# CELLULAR-GUIDED GRAPH GENERATIVE MODEL

**Yiming Huang & Tolga Birdal**
Imperial College London
{y.huang24, t.birdal}@imperial.ac.uk

## ABSTRACT

Graph generation is a critical yet challenging task as empirical analyses require a deep understanding of complex, non-Euclidean structures. Although diffusion models have recently made significant achievements in graph generation, these models typically adapt from the frameworks designed for image generation, making them ill-suited for capturing the topological properties of graphs. In this work, we propose a Cellular-Guided Graph Generative (CG3) model that follows a coarse-to-fine generation curriculum and is guided by cellular information, enabling the progressive generation of authentic graphs with inherent topological structures. Experimental results show that our method is able to generate molecules that lie close to the training distribution yet do not violate the chemical valency rule, demonstrating the effectiveness of CG3 in modelling the higher-order relationships.

## 1 INTRODUCTION

Graphs provide an elegant abstraction for representing complex empirical phenomena (Zeng et al., 2024b). The study of graph generation aims to synthesize graphs that align with the observed distribution. Recently, diffusion-based models have achieved remarkable success in image generation by learning a model to denoise a noisy sample (Ho et al., 2020; Song et al., 2021). With their emergence, applying diffusion models to graphs with complex topological structures has garnered significant scientific interest (Niu et al., 2020; Jo et al., 2022; Vignac et al., 2023).

Despite these advances, existing graph generative models typically inherit the frameworks designed for image generation, which fundamentally limits their ability to capture the intrinsic topological properties of networks. Notably, networks exhibit higher-order structures, such as motifs, simplices, and cells, which capture multi-way interactions and critical topological dependencies beyond pairwise relationships (Battiston et al., 2020; Zeng et al., 2024a; Papamarkou et al., 2024). These structures are vital for representing complex phenomena in domains like molecular graphs, social networks, and protein interactions. However, current methods are ineffective at modelling the topological properties of higher-order systems since learning to denoise the noisy samples does not explicitly preserve the intricate structural dependencies required for generating realistic graphs.

Moreover, the image corrupted by Gaussian noise retains recognizable numerical patterns during the early and middle stages of forward diffusion. By contrast, the graph adjacency matrix quickly degrades into a dense matrix with uniformly distributed entries within a few diffusion steps. In addition, directly applying diffusion-based generative models to graph topology generation by injecting isotropic Gaussian noise to adjacency matrices is harmful as it destroys critical graph properties such as sparsity and connectivity. Therefore, a graph-friendly diffusion process should also retain meaningful intermediate states and trajectories and avoid inappropriate noise addition.

Motivated by these principles and advances in topological deep learning (Hajij et al., 2022; Papamarkou et al., 2024; Huang et al., 2024), we propose the **Cellular-Guided Graph Generative** (CG3) framework, illustrated in Figure 1. CG3 introduces a coarse-to-fine generation curriculum, preserving higher-order topologies throughout the diffusion process to better capture complex graph structures. Specifically, we decompose the graph generation task into manageable sub-tasks, beginning by generating higher-order graph skeletons that capture core structures, which are then refined to include pairwise interactions and finer details, resulting in complete graphs with both topological and semantic fidelity. Additionally, CG3 integrates diffusion bridge and spectral diffusion to ensure effective generation and adherence to the aforementioned graph generation principles. Our theoretical analysis reveals that CG3 converges more rapidly in score matching and achieves sharper reconstruction error bounds than classical approaches, offering strong theoretical support for the proposed framework. Furthermore, our framework promises to enhance interpretability by enabling the analysis of different topological guides' performance in the generation process.

## 2 Cellular-Guided Diffusion Framework

We now present our *Cellular-Guided Graph Generative* (CG3) model, which enhances graph generation by exploiting cell structures. We begin by detailing a coarse-to-fine generation curriculum that incrementally constructs graphs, followed by the introduction of three essential supporting techniques: the diffusion bridge, spectral diffusion, and a denoising model, respectively.

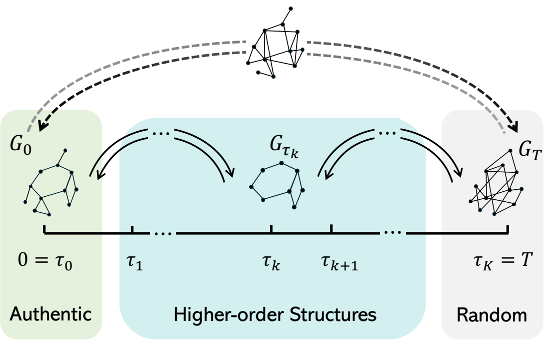


Figure 1: **Overview of CG3.** The dashed line above illustrates the classical generation process, where graphs quickly degrade into random structures. In contrast, as shown in the coloured region below, CG3 adopts a coarse-to-fine generation curriculum based on the diffusion bridge, explicitly learning higher-order structures during intermediate steps.

**Coarse-to-fine Generation.** Drawing inspiration from curriculum learning (Soviany et al., 2022; Xiao & Blanco, 2022; Karami, 2024), we believe that an ideal graph generation curriculum should be composed of multiple easy-to-learn and meaningful intermediate steps. Higher-order structures, especially cells, encapsulate rich structural properties beyond pairwise interactions that are crucial for various empirical systems (Huang et al., 2024; Wu et al., 2024). As a graph-friendly generation framework, CG3 incorporates cell structures during the intermediate stages of forward diffusion and reverse generative processes, thereby realizing a coarse-to-fine generation curriculum.

