

Generative Diffusion Models on Graphs: Methods and Applications

Chengyi Liu¹, Wenqi Fan^{1*}, Yunqing Liu¹, Jiatong Li¹,
Hang Li², Hui Liu², Jiliang Tang² and Qing Li¹

¹The Hong Kong Polytechnic University

²Michigan State University

wenqifan03@gmail.com, {chengyi.liu, yunqing617.liu, jiatong.li}connect.polyu.hk,
{lihang4, liuhui7, tangjili}@msu.edu, csqli@comp.polyu.edu.hk

Abstract

Diffusion models, as a novel generative paradigm, have achieved remarkable success in various image generation tasks such as image inpainting, image-to-text translation, and video generation. Graph generation is a crucial computational task on graphs with numerous real-world applications. It aims to learn the distribution of given graphs and then generate new graphs. Given the great success of diffusion models in image generation, increasing efforts have been made to leverage these techniques to advance graph generation in recent years. In this paper, we first provide a comprehensive overview of generative diffusion models on graphs. In particular, we review representative algorithms for three variants of graph diffusion models, i.e., Score Matching with Langevin Dynamics (SMLD), Denoising Diffusion Probabilistic Model (DDPM), and Score-based Generative Model (SGM). Then, we summarize the major applications of generative diffusion models on graphs with a specific focus on molecule and protein modeling. Finally, we discuss promising directions in generative diffusion models on graph-structured data.

1 Introduction

Graphs can represent the rich variety of relationships (*i.e.*, edges) between real-world entities (*i.e.*, nodes). They have been widely used in a diversity of domains [Xia *et al.*, 2021], such as social networks [Derr *et al.*, 2020; Fan *et al.*, 2019a], molecular graph structure [Wu *et al.*, 2022b], and recommender systems [Fan *et al.*, 2022a; Fan *et al.*, 2020], aiming to model association information and structural patterns among various real-world objects [Barabási, 2013]. Due to the prevalence of graphs, graph generative models, with the goal of learning the given graph distributions and generating novel graphs, have attracted significant attention in various applications [Zhang *et al.*, 2020], such as drug discovery and semantic parsing in NLP. Typically, most existing graph generative methods can be classified into two groups: autoregressive generation and one-shot generation [Jo *et al.*, 2022]. Particularly, autoregressive generation methods are designed

to generate desired graphs in a sequential process, while one-shot generation methods generate the entire graph with topology structure and node/edge feature in one single step. In general, graph generation faces three fundamental challenges - (1) *Discreteness*: The graph structure is naturally discrete, resulting in calculation difficulties of models' gradients [Guo and Zhao, 2022; Zhang *et al.*, 2020]. To this end, the most widely used optimization approaches cannot be directly applied to perform back-propagation training for graph generation in an end-to-end fashion. (2) *Complex Intrinsic Dependencies*: Unlike image data, nodes are not independent and identically distributed (or i.i.d.). In other words, these complex graph structural dependencies are inherent relationships among instances (e.g., nodes and edges) [Niu *et al.*, 2020; Guo and Zhao, 2022]. Such complexity of graph structure introduces tremendous challenges in generating desired graphs. (3) *Permutation Invariant*: Since nodes are naturally unordered in most graphs, there are at most $N!$ different equivalent adjacency matrices representing the same graph with N nodes [Niu *et al.*, 2020].

Traditional graph generation methods rely on leveraging hand-crafted graph statistics (e.g., degrees and clustering coefficients properties), and learning kernel functions or engineered features to model the structural information [Xia *et al.*, 2021]. Driven by recent advances in Deep Neural Networks (DNNs) techniques, deep generative models, such as variational autoencoder (VAE) [Simonovsky and Komodakis, 2018], Generative Adversarial Networks (GAN) [De Cao and Kipf, 2018], and normalizing flows [Luo *et al.*, 2021b], have largely improved the generation performance for graph-structured data. For example, GraphVAE trains two graph neural networks (GNNs) as an encoder and a decoder to estimate the graph distribution [Simonovsky and Komodakis, 2018], and MolGAN introduces a GAN-based framework for molecular generation [De Cao and Kipf, 2018]. Although these deep generative methods have achieved promising results, most of them still have several limitations. For example, VAE approaches struggle with the estimation of posterior to generate realistic large-scale graphs and require expensive computation to achieve permutation invariance because of the likelihood-based method [Bond-Taylor *et al.*, 2021]. GAN methods are more prone to mode collapse with graph-structured data and require additional computation to train a discriminator [De Cao and Kipf, 2018; Wang *et al.*, 2018]. The flow-based generative models are hard to fully learn the structural information of graphs be-

*Wenqi Fan is the corresponding author.

cause of the constraints on the specialized architectures [Cornish *et al.*, 2020]. Thus, it is desirable to have a novel generative paradigm for deep generation techniques on graphs.

In recent years, denoising diffusion models have become an emerging generative paradigm to enhance generative capabilities in the image domain [Cao *et al.*, 2022; Yang *et al.*, 2022]. More specifically, inspired by the theory of non-equilibrium thermodynamics, the diffusion generative paradigm can be modelled as Markov chains trained with variational inference [Yang *et al.*, 2022], consisting of two main stages, namely, a forward diffusion and a reverse diffusion. The main idea is that they first develop a noise model to perturb the original input data by adding noise (i.e., generally Gaussian noise) and then train a learnable reverse process to recover the original input data from the noise. Enhanced by the solid theoretical foundation, the probabilistic parameters of the diffusion models are easy-to-tractable, making tremendous success in a wide range of tasks [Cao *et al.*, 2022; Yang *et al.*, 2022].

Recent surveys on deep diffusion models have focused on the image domain [Cao *et al.*, 2022; Yang *et al.*, 2022; Croitoru *et al.*, 2023]. Therefore, in this survey, we provide a comprehensive overview of the advanced techniques of deep graph diffusion models. More specifically, we first briefly introduce the basic ideas of the deep generative models on graphs along with three main paradigms in diffusion models and their key applications i.e., molecule generation and protein modeling. At last, we discuss the future research directions for diffusion models on graphs. To the best of our knowledge, this survey is the very first to summarize the literature in this novel and fast-developing research area.

2 Preliminaries

The general architecture of these deep generative models on graphs is illustrated in Figure 1. In general, a graph is represented as $\mathbf{G} = (\mathbf{X}, \mathbf{A})$, consisting of N nodes. $\mathbf{A} \in \mathbb{R}^{N \times N}$ is the adjacency matrix, where $\mathbf{A}_{ij} = 1$ when node v_i and v_j are connected, and 0 otherwise. $\mathbf{X} \in \mathbb{R}^{N \times d}$ denotes the node feature with dimension d . Under diffusion context, \mathbf{G}_0 refers to the original input graph, while \mathbf{G}_t refers to the noise graph at the t time step.

2.1 Deep Generative Models on Graphs

Variational Autoencoders (VAEs). As the very first deep generative model, variational autoencoders have been successfully applied to graphs, where VAE aims to train a probabilistic graph encoder $q_\phi(\mathbf{z}|\mathbf{G})$ to map the graph space to a low-dimensional continuous embedding \mathbf{z} , and a graph decoder $p_\theta(\mathbf{G}|\mathbf{z})$ to reconstruct new data given the sampling from \mathbf{z} [Simonovsky and Komodakis, 2018].

Generative Adversarial Networks (GAN). GAN is to implicitly learn the graph data distribution with the min-max game theory [Wang *et al.*, 2018] with two neural networks: generator and discriminator. Specifically, the generator attempts to learn the graph distribution and generate new graphs, while the discriminator tries to distinguish the real graph from the generated graph. Due to the discrete nature of graphs, most GAN-based methods are optimized by reinforcement learning techniques.

Normalizing Flows. The normalizing flow leverages a sequence of invertible functions $f(\mathbf{x})$ to map the graph sam-

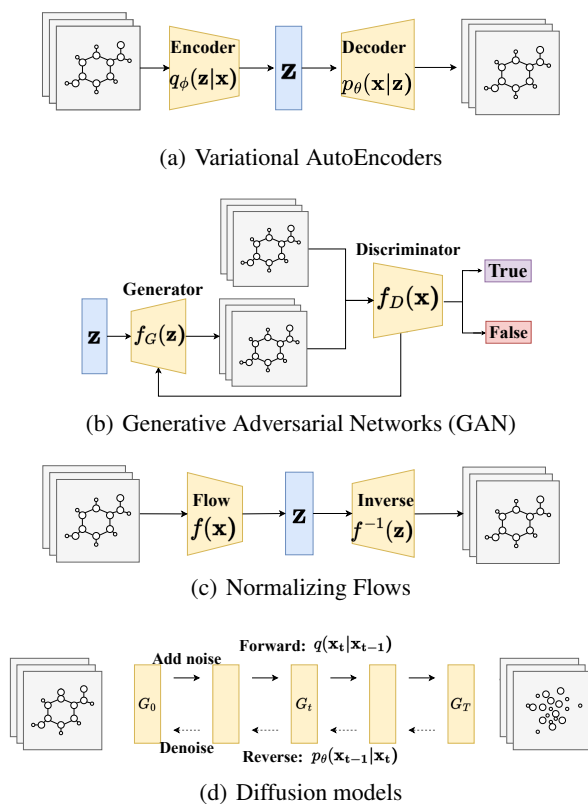


Figure 1: Deep Generative Models on Graphs.

ples (i.e., adjacency matrices and/or edge features) to latent variables \mathbf{z} and learns the graph distribution by tracking the change of density with Jacobian matrix [Liu *et al.*, 2019]. The inverse function $f^{-1}(\mathbf{z})$ yields new samples from latent variables by reversing mapping $f(\mathbf{x})$. The function f specifies an expressive bijective map which supports a tractable computation of the Jacobian determinant [Cornish *et al.*, 2020]. Generally, the training process would estimate the log-likelihoods of each graph sample and update the parameter of $f^{-1}(\mathbf{z})$ by maximizing log-likelihoods with gradient descent.

Limitations. Despite the great success, most existing deep generative models are still facing challenges in graph generation. For instance, VAE models generate graphs based on likelihood, which requires a massive graph-matching process or an explicit estimation of the likelihood of each possible node alignment when achieving permutation invariant [Simonovsky and Komodakis, 2018]. In practice, GAN-based generative models on graphs easily fall into mode collapse, which can limit both the scale and novelty of generated graphs [De Cao and Kipf, 2018]. As for normalizing flow, the bijective model structure limits its ability to capture large-scale node-edge dependencies [Cornish *et al.*, 2020].

2.2 Diffusion Models

In general, there are three paradigms of diffusion models: *Score Matching with Langevin Dynamics (SMLD)*, *Denoising Diffusion Probabilistic Model (DDPM)* and *Score-based Generative Model (SGM)*. SMLD and DDPM leverage the score matching idea and nonequilibrium thermodynamics respectively to learn different reverse functions of the diffusion

process. SGM generalizes the discrete diffusion steps into the continuous scenario and further models the diffusion process with the Stochastic Differential Equations (SDE).

SMLD. As the first representative version of the diffusion model, SMLD [Song and Ermon, 2019] proposes a novel generative model mechanism that first progressively adds random noise to the data distribution to a predefined prior (usually Gaussian noise), and then reverses the diffusion process by learning the gradient of the data distribution $\nabla_{\mathbf{x}} \log p(\mathbf{x})$. The SMLD perturbs the original distribution with a sequence of random Gaussian noises of incremental scales that can be modelled as $q_{\alpha}(\tilde{\mathbf{x}}|\mathbf{x}) := \mathcal{N}(\tilde{\mathbf{x}}|\mathbf{x}, \sigma^2 I)$. This noise scheme facilitates an accurate score matching by preventing the noised distribution from a low-dimensional manifold and providing sufficient training data in low data density regions with large-scale noise. In the gradual denoising process, SMLD proposes a Noise Conditional Score Network (NCSN) $s_{\theta}(\mathbf{x}_t, \sigma)$ to approximate the score jointly. With the annealed Langevin dynamics, NCSN is able to yield new samples from the Gaussian distribution as follows:

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \frac{\epsilon}{2} \nabla_{\mathbf{x}} \log p(\mathbf{x}_{t-1}) + \sqrt{\epsilon} \mathbf{z}_t,$$

where $\mathbf{z}_t \sim \mathcal{N}(0, I)$.

DDPM. Enhanced by the variational inference, the denoising diffusion probabilistic model (DDPM) [Ho *et al.*, 2020] constructs two parameterized Markov chains to diffuse the data with predefined noise and reconstruct the desired samples from the noise. In the forward chain, the DDPM gradually perturbs the raw data distribution $\mathbf{x}_0 \sim q(\mathbf{x}_0)$ to converge to the standard Gaussian distribution \mathbf{z}_t under a pre-designed mechanism. Meanwhile, the reverse chain seeks to train a parameterized Gaussian transition kernel to recover the unperturbed data distribution. Mathematically, the forward process q can be defined as follows:

$$\begin{aligned} q(\mathbf{x}_t|\mathbf{x}_{t-1}) &= \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t} \mathbf{x}_{t-1}, \beta_t I), \\ q(\mathbf{x}_{1:T}|\mathbf{x}_0) &= \prod_{t=1}^T q(\mathbf{x}_t|\mathbf{x}_{t-1}), \end{aligned} \quad (1)$$

where $\beta_t \in (0, 1)$ represents the variance of the Gaussian noise added at time step t . With $\alpha_t = 1 - \beta_t$, $\bar{\alpha}_t = \prod_{i=1}^t \alpha_i$, the marginal can be written as:

$$\begin{aligned} q(\mathbf{x}_t|\mathbf{x}_0) &= \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t} \mathbf{x}_0, (1 - \bar{\alpha}_t) I), \\ \mathbf{x}_t &= \sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon, \end{aligned} \quad (2)$$

where ϵ denotes the Gaussian noise. These equations enable the DDPM to sample the noised latent \mathbf{x}_t at an arbitrary step conditioned on \mathbf{x}_0 [Vignac *et al.*, 2023]. The reverse Gaussian transitions p_{θ} parameterized by θ can be defined as:

$$\begin{aligned} p_{\theta}(\mathbf{x}_{0:T}) &= p(\mathbf{x}_T) \prod_{t=1}^T p_{\theta}(\mathbf{x}_{t-1}|\mathbf{x}_t), \\ p_{\theta}(\mathbf{x}_{t-1}|\mathbf{x}_t) &= \mathcal{N}(\mathbf{x}_{t-1}; \mu_{\theta}(\mathbf{x}_t, t), \Sigma_{\theta}(\mathbf{x}_t, t)). \end{aligned} \quad (3)$$

The neural network will be trained to optimize the variational upper bound on negative log-likelihood, which can be estimated by the Monte-Carlo algorithm [Vignac *et al.*, 2023]. As a result, the DDPM would sample from the limit distribution, and then recursively generate samples \mathbf{x}_t using the learned reverse chain.

SGM. The score SDE formula describes the diffusion process in continuous time steps with a standard Wiener process.

The forward diffusion process in infinitesimal time can be formally represented as [Song *et al.*, 2021]:

$$d\mathbf{x} = f(\mathbf{x}, t)dt + g(t)d\mathbf{w}, \quad (4)$$

where \mathbf{w} denotes a standard Wiener process (a.k.a., Brownian motion), and $g(\cdot)$ denotes the diffusion coefficient, which is assumed to be a scalar independent of \mathbf{x} . The reverse-time SDE describes the diffusion process running backwards in time to generate new samples from the known prior \mathbf{x}_T , which is shown as follows:

$$d\mathbf{x} = [f(\mathbf{x}, t) - g(t)^2 \nabla_{\mathbf{x}} \log p_t(\mathbf{x})]dt + g(t)d\bar{\mathbf{w}}. \quad (5)$$

The only unknown information in reverse-time SDE is the score function $\nabla_{\mathbf{x}} \log p(\mathbf{x})$, which can be approximated by a time-dependent score-based model $s_{\theta}(\mathbf{x}, t)$ by optimizing denoising score-matching objective:

$$\mathbb{E}_{t, \mathbf{x}_0, \mathbf{x}_t} \left[\lambda(t) \|s_{\theta}(\mathbf{x}_t, t) - \nabla_{\mathbf{x}_t} \log p_{0t}(\mathbf{x}_t|\mathbf{x}_0)\|^2 \right], \quad (6)$$

where $p_{0t}(\mathbf{x}_t|\mathbf{x}_0)$ denotes the probability distribution of \mathbf{x}_t conditioned on \mathbf{x}_0 , t is uniformly sampled over $[0, T]$, $\mathbf{x}_0 \sim p_0(\mathbf{x})$ and $\mathbf{x}_t \sim p_{0t}(\mathbf{x}_t|\mathbf{x}_0)$. The score-based diffusion model through SDE unifies the SMLD and DDPM into a continuous version.

Furthermore, the probability flow ODE [Song *et al.*, 2021] designs a deterministic reverse-time diffusion process, whose marginal probability density is identical to the SDE formula. This method largely accelerates the sampling process considering that it allows to perform Gaussian sampling at adaptive time intervals with discretization strategy rather than at successive time steps, and thus the number of estimations of score function is reduced. The reverse-time ODE is defined as below:

$$d\mathbf{x} = [f(\mathbf{x}, t) - \frac{1}{2}g(t)^2 \nabla_{\mathbf{x}} \log p_t(\mathbf{x})]dt.$$

3 Generative Diffusion Models on Graphs

3.1 SMLD on Graphs

EDP-GNN [Niu *et al.*, 2020] is the very first score matching based diffusion method for undirected graph generation. Through fitting a neural network to estimate the different scales of Gaussian noise added to the upper triangular segment of the symmetric adjacency matrices, EDP-GNN learns the score function of the adjacency matrices distributions of the graphs. By using a similar annealed Langevin dynamics implementation as SMLD [Song and Ermon, 2019], the adjacency matrices are generated from the sampled Gaussian noise. Inspired by the GIN method [Xu *et al.*, 2019], EDP-GNN also introduces a multi-channel GNN layer to obtain node features with the message-passing mechanism and an MLP output layer including a noise-conditioned term to prevent separately training the score network at each noise scale. ConfGF [Shi *et al.*, 2021] is the first work adapting the SMLD-based diffusion work to the molecular confirmation generation problem. Unlike EDP-GNN, whose target is to generate adjacency matrices, ConfGF focuses on generating atomic coordinates (node feature) \mathbf{R} given the molecular graph \mathbf{G} . Due to the roto-translation equivalent property, ConfGF maps a set of atomic coordinates to a set of interatomic distances l . By injecting the Gaussian noise over l , ConfGF learns the score function of interatomic distance distributions. Similar to EDP-GNN, the score function is later combined with the annealed Langevin dynamics to generate new atomic coordinate samples.

3.2 DDPM on Graphs

The adaptation of denoising diffusion probabilistic models on graphs is mainly focusing on designing the appropriate transition kernel of the Markov chain. The previous diffusion models usually embed the graphs in continuous space, which might lead to structural information loss. Haefeli *et al.* propose a denoising diffusion kernel to discretely perturb the data distribution. At each diffusion step, each row of the graphs' adjacency matrices is encoded in a one-hot manner and multiplied with a double stochastic matrix \mathbf{Q}_t . In the reverse process, the model includes a re-weighted ELBO as the loss function to obtain stable training. With discrete noise, the sampling process is largely accelerated. Furthermore, DiGress [Vignac *et al.*, 2023] extends the DDPM algorithm to generate graphs with categorical node and edge attributes. The conditional probabilities for the noisy graphs can be defined as follows:

$$\begin{aligned} q(\mathbf{G}_t|\mathbf{G}_{t-1}) &= (\mathbf{X}_{t-1}\mathbf{Q}_t^{\mathbf{X}}, \mathbf{E}_{t-1}\mathbf{Q}_t^{\mathbf{E}}), \\ q(\mathbf{G}_t|\mathbf{G}) &= (\mathbf{X}\mathbf{Q}_t^{\mathbf{X}}, \mathbf{E}\mathbf{Q}_t^{\mathbf{E}}), \end{aligned} \quad (7)$$

where $\mathbf{G}_t = (\mathbf{X}_t, \mathbf{E}_t)$ refers to the noisy graph composed of a node attribute matrix \mathbf{X}_t and an edge attribute tensor \mathbf{E}_t . $\mathbf{Q}_t^{\mathbf{X}}$ and $\mathbf{Q}_t^{\mathbf{E}}$ refer to the noise added to the node and edge, respectively. This Markov formulation allows sampling directly at an arbitrary time step without computing the previous steps. In the denoising process, DiGress incorporates the cross-entropy to measure the distance between the predicted distribution and the input graph distribution with respect to node and edge, so as to train the parameterized graph transformer network ϕ_θ . Thus, the modeling of graph distribution is simplified to a sequence of classification. In addition, operating on discrete steps allows DiGress to leverage various graph descriptors, such as spectral features, to guide the diffusion process. Overall, DiGress is capable of yielding realistic large-scale graphs depending on the overall or partial graph. Moreover, the E(3) Equivariant Diffusion Model (EDMs) is able to operate on both continuous and categorical features of the graph by training an equivariant network [Hoogetboom *et al.*, 2022]. The EDMs jointly inject the Gaussian noise to the latent variables $\mathbf{z}_t = [\mathbf{z}_t^{\mathbf{x}}, \mathbf{z}_t^{\mathbf{h}}]$ of nodes coordinates (continuous) \mathbf{x}_i and the other features \mathbf{h} (categorical). As an extension of EDMs, Equivariant Energy-Guided SDE (EEGSDE) [Bao *et al.*, 2023] introduces a novel property prediction network that serves as a guiding mechanism in the generative process. This network is concurrently trained alongside the reverse diffusion process, with its gradient serving as an additional force to enhance the overall performance. Current generative models struggle to effectively capture the complexities of interatomic forces and the presence of numerous local constraints. To address this issue, the proposed approach in MDM [Huang *et al.*, 2022b] utilizes augmented potential interatomic forces and incorporates dual equivariant encoders to effectively encode the varying strengths of interatomic forces. Additionally, a distributional controlling variable is introduced to ensure thorough exploration and enhance generation diversity during each diffusion/reverse step. Although the diffusion method is initially designed for a one-shot generative manner, the GRAPHARM model proposes an autoregressive diffusion model to generate graphs by sequentially predicting each row in the adjacency matrix [Kong *et al.*, 2023]. The GRAPHARM masks nodes and corresponding edges in forward diffusion whose order is determined by a

diffusion ordering network $q_\phi(\sigma|\mathbf{G}_0)$. In the reverse process, the GRAPHARM denoises only one node at each step (i.e., sequentially generates one row in the adjacency matrix) with the help of the graph attention network [Liao *et al.*, 2019].

3.3 SGM on Graphs

Although EDP-GNN develops a score-based generative model to derive the adjacency matrix of the graph, the estimation for the score function depends on the noise scales at the discrete steps, which restricts its capacity to produce large-scale graphs. GraphGDP [Huang *et al.*, 2022a] leverages the variance-preserving SDE to disturb the adjacency matrix to random graphs. In the reverse process, Position-enhanced Graph Score Network (PGSN) incorporates the features of both nodes and edges and graph position information for permutation equivariant score estimation. Notably, GraphGDP defines a transformation in the forward process to associate continuous distribution with discrete graphs, which allows the model to learn additional graph information of intermediate diffusion steps. Moreover, GDSS [Jo *et al.*, 2022] proposes a continuous-time SDE system to model the diffusion process over nodes and edges simultaneously, where Gaussian noise is directly added to the adjacency matrix and node features. The forward diffusion process on the weighted graph \mathbf{G} at each infinitesimal time step can be modelled as:

$$d\mathbf{G}_t = f_t(\mathbf{G}_t)dt + g_t(\mathbf{G}_t)d\mathbf{w}, \quad \mathbf{G}_0 \sim p_{data}, \quad (8)$$

where f_t represents the linear drift coefficient. To reduce the computation in the reverse diffusion process, GDSS introduces a reverse-time SDE system with respect to nodes and edges as:

$$\begin{cases} d\mathbf{X}_t = [f_{1,t}(\mathbf{X}_t) - g_{1,t}^2 \nabla_{\mathbf{X}_t} \log p_t(\mathbf{X}_t, \mathbf{A}_t)]d\tilde{t} + g_{1,t}d\tilde{\mathbf{w}}_1, \\ d\mathbf{A}_t = [f_{2,t}(\mathbf{A}_t) - g_{2,t}^2 \nabla_{\mathbf{A}_t} \log p_t(\mathbf{X}_t, \mathbf{A}_t)]d\tilde{t} + g_{2,t}d\tilde{\mathbf{w}}_2, \end{cases} \quad (9)$$

where $\nabla_{\mathbf{X}_t} \log p_t(\mathbf{X}_t, \mathbf{A}_t)$ and $\nabla_{\mathbf{A}_t} \log p_t(\mathbf{X}_t, \mathbf{A}_t)$ are the partial score functions, indicating the gradients of the joint log-density connecting the adjacency matrix \mathbf{A} and node feature matrix \mathbf{X} . A corresponding object function is proposed to jointly estimate the log density of nodes and edges:

$$\begin{cases} \min_\theta \mathbb{E}_t \left\{ \lambda_1(t) \mathbb{E}_{\mathbf{G}_0} \mathbb{E}_{\mathbf{G}_t|\mathbf{G}_0} \| s_{\theta,t}(\mathbf{G}_t) - \nabla_{\mathbf{X}_t} \log p_{0t}(\mathbf{X}_t|\mathbf{X}_0) \|_2^2 \right\} \\ \min_\phi \mathbb{E}_t \left\{ \lambda_2(t) \mathbb{E}_{\mathbf{G}_0} \mathbb{E}_{\mathbf{G}_t|\mathbf{G}_0} \| s_{\phi,t}(\mathbf{G}_t) - \nabla_{\mathbf{A}_t} \log p_{0t}(\mathbf{A}_t|\mathbf{A}_0) \|_2^2 \right\} \end{cases} \quad (10)$$

where $s_{\theta,t}(\mathbf{G}_t)$ and $s_{\phi,t}(\mathbf{G}_t)$ are the MLP advanced by the graph multi-head attention blocks [Baek *et al.*, 2021a] to learn the long-term relationship. The GDSS approximates the expectation in Eq.(10) with Monte Carlo estimation, which requires fewer computation and sampling steps compared to the Langevin dynamics [Jo *et al.*, 2022]. To further improve the generation quality, GDSS proposes an integrator to correct the estimated partial score by the score-based MCMC estimation. Note that GDSS is the very first diffusion framework that enables the generation of a whole graph based on node-edge dependency. However, the standard diffusion process would eliminate the features of the sparse graphs in a few steps, which may cause the score estimation uninformative in the reverse process. To address such limitation, a Graph Spectral Diffusion Model (GSDM) [Luo *et al.*, 2022b] is introduced to perform the low-rank Gaussian noise insertion, which can gradually perturb the data distribution with less computation consumption while achieving

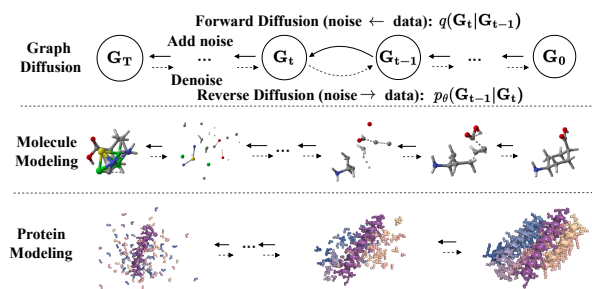


Figure 2: An illustration of diffusion models on molecular and protein modeling. The forward diffusion process involves the gradual addition of noise from the fixed posterior distribution $q(\mathbf{G}_t|\mathbf{G}_{t-1})$ to the input graph \mathbf{G}_0 (representing a molecule or protein) over a period of time T steps, ultimately resulting in the destruction of the molecule or protein structure. In contrast, the reverse diffusion process samples an initial graph \mathbf{G}_T from a standard Gaussian distribution and gradually refines the graph’s structure by using Markov kernels $p_\theta(\mathbf{G}_{t-1}|\mathbf{G}_T)$.

higher generation quality. To be specific, in the diffusion process, GSDM performs spectral decomposition on the adjacency matrix \mathbf{A} to obtain diagonal eigenvalue matrix $\mathbf{\Lambda}$ and eigenvectors \mathbf{U} (i.e., $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$). Meanwhile, since the top- k diagonal entries in $\mathbf{\Lambda}$ can maintain most of the graph information, GSDM conducts the diffusion on the corresponding top- k largest eigenvalues in \mathbf{A} for efficient computation. In addition, SGGM introduces a latent-based generative framework on the graph [Yang *et al.*, 2023], which first encodes the high-dimensional discrete space to low-dimensional topology-injected latent space via a pre-trained variational graph autoencoder and then adopts score-based generative model for the graph generation.

4 Applications

In this section, we aim to review key applications for diffusion models on graphs, such as molecule and protein generations in chemistry and biology domains. The Figure 2 illustrates diffusion models on molecule and protein generation tasks. In Table 1, we summarize these applications.

4.1 Molecule Modeling

The goal of molecule modeling is to employ graph learning techniques for the purpose of representing to better perform downstream tasks, such as molecule generation [Hooeboom *et al.*, 2022]. In general, molecules can be naturally represented as graphs (e.g., atom-bond graphs), in which atoms are represented as nodes and chemical bonds are represented as edges. As such, graph learning techniques can be applied to analyze and manipulate molecular structures for various downstream tasks, such as drug discovery, computational chemistry, materials science, bioinformatics, etc. Furthermore, molecular graph modeling can be used to generate new molecules with desired properties by using advanced graph learning techniques such as VAE [Ma and Zhang, 2021], GNNs [Han *et al.*, 2022], and reinforcement learning (RL) [Zhou *et al.*, 2019]. Particularly, molecule modeling can be further classified into two tasks, namely, molecule conformation generation and molecular docking.

Tasks	Applications	Frame	Representative Methods
Molecule Modeling	Molecule Conformation Generation	SMLD	MDM [Huang <i>et al.</i> , 2022b] GeoDiff [Xu <i>et al.</i> , 2022], EDMs [Hooeboom <i>et al.</i> , 2022], EEGSDE [Bao <i>et al.</i> , 2023], DiGress [Vignac <i>et al.</i> , 2023]
		DDPM	Torsional Diffusion [Jing <i>et al.</i> , 2022], MOOD [Lee <i>et al.</i> , 2022], GDSS [Jo <i>et al.</i> , 2022], DGSM [Luo <i>et al.</i> , 2021a], DiffBridges [Wu <i>et al.</i> , 2022b]
		SGM	
	Molecular Docking	DDPM	FragDiff [Peng <i>et al.</i> , 2023], DiffLink [Igashov <i>et al.</i> , 2022], TargetDiff [Guan <i>et al.</i> , 2023], DiffBP [Lin <i>et al.</i> , 2022]
		SGM	DiffDock [Corso <i>et al.</i> , 2022]
Protein Modeling	Protein Generation	DDPM	SMCDiff [Trippe <i>et al.</i> , 2023], SiamDiff [Zhang <i>et al.</i> , 2021], DiffFold [Wu <i>et al.</i> , 2022a], ProSSDG [Anand and Achim, 2022], DiffAntigen [Luo <i>et al.</i> , 2022a], RFdiffusion [Watson <i>et al.</i> , 2022]
		SGM	ProteinSGM [Luo <i>et al.</i> , 2022a]
	Protein-ligand Complex Structure Prediction	DDPM	DiffEE [Nakata <i>et al.</i> , 2022]
		SGM	NeuralPLexer [Qiao <i>et al.</i> , 2022]

Table 1: A summary of representative applications for generative diffusion method on graphs.

Molecule Conformation Generation

A molecule can be represented by three-dimensional geometry or conformation, in which atoms can be denoted as their Cartesian coordinates. The biological and physical characteristics of the molecule are significantly influenced by its three-dimensional structure. Meanwhile, molecular conformations possess roto-translational invariance. As a result, several techniques avoid directly modeling atomic coordinates by using intermediary geometric variables, such as atomic distances, bond angles, and torsion angles, which also have roto-translational invariance [Shi *et al.*, 2021]. However, as they seek to indirectly model these intermediate geometric variables, they may be subject to limitations in either the training or inference process. To address this issue, GeoDiff [Xu *et al.*, 2022] treats atoms as particles in a thermodynamic system and simulates the diffusion process in nonequilibrium thermodynamics. By learning to reverse the diffusion process, the sampled molecules gradually diffuse backwards into the target conformation. Dynamic Graph Score Matching (DGSM) [Luo *et al.*, 2021a] uses score matching based on a 2D molecular graph to generate the conformation structure of a molecule by modeling the local and long-range interactions of atoms.

Torsional Diffusion [Jing *et al.*, 2022] defines the diffusion over a torus to represent torsion angles, providing a more natural parameterization of conformers. Besides, Torsional Diffusion leverages the torus as a representation of torsion angles for conformations generation. By incorporating a probabilistic analysis that enables probability calculations and atom classification, E(3) Equivariant Diffusion Model (EDMs) [Hooeboom *et al.*, 2022] enables the model to learn the denoising process in continuous coordinates and improve the generation of molecular conformations. As an extension of EDMs, equivariant energy-guided stochastic differential equations (EEGSDE) [Bao *et al.*, 2023] adds energy functions to the model as a guide to fully learn the geometric sym-

metry of molecules to generate 3D molecular conformations.

In addition to adding energy guidance and learning the atomic coordinates of molecules and torsion angles, other domain knowledge can be introduced into the model for enhancing the molecular representation learning. For example, to model interatomic interactions in molecular representation, MDM [Huang *et al.*, 2022b] considers the role of atomic spacing in interatomic forces for molecular representation. More specifically, MDM treats atomic pairs with atomic spacing below a certain threshold as covalent bonds, as chemical bonds govern interatomic forces when atoms are close enough to each other. For atomic pairs with atomic spacing above a certain threshold, the van der Waals force is proposed to govern the interatomic forces. Additionally, to enhance the diversity of molecule generation, they introduce latent variables that are interpreted as control representations in each diffusion/reverse step of the diffusion model. DiffBridges [Wu *et al.*, 2022b] designs an energy function with physical information and a statistical prior for molecule generation. It differs from other methods in its incorporation of physical prior into bridges, as opposed to learning diffusions as a combination of forward-time diffusion bridges.

Transformer architectures have achieved notable success in the fields of computer vision (CV) and natural language processing (NLP), such as ViT [Dosovitskiy *et al.*, 2020], BERT [Kenton and Toutanova, 2019], GPT [Brown *et al.*, 2020]. Similarly, graph transformer [Dwivedi and Bresson, 2021] is incorporated into discrete diffusion model with discrete diffusion [Austin *et al.*, 2021] for graph generation. In particular, DiGress uses graph-based architectures and a noise model that preserves the marginal distribution of node and edge types [Vignac *et al.*, 2023], rather than using uniform noise as a starting point. A denoising network is then augmented with structural features, enabling conditional generation through guidance procedures.

Since most existing molecular generation approaches generate graphs that are likely similar to training samples, Molecular Out-of-distribution Diffusion (MOOD) [Lee *et al.*, 2022] incorporates out-of-distribution (OOD) control in the generative stochastic differential equation (SDE), to generate a new molecule graph that is distinct from those in the training set. GDSS [Jo *et al.*, 2022] learns the underlying distribution of graphs by deriving score matching objectives tailored for the proposed diffusion process, enabling the estimation of the gradient of the joint logdensity w.r.t. each component.

Molecular Docking

Molecular docking is a computational method for predicting the preferred orientation of one molecule to a second molecule (typically a protein) when binding to each other. It's used in drug discovery to find the best fit of a small molecule into the active site of a target protein.

Autoregressive models are widely adopted to generate 3D molecules for the protein binding pocket [Shin *et al.*, 2021]. However, autoregressive models might struggle with capturing complