformation via Dynamic Graph Score Matching. In Ranzato, M.; Beygelzimer, A.; Dauphin, Y.; Liang, P.; and Vaughan, J. W., eds., *Advances in Neural Information Processing Systems*, volume 34, 19784–19795. Curran Associates, Inc.

Léonard, C. 2014. A survey of the Schrödinger problem and some of its connections with optimal transport.

Mansimov, E.; Mahmood, O.; Kang, S.; and Cho, K. 2019. Molecular Geometry Prediction using a Deep Generative Graph Neural Network. *Scientific Reports*, 9(1): 20381.

Nichol, A. Q.; and Dhariwal, P. 2021. Improved Denoising Diffusion Probabilistic Models. In Meila, M.; and Zhang, T., eds., *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, 8162–8171. PMLR.

Noé, F.; Olsson, S.; Köhler, J.; and Wu, H. 2019. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science*, 365(6457): eaaw1147.

Pracht, P.; Bohle, F.; and Grimme, S. 2020. Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Phys. Chem. Chem. Phys.*, 22: 7169–7192.

Riniker, S.; and Landrum, G. A. 2015. Better Informed Distance Geometry: Using What We Know To Improve Conformation Generation. *Journal of Chemical Information and Modeling*, 55(12): 2562–2574.

Schrödinger, E. 1932. Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. *Annales de l'institut Henri Poincaré*, 2(4): 269–310.

Shi, C.; Luo, S.; Xu, M.; and Tang, J. 2021. Learning Gradient Fields for Molecular Conformation Generation. In Meila, M.; and Zhang, T., eds., *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, 9558–9568. PMLR.

Simm, G. N. C.; and Hernández-Lobato, J. M. 2020. A Generative Model for Molecular Distance Geometry. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*,

volume 119 of *Proceedings of Machine Learning Research*, 8949–8958. PMLR.

Xu, C.; Lu, Y.; Deng, X.; and Yu, P. 2023. Prediction of Molecular Conformation using Deep Generative Neural Networks. *Chinese Journal of Chemistry*, n/a(n/a).

Xu, M.; Luo, S.; Bengio, Y.; Peng, J.; and Tang, J. 2021a. Learning Neural Generative Dynamics for Molecular Conformation Generation. *CoRR*, abs/2102.10240.

Xu, M.; Wang, W.; Luo, S.; Shi, C.; Bengio, Y.; Gomez-Bombarelli, R.; and Tang, J. 2021b. An End-to-End Framework for Molecular Conformation Generation via Bilevel Programming. In Meila, M.; and Zhang, T., eds., *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, 11537–11547. PMLR.

Xu, M.; Yu, L.; Song, Y.; Shi, C.; Ermon, S.; and Tang, J. 2022. GeoDiff: A Geometric Diffusion Model for Molecular Conformation Generation. In *International Conference on Learning Representations*.

Zhou, G.; Gao, Z.; Wei, Z.; Zheng, H.; and Ke, G. 2023. Do Deep Learning Methods Really Perform Better in Molecular Conformation Generation? arXiv:2302.07061.

Zhu, J.; Xia, Y.; Liu, C.; Wu, L.; Xie, S.; Wang, Y.; Wang, T.; Qin, T.; Zhou, W.; Li, H.; Liu, H.; and Liu, T.-Y. 2022. Direct Molecular Conformation Generation. arXiv:2202.01356.