To implement our coarse-to-fine generation curriculum, we introduce a key operation termed cell complex filtering (CCF). CCF generates an intermediate state of a graph by pruning nodes and edges that do not belong to a given cell complex. This operation plays a crucial role in decomposing the graph generation task into manageable sub-tasks, with the filtered results serving as natural intermediaries in hierarchical graph generation.

The overall framework of CG3 is depicted in Figure 1. Specifically, the forward and reverse diffusion processes in CG3 are divided into $K$ hierarchical time windows, denoted as $\{[\tau_{k-1}, \tau_k]\}_{k=1}^{K}$, where $0 = \tau_0 < \cdots < \tau_{k-1} < \tau_k < \cdots < \tau_K = T$. The process begins by generating coarse-grained higher-order skeletons, which are then refined into finer pairwise relationships, simplifying the modeling of complex graph distributions. Formally, our generation process factorizes the joint distribution of the final graph $\boldsymbol{G}_0$ into a product of conditional distributions across these time windows: $p(\boldsymbol{G}_0) = \prod_{k=1}^{K} p(\boldsymbol{G}_{\tau_{k-1}} \mid \boldsymbol{G}_{\tau_k})$. Here, intermediate states $\boldsymbol{G}_{\tau_k}$ correspond to different levels of cell filtering applied to the graph and the ordering reflects a coarse-to-fine generation process. This approach aligns intermediate stages of the diffusion process with realistic graph representations, reflecting the hierarchical nature of many real-world graphs and enabling smoother training and improved sampling performance.

**Diffusion Bridge Process.** We implement guided diffusion based on diffusion bridge process (Ahmad, 1988; Luo et al., 2023b). Over time, the marginal distribution of the generalized Ornstein-Uhlenbeck (GOU) process stabilizes around a fixed mean and variance, making it well-suited for stochastic modelling with terminal constraints. The GOU process is governed by:

$$\mathrm{d}\boldsymbol{G}_t = \theta_t(\boldsymbol{\mu} - \boldsymbol{G}_t)\mathrm{d}t + g_t(\boldsymbol{G}_t)\mathrm{d}\boldsymbol{W}_t, \tag{1}$$

where $\boldsymbol{\mu} = \boldsymbol{G}_{\tau_k}$ is the target terminal state, $\theta_t$ denotes a scalar drift coefficient and $g_t$ represents the diffusion coefficient. To ensure the process remains analytically tractable, $\theta_t$ and $g_t$ are constrained by the relationship $g_t^2/\theta_t = 2\sigma^2$ (Luo et al., 2023b), where $\sigma^2$ is a given constant scalar. Under these conditions, its transition probability admits a closed-form solution:

$$p(\boldsymbol{G}_t \mid \boldsymbol{G}_s) = \mathcal{N}(\mathbf{m}_{s:t}, v_{s:t}^2\boldsymbol{I}) = \mathcal{N}\left(\boldsymbol{\mu} + (\boldsymbol{G}_s - \boldsymbol{\mu})\, e^{-\bar{\theta}_{s:t}}, \sigma^2(1 - e^{-2\bar{\theta}_{s:t}})\boldsymbol{I}\right). \tag{2}$$

Here, $\bar{\theta}_{s:t} = \int_s^t \theta_z dz$, and for notional simplicity, $\bar{\theta}_{0:t}$ is replaced by $\bar{\theta}_t$ when $s = 0$. At time $t$ progress, $p(\boldsymbol{G}_t)$ gradually approaches a Gaussian distribution characterized by mean $\boldsymbol{\mu}$ and variance $\sigma^2$, indicating that the GOU process exhibits the mean-reverting property.

The Doob's $h$-transform (Doob & Doob, 1984) can modify an SDE such that it passes through a specified endpoint. When applied to the GOU process, this eliminates variance in the terminal state, driving the diffusion toward a Dirac distribution centered at $\boldsymbol{G}_{\tau_k}$ (Heng et al., 2021; Yue et al., 2024).

**Proposition 1.** *Let $\boldsymbol{G}_t$ evolve according to the GOU process in Eq.* (1)*, subject to the terminal conditional $\boldsymbol{\mu} = \boldsymbol{G}_{\tau_k}$. Then, conditional marginal distribution $p(\boldsymbol{G}_t \mid \boldsymbol{G}_{\tau_k})$ evolves according to:*

$$\mathrm{d}\boldsymbol{G}_t = \theta_t \left(1 + \frac{2}{e^{2\bar{\theta}_{t:\tau_k}} - 1}\right)(\boldsymbol{G}_{\tau_k} - \boldsymbol{G}_t)\mathrm{d}t + g_{k,t}\mathrm{d}\boldsymbol{W}_t. \tag{3}$$

*The conditional transition probability has an analytical form as*

$$p(\boldsymbol{G}_t \mid \boldsymbol{G}_{\tau_{k-1}}, \boldsymbol{G}_{\tau_k}) = \mathcal{N}(\bar{\mathbf{m}}_t, \bar{v}_t^2 \boldsymbol{I}) = \mathcal{N}(\boldsymbol{G}_{\tau_k} + (\boldsymbol{G}_{\tau_{k-1}} - \boldsymbol{G}_{\tau_k})e^{-\bar{\theta}_{\tau_{k-1}:t}} \frac{v_{t:\tau_k}^2}{v_{\tau_{k-1}:\tau_k}^2}, \frac{v_{\tau_{k-1}:t}^2 v_{t:\tau_k}^2}{v_{\tau_{k-1}:\tau_k}^2}). \tag{4}$$

We can directly use the closed-form solution in Proposition 1 for one-step forward sampling without performing multi-step forward iteration using the SDE.

**Spectral Diffusion.** Generating graph adjacency matrices presents several significant challenges. Firstly, the non-uniqueness of graph representations means that a graph with $n$ vertices can be equivalently modelled by up to $n!$ distinct adjacency matrices. This ambiguity requires a generative model to assign probabilities uniformly across all equivalent adjacencies. Additionally, sparsity distinguishes graphs from densely distributed image data, causing adjacency score functions to reside on a low-dimensional manifold. Consequently, noise injected into out-of-support regions of the full adjacency space severely impairs the score-matching process. Even for densely connected graphs, isotropic noise distorts global message-passing patterns by encouraging interactions on sparsely connected regions. Moreover, the adjacencies scale quadratically with the number of nodes, making the direct generation of adjacencies computationally prohibitive for large-scale graphs.

To address these challenges, inspired by Martinkus et al. (2022); Luo et al. (2023a), we introduce noise in the eigenvalue domain of the graph Laplacian matrix $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A}$, instead of the adjacency matrix $\boldsymbol{A}$, where $\boldsymbol{D}$ denotes the diagonal degree matrix. As a symmetric positive semi-definite matrix, the graph Laplacian can be diagonalized as $\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top$. Here, the orthogonal matrix $\boldsymbol{U} = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]$ comprises the eigenvectors, and the diagonal matrix $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$ holds the corresponding eigenvalues. Therefore, the target graph distribution $p(\boldsymbol{G}_0)$ represents a joint distribution of $\boldsymbol{X}_0$ and $\boldsymbol{\Lambda}_0$, exploiting the permutation invariance and structural robustness of the Laplacian spectrum. Consequently, we split the forward and reverse SDE into two parts that share drift and diffusion coefficients as

$$\begin{cases} \mathrm{d}\boldsymbol{X}_t = \mathbf{f}_{k,t}(\boldsymbol{X}_t)\,\mathrm{d}t + g_{k,t}\mathrm{d}\boldsymbol{W}_t^1 \\ \mathrm{d}\boldsymbol{\Lambda}_t = \mathbf{f}_{k,t}(\boldsymbol{\Lambda}_t)\,\mathrm{d}t + g_{k,t}\mathrm{d}\boldsymbol{W}_t^2 \end{cases}, \quad \begin{cases} \mathrm{d}\boldsymbol{X}_t = \left[\mathbf{f}_{k,t}(\boldsymbol{X}_t) - g_{k,t}^2 \nabla_{\boldsymbol{X}} \log p_t(\boldsymbol{G}_t | \boldsymbol{G}_{\tau_k})\right] \mathrm{d}\bar{t} + g_{k,t}\mathrm{d}\bar{\boldsymbol{W}}_t^1 \\ \mathrm{d}\boldsymbol{\Lambda}_t = \left[\mathbf{f}_{k,t}(\boldsymbol{\Lambda}_t) - g_{k,t}^2 \nabla_{\boldsymbol{\Lambda}} \log p_t(\boldsymbol{G}_t | \boldsymbol{G}_{\tau_k})\right] \mathrm{d}\bar{t} + g_{k,t}\mathrm{d}\bar{\boldsymbol{W}}_t^2 \end{cases}.$$

Here, the superscript of $\boldsymbol{X}_t^{(k)}$ and $\boldsymbol{\Lambda}_t^{(k)}$ are dropped for simplicity, and $\mathbf{f}_{k,t}$ is defined as in Eq. (3).

To approximate the score functions $\nabla_{\boldsymbol{X}_t} \log p_t(\boldsymbol{G}_t | \boldsymbol{G}_{\tau_k})$ and $\nabla_{\boldsymbol{\Lambda}_t} \log p_t(\boldsymbol{G}_t | \boldsymbol{G}_{\tau_k})$, we employ a neural network $\boldsymbol{s}_{\boldsymbol{\theta}}^{(k)}(\boldsymbol{G}_t, \boldsymbol{G}_{\tau_k}, t)$, composed of a node ($\boldsymbol{s}_{\boldsymbol{\theta},\boldsymbol{X}}^{(k)}$) and a spectrum ($\boldsymbol{s}_{\boldsymbol{\theta},\boldsymbol{\Lambda}}^{(k)}$) output, respectively. The model is optimized by minimizing the loss function:

$$\ell^{(k)}(\boldsymbol{\theta}) = \mathbb{E}_{t,\boldsymbol{G}_t,\boldsymbol{G}_{\tau_{k-1}},\boldsymbol{G}_{\tau_k}} \{\omega(t)[c_1 \| \boldsymbol{s}_{\boldsymbol{\theta},\boldsymbol{X}}^{(k)} - \nabla_{\boldsymbol{X}} \log p_t(\boldsymbol{G}_t | \boldsymbol{G}_{\tau_k})\|_2^2 + c_2 || \boldsymbol{s}_{\boldsymbol{\theta},\boldsymbol{\Lambda}}^{(k)} - \nabla_{\boldsymbol{\Lambda}} \log p_t(\boldsymbol{G}_t | \boldsymbol{G}_{\tau_k})||_2^2]\}, \tag{5}$$

where $\omega(t)$ is a weighting function, and $c_1, c_2$ controls the relative importance.

We sample $(\boldsymbol{X}_{\tau_K}, \boldsymbol{\Lambda}_{\tau_K})$ from the prior distribution and uniformly sample $\boldsymbol{U}_0$ from the observed eigenvector matrices. The generation process involves multi-step diffusion to produce samples $(\hat{\boldsymbol{X}}_{\tau_{K-1}}, \hat{\boldsymbol{\Lambda}}_{\tau_{K-1}}), \cdots, (\hat{\boldsymbol{X}}_{\tau_1}, \hat{\boldsymbol{\Lambda}}_{\tau_1}), (\hat{\boldsymbol{X}}_0, \hat{\boldsymbol{\Lambda}}_0)$ in sequence by reversing the diffusion bridge, where the endpoint of one generation step serves as the starting point for the next. Finally, plausible samples with higher-order structures $\hat{\boldsymbol{G}}_0 = (\hat{\boldsymbol{X}}_0, \hat{\boldsymbol{L}}_0 = \boldsymbol{U}_0 \hat{\boldsymbol{\Lambda}}_0 \boldsymbol{U}_0^\top)$ can be reconstructed.

**Denoising Network Architecture.** We design a neural network $\boldsymbol{s}_{\boldsymbol{\theta}}^{(k)}(\boldsymbol{G}_t, \boldsymbol{G}_{\tau_k}, t)$ to estimate score functions, which comprises two different graph processing modules: a standard graph convolution network (GCN) (Kipf & Welling, 2017) for local feature aggregation and a graph transformer network (ATTN) (Dwivedi & Bresson, 2021; Vignac et al., 2023) for global information extraction. The outputs of these modules are fused with time information through a Feature-wise Linear Modulation (FiLM) layer (Perez et al., 2018). The resulting representations are concatenated to form a unified hidden embedding. This hidden embedding is further processed through separate multilayer perceptrons (MLPs) to produce predictions for $\nabla_{\boldsymbol{X}} \log p(\boldsymbol{G}_t|\boldsymbol{G}_{\tau_k})$ and $\nabla_{\boldsymbol{\Lambda}} \log p(\boldsymbol{G}_t|\boldsymbol{G}_{\tau_k})$, respectively. It is worth noting that our graph noise prediction model is permutation equivalent as each component of our model avoids any node ordering-dependent operations.

## 3 Experiment

To assess the capability of the proposed graph generation method, we conduct evaluations on two well-known molecular datasets: QM9 (Ramakrishnan et al., 2014) and ZINC250k (Irwin et al., 2012), and obtain the intermediate higher-order skeletons using 2-cell complex filtering. We evaluate the quality of 10, 000 generated molecules with three metrics: Neighborhood Subgraph Pairwise Distance Kernel (**NSPDK**) MMD (Costa & Grave, 2010), Fréchet ChemNet Distance (**FCD**) (Preuer et al., 2018) and Validity without correction (**Val. w/o corr.**) (Jo et al., 2022). We compare our model against state-of-the-art molecular generation models, including GraphAF (Shi et al., 2020), GraphDF (Luo et al., 2021), MoFlow (Zang & Wang, 2020), EDP-GNN (Niu et al., 2020), GraphEBM (Liu et al., 2021), GDSS (Jo et al., 2022), and DiGress (Vignac et al., 2023). For a fair comparison, as recommended by Jo et al. (2022), we extend GraphAF and GraphDF to account for formal charges in the molecular generation, termed GraphAF+FC and GraphDF+FC, respectively.

Table 1 indicates that CG3 consistently outperforms both auto-regressive and one-shot models. Notably, the dramatic decrease in NSPDK and FCD implies that CG3 is able to generate molecules with data distributions close to those of the real molecules in both the chemical and graph space. We provide the visualizations of the generated molecules in Appendix B. These results demonstrate CG3 's ability to model the intricate interdependencies between nodes and edges effectively.

Table 1: Comparison of different methods based on molecular datasets.

| Methods | QM9 | | | ZINC250k | | |
|---|---|---|---|---|---|---|
| | NSPDK↓ | FCD↓ | Val. w/o corr.↑ | NSPDK↓ | FCD↓ | Val. w/o corr.↑ |
| GraphAF | 0.020 | 5.268 | 67.00 | 0.044 | 16.289 | 68.00 |
| GraphAF+FC | 0.021 | 5.625 | 74.43 | 0.044 | 16.023 | 68.47 |
| GraphDF | 0.063 | 10.816 | 82.67 | 0.176 | 34.202 | 89.03 |
| GraphDF+FC | 0.064 | 10.928 | 93.88 | 0.177 | 33.546 | 90.61 |
| MoFlow | 0.017 | 4.467 | 91.36 | 0.046 | 20.931 | 63.11 |
| EDP-GNN | 0.005 | 2.680 | 47.52 | 0.049 | 16.737 | 82.97 |
| GraphEBM | 0.003 | 6.143 | 8.22 | 0.212 | 35.471 | 5.29 |
| GDSS | 0.003 | 2.900 | 95.72 | 0.019 | 14.656 | 97.01 |
| DiGress | 0.0005 | 0.360 | **99.00** | 0.082 | 23.060 | 91.02 |
| **CG3 (Ours)** | **0.0003** | **0.172** | 98.74 | **0.001** | **1.533** | **98.56** |

## 4 Conclusion

We introduce a coarse-to-fine graph generation framework, called CG3, that explicitly exploits higher-order topological information. CG3 decomposes the complicated generation process into multiple easy-to-learn generation sub-steps, which are realized with the help of a generalized Ornstein-Uhlenbeck bridge process. Moreover, our model outperforms existing molecular graph generative methods in both graph space and chemical space for distribution learning. Additionally, our framework promises to improve interpretability by enabling the analysis of different topological guides' performance in the generation process. We believe that our work suggests a new formulation of graph generative models.

## ACKNOWLEDGMENTS

T. Birdal acknowledges support from the Engineering and Physical Sciences Research Council [grant EP/X011364/1]. T. Birdal was supported by a UKRI Future Leaders Fellowship [grant number MR/Y018818/1] as well as a Royal Society Research Grant RG/R1/241402.

## REFERENCES

R Ahmad. Introduction to stochastic differential equations. *Journal of the Royal Statistical Society Series C*, 37(3):446–446, 1988.

Federico Battiston, Giulia Cencetti, Iacopo Iacopini, Vito Latora, Maxime Lucas, Alice Patania, Jean-Gabriel Young, and Giovanni Petri. Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports*, 874:1–92, 2020. ISSN 0370-1573.

Fabrizio Costa and Kurt De Grave. Fast neighborhood subgraph pairwise distance kernel. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, pp. 255–262, 2010.

Joseph L Doob and JI Doob. *Classical potential theory and its probabilistic counterpart*, volume 262. Springer, 1984.

Vijay Prakash Dwivedi and Xavier Bresson. A generalization of transformer networks to graphs. In *AAAI Workshop on Deep Learning on Graphs: Methods and Applications*, 2021.

Mustafa Hajij, Ghada Zamzmi, Theodore Papamarkou, Nina Miolane, Aldo Guzmán-Sáenz, Karthikeyan Natesan Ramamurthy, Tolga Birdal, Tamal K Dey, Soham Mukherjee, Shreyas N Samaga, et al. Topological deep learning: Going beyond graph data. *arXiv preprint arXiv:2206.00606*, 2022.

Jeremy Heng, Valentin De Bortoli, Arnaud Doucet, and James Thornton. Simulating diffusion bridges with score matching. *arXiv preprint arXiv:2111.07243*, 2021.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*, volume 33, pp. 6840–6851, 2020.

Yiming Huang, Yujie Zeng, Qiang Wu, and Linyuan Lü. Higher-order graph convolutional network with flower-petals laplacians on simplicial complexes. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pp. 12653–12661, 2024.

John J Irwin, Teague Sterling, Michael M Mysinger, Erin S Bolstad, and Ryan G Coleman. Zinc: a free tool to discover chemistry for biology. *Journal of chemical information and modeling*, 52 (7):1757–1768, 2012.

Jaehyeong Jo, Seul Lee, and Sung Ju Hwang. Score-based generative modeling of graphs via the system of stochastic differential equations. In *International Conference on Machine Learning (ICML)*, pp. 10362–10383. PMLR, 2022.

Mahdi Karami. Higen: Hierarchical graph generative networks. In *Proceedings of International Conference on Learning Representations (ICLR)*, 2024.

Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *Proceedings of International Conference on Learning Representations (ICLR)*, 2017.

Meng Liu, Keqiang Yan, Bora Oztekin, and Shuiwang Ji. Graphebm: Molecular graph generation with energy-based models. In *Energy Based Models Workshop-ICLR*, 2021.

Tianze Luo, Zhanfeng Mo, and Sinno Jialin Pan. Fast graph generation via spectral diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2023a.

Youzhi Luo, Keqiang Yan, and Shuiwang Ji. Graphdf: A discrete flow model for molecular graph generation. In *International Conference on Machine Learning (ICML)*, pp. 7192–7203. PMLR, 2021.

Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B Schön. Image restoration with mean-reverting stochastic differential equations. In *International Conference on Machine Learning (ICML)*, pp. 23045–23066, 2023b.

Karolis Martinkus, Andreas Loukas, Nathanaël Perraudin, and Roger Wattenhofer. Spectre: Spectral conditioning helps to overcome the expressivity limits of one-shot graph generators. In *International Conference on Machine Learning (ICML)*, pp. 15159–15179. PMLR, 2022.

Chenhao Niu, Yang Song, Jiaming Song, Shengjia Zhao, Aditya Grover, and Stefano Ermon. Permutation invariant graph generation via score-based generative modeling. In *International Conference on Artificial Intelligence and Statistics*, pp. 4474–4484. PMLR, 2020.

Theodore Papamarkou, Tolga Birdal, Michael M. Bronstein, Gunnar E. Carlsson, Justin Curry, Yue Gao, Mustafa Hajij, Roland Kwitt, Pietro Lio, Paolo Di Lorenzo, Vasileios Maroulas, Nina Miolane, Farzana Nasrin, Karthikeyan Natesan Ramamurthy, Bastian Rieck, Simone Scardapane, Michael T Schaub, Petar Veličković, Bei Wang, Yusu Wang, Guowei Wei, and Ghada Zamzmi. Position: Topological deep learning is the new frontier for relational learning. In *International Conference on Machine Learning (ICML)*, volume 235, pp. 39529–39555, 2024.

Ethan Perez, Florian Strub, Harm De Vries, Vincent Dumoulin, and Aaron Courville. Film: Visual reasoning with a general conditioning layer. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 32, 2018.

Kristina Preuer, Philipp Renz, Thomas Unterthiner, Sepp Hochreiter, and Gunter Klambauer. Fréchet chemnet distance: a metric for generative models for molecules in drug discovery. *Journal of chemical information and modeling*, 58(9):1736–1741, 2018.

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.

Paavo Salminen. On conditional ornstein–uhlenbeck processes. *Advances in applied probability*, 16 (4):920–922, 1984.

Chence Shi, Minkai Xu, Zhaocheng Zhu, Weinan Zhang, Ming Zhang, and Jian Tang. Graphaf: a flow-based autoregressive model for molecular graph generation. In *Proceedings of International Conference on Learning Representations (ICLR)*, 2020.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *Proceedings of International Conference on Learning Representations (ICLR)*, 2021.

Petru Soviany, Radu Tudor Ionescu, Paolo Rota, and Nicu Sebe. Curriculum learning: A survey. *International Journal of Computer Vision*, 130(6):1526–1565, 2022.

Clément Vignac, Igor Krawczuk, Antoine Siraudin, Bohan Wang, Volkan Cevher, and Pascal Frossard. Digress: Discrete denoising diffusion for graph generation. In *Proceedings of International Conference on Learning Representations (ICLR)*, 2023.

Yingyi Wu, Xinpeng Xie, Chenyang Yu, Zhaomin Xiao, Jinran Zhang, Zhuoer Xu, and Zhelu Mai. A survey on origin-destination flow prediction. In *2024 11th International Conference on Soft Computing & Machine Intelligence*, pp. 48–52, 2024.

Zhaomin Xiao and Eduardo Blanco. Are people located in the places they mention in their tweets? a multimodal approach. In *International Conference on Computational Linguistics*, pp. 2561–2571, 2022.

Conghan Yue, Zhengwei Peng, Junlong Ma, Shiyan Du, Pengxu Wei, and Dongyu Zhang. Image restoration through generalized ornstein-uhlenbeck bridge. In *International Conference on Machine Learning (ICML)*, 2024.

Chengxi Zang and Fei Wang. Moflow: an invertible flow model for generating molecular graphs. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 617–626, 2020.

Yujie Zeng, Yiming Huang, Xiao-Long Ren, and Linyuan Lü. Identifying vital nodes through augmented random walks on higher-order networks. *Information Sciences*, pp. 121067, 2024a.

Yujie Zeng, Yiming Huang, Qiang Wu, and Linyuan Lü. Influential simplices mining via simplicial convolutional networks. *Information Processing & Management*, 61(5):103813, 2024b.

## A Formal Statements and Proofs

In this section, we derive the generalized Ornstein-Uhlenbeck (GOU) bridge process using Doob's $h$-transform (Doob & Doob, 1984).

Recall that the generalized Ornstein-Uhlenbeck (GOU) process is the time-varying OU process. It is a stationary Gaussian-Markov process whose marginal distribution gradually tends towards a stable mean and variance over time. The GOU process $\mathbb{Q}$ is generally defined as follows (Ahmad, 1988; Luo et al., 2023b):

$$\mathbb{Q} : \mathrm{d}\boldsymbol{G}_t = \theta_t \left(\boldsymbol{\mu} - \boldsymbol{G}_t\right) \mathrm{d}t + g_t \mathrm{d}\boldsymbol{W}_t, \tag{6}$$

where $\boldsymbol{\mu}$ is a given state vector, $\theta_t$ denotes a scalar drift coefficient and $g_t$ represents the diffusion coefficient. At the same time, we require $\theta_t, g_t$ to satisfy the specified relationship $2\sigma^2 = g_t^2/\theta_t$, where $\sigma^2$ is a given constant scalar. As a result, its transition probability possesses a closed-form analytical solution:

$$\begin{aligned} p\left(\boldsymbol{G}_t \mid \boldsymbol{G}_s\right) &= \mathcal{N}(\mathbf{m}_{s:t}, v_{s:t}^2 \boldsymbol{I}), \\ \mathbf{m}_{s:t} &= \boldsymbol{\mu} + \left(\boldsymbol{G}_s - \boldsymbol{\mu}\right) e^{-\bar{\theta}_{s:t}}, \\ v_{s:t}^2 &= \sigma^2 \left(1 - e^{-2\bar{\theta}_{s:t}}\right). \end{aligned} \tag{7}$$

Here, $\bar{\theta}_{s:t} = \int_s^t \theta_z dz$. When the starting time $t = 0$, we substitute $\bar{\theta}_{0:t}$ with $\bar{\theta}_t$ for notation simplicity.

**Proposition 1.** *Let $\boldsymbol{G}_t$ evolve according to the generalized OU process in Eq.* (1)*, subject to the terminal conditional $\boldsymbol{\mu} = \boldsymbol{G}_{\tau_k}$. The conditional marginal distribution $p(\boldsymbol{G}_t \mid \boldsymbol{G}_{\tau_k})$ then evolves according to the following SDE:*

$$\mathrm{d}\boldsymbol{G}_t = \theta_t \left(1 + \frac{2}{e^{2\bar{\theta}_{t:\tau_k}} - 1}\right) (\boldsymbol{G}_{\tau_k} - \boldsymbol{G}_t)\mathrm{d}t + g_{k,t}\mathrm{d}\boldsymbol{W}_t. \tag{8}$$

*The conditional transition probability $p(\boldsymbol{G}_t \mid \boldsymbol{G}_{\tau_{k-1}}, \boldsymbol{G}_{\tau_k})$ has analytical form as follows:*

$$\begin{aligned} &p(\boldsymbol{G}_t \mid \boldsymbol{G}_{\tau_{k-1}}, \boldsymbol{G}_{\tau_k}) = \mathcal{N}(\bar{\mathbf{m}}_t, \bar{v}_t^2 \boldsymbol{I}), \\ &\bar{\mathbf{m}}_t = \boldsymbol{G}_{\tau_k} + (\boldsymbol{G}_{\tau_{k-1}} - \boldsymbol{G}_{\tau_k}) e^{-\bar{\theta}_{\tau_{k-1}:t}} \frac{v_{t:\tau_k}^2}{v_{\tau_{k-1}:\tau_k}^2}, \\ &\bar{v}_t^2 = v_{\tau_{k-1}:t}^2 v_{t:\tau_k}^2 / v_{\tau_{k-1}:\tau_k}^2. \end{aligned} \tag{9}$$

*Here, $\bar{\theta}_{a:b} = \int_a^b \theta_s \mathrm{d}s$, and $v_{a:b} = \sigma^2(1 - e^{-2\bar{\theta}_{a:b}})$.*

*Proof.* To simplify the notion, in the $k$-th generation step, we adopt the following conventions: $T = \tau_k$, $\mathbf{x}_t = \boldsymbol{G}_t^{(k)}$, $0 = \tau_{k-1}$, $\mathbf{x}_0 = \boldsymbol{G}_{\tau_{k-1}}$, $\mathbf{x}_T = \boldsymbol{G}_{\tau_k}$.

From Eq. (2), we can derive the following conditional distribution

$$p(\mathbf{x}_T \mid \mathbf{x}_t) = \mathcal{N}(\mathbf{x}_T + (\mathbf{x}_t - \mathbf{x}_T) e^{\bar{\theta}_{t:T}}, v_{t:T}^2 \boldsymbol{I}). \tag{10}$$

According to the definition of Doob's $h$-transform, the $h$-function can be directly computed as:

$$\begin{aligned} \boldsymbol{h}(\mathbf{x}_t, t, \mathbf{x}_T, T) &= \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_T \mid \mathbf{x}_t) \\ &= -\nabla_{\mathbf{x}_t} \left[ \frac{(\mathbf{x}_t - \mathbf{x}_T)^2 e^{-2\bar{\theta}_{t:T}}}{2 v_{t:T}^2} + const \right] \\ &= (\mathbf{x}_T - \mathbf{x}_t) \frac{e^{-2\bar{\theta}_{t:T}}}{v_{t:T}^2} \\ &= (\mathbf{x}_T - \mathbf{x}_t) \sigma^{-2} / (e^{2\bar{\theta}_{t:T}} - 1). \end{aligned} \tag{11}$$

Then the Doob's $h$-transform yields the representation of an endpoint $\mathbf{x}_T$ conditioned process defined by the following SDE:

$$\begin{aligned} \mathrm{d}\mathbf{x}_t &= \left[f(\mathbf{x}_t, t) + g_t^2 \boldsymbol{h}(\mathbf{x}_t, t, \mathbf{x}_T, T)\right] \mathrm{d}t + g_t \mathrm{d}\mathbf{w}_t \\ &= \left(\theta_t + \frac{g_t^2}{\sigma^2 (e^{2\bar{\theta}_{t:T}} - 1)}\right) (\mathbf{x}_T - \mathbf{x}_t)\mathrm{d}t + g_t \mathrm{d}\mathbf{w}_t \\ &= \theta_t \left(1 + \frac{2}{e^{2\bar{\theta}_{t:T}} - 1}\right) (\mathbf{x}_T - \mathbf{x}_t)\mathrm{d}t + g_t \mathrm{d}\mathbf{w}_t. \end{aligned} \tag{12}$$

Given that the joint distribution of $[\mathbf{x}_0, \mathbf{x}_t, \mathbf{x}_T]$ is multivariate normal, the conditional distribution $p(\mathbf{x}_t \mid \mathbf{x}_0, \mathbf{x}_T)$ is also Gaussian:

$$p(\mathbf{x}_t \mid \mathbf{x}_0, \mathbf{x}_T) = \mathcal{N}(\bar{\mathbf{m}}_t, \bar{v}_t^2 \boldsymbol{I}), \tag{13}$$

where the mean $\bar{\mathbf{m}}_t$ and variance $\bar{v}_t^2$ are determined using the conditional formulas for multivariate normal variables:

$$\begin{aligned} \bar{\mathbf{m}}_t &= \mathbb{E}[\mathbf{x}_t \mid \mathbf{x}_0 \mid \mathbf{x}_T] = \mathbb{E}[\mathbf{x}_t \mid \mathbf{x}_0] + \mathrm{Cov}(\mathbf{x}_t, \mathbf{x}_T \mid \mathbf{x}_0)\mathrm{Var}(\mathbf{x}_T \mid \mathbf{x}_0)^{-1}(\mathbf{x}_T - \mathbb{E}[\mathbf{x}_T \mid \mathbf{x}_0]), \\ \bar{v}_t^2 &= \mathrm{Var}(\mathbf{x}_t \mid \mathbf{x}_0 \mid \mathbf{x}_T) = \mathrm{Var}(\mathbf{x}_t \mid \mathbf{x}_0) - \mathrm{Cov}(\mathbf{x}_t, \mathbf{x}_T \mid \mathbf{x}_0)\mathrm{Var}(\mathbf{x}_T \mid \mathbf{x}_0)^{-1}\mathrm{Cov}(\mathbf{x}_T, \mathbf{x}_t \mid \mathbf{x}_0). \end{aligned} \tag{14}$$

Notice that

$$\mathrm{Cov}(\mathbf{x}_t, \mathbf{x}_T \mid \mathbf{x}_0) = \mathrm{Cov}\left(\mathbf{x}_t, (\mathbf{x}_t - \mathbf{x}_T)e^{-\bar{\theta}_{t:T}} \mid \mathbf{x}_0\right) = e^{-\bar{\theta}t:T}\mathrm{Var}(\mathbf{x}_t \mid \mathbf{x}_0). \tag{15}$$

By substituting this and the results in Eq. (2) into Eq. (14), we can obtain

$$\begin{aligned} \bar{\mathbf{m}}_t &= \left(\mathbf{x}_T + (\mathbf{x}_0 - \mathbf{x}_T)e^{-\bar{\theta}_t}\right) + \left(e^{-\bar{\theta}_{t:T}} v_t^2\right)/v_T^2 \cdot \left(\mathbf{x}_T - \mathbf{x}_T - (\mathbf{x}_0 - \mathbf{x}_T)e^{-\bar{\theta}_T}\right) \\ &= \mathbf{x}_T + (\mathbf{x}_0 - \mathbf{x}_T)\left(e^{-\bar{\theta}_t} - e^{-\bar{\theta}_{t:T}} e^{-\bar{\theta}_T} v_t^2/v_T^2\right) \\ &= \mathbf{x}_T + (\mathbf{x}_0 - \mathbf{x}_T)e^{-\bar{\theta}_t}\left(\frac{1 - e^{-2\bar{\theta}_T} - e^{-2\bar{\theta}_{t:T}}(1 - e^{-2\bar{\theta}_t})}{1 - e^{-2\bar{\theta}_T}}\right) \\ &= \mathbf{x}_T + (\mathbf{x}_0 - \mathbf{x}_T)e^{-\bar{\theta}_t} v_{t:T}^2/v_T^2, \end{aligned} \tag{16}$$

and

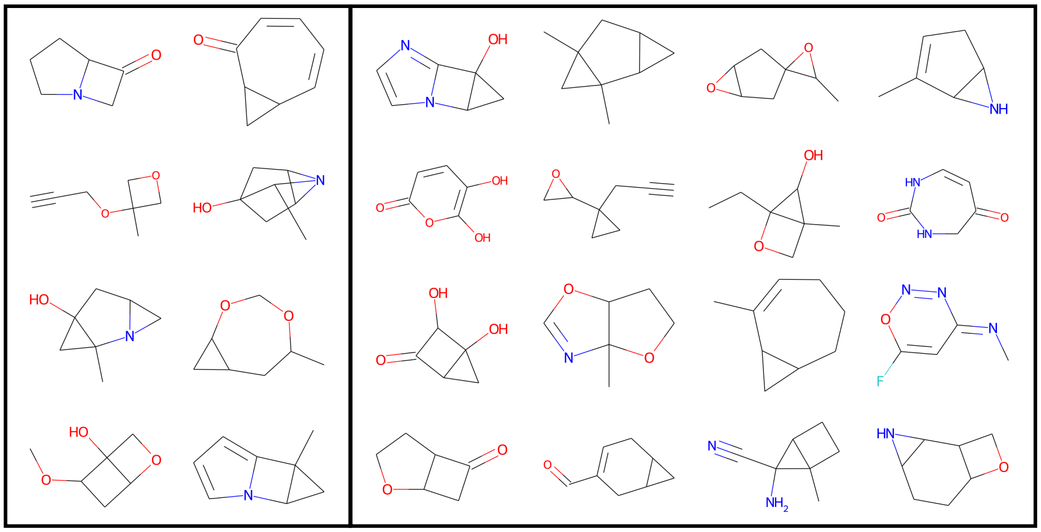
$$\begin{aligned} \bar{v}_t^2 &= v_t^2 - \left(e^{-\bar{\theta}_{t:T}} v_t^2\right)^2 /v_T^2 \\ &= \frac{v_t^2}{v_T^2}(v_T^2 - e^{-2\bar{\theta}_{t:T}} v_t^2) \\ &= \frac{v_t^2}{v_T^2}\sigma^2\left(1 - e^{-2\bar{\theta}_T} - e^{-2\bar{\theta}_{t:T}}(1 - e^{-\bar{2\theta}_t})\right) \\ &= v_t^2 v_{t:T}^2/v_T^2. \end{aligned} \tag{17}$$

Finally, we conclude the proof by reverting to the original notations. □

Note that the generalized OU bridge process, also referred to as the conditional GOU process, has been studied theoretically in previous works (Salminen, 1984; Heng et al., 2021; Yue et al., 2024). However, we are the first to demonstrate its effectiveness in explicitly learning higher-order structures within the graph generation process.

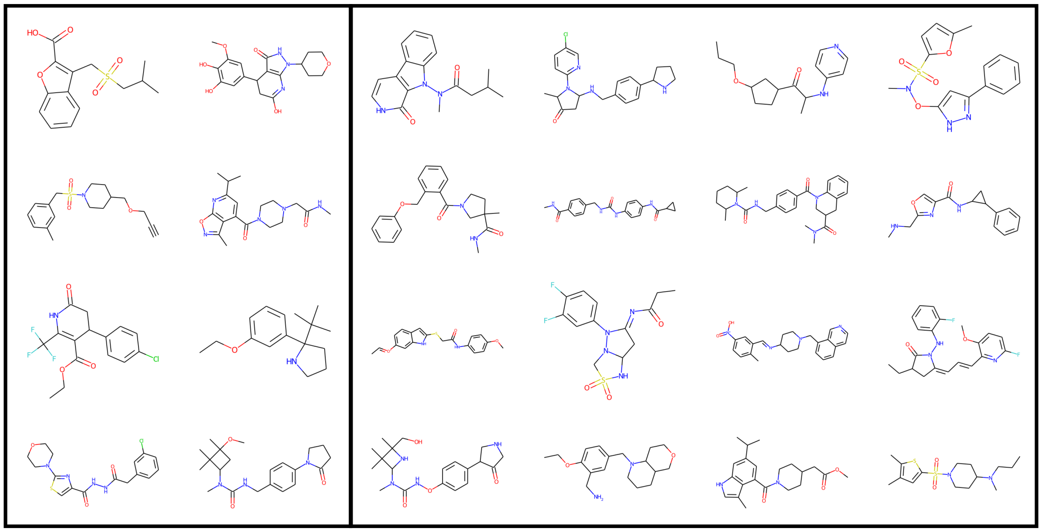
## B Visualization Results

In this section, we additionally provide the visualizations of the generated molecular graphs. Figures 2 and 3 illustrate non-curated samples generated for QM9 and Zinc250k dataset, respectively. It can be observed that CG3 is capable of generating high-quality samples that closely resemble the topological properties of empirical data while preserving essential structural details.



**(a) Training set** **(b) Ours**

Figure 2: Visualization of random samples taken from the CG3 trained on the QM9 dataset.



**(a) Training set** **(b) Ours**

Figure 3: Visualization of random samples taken from the CG3 trained on the Zinc250k dataset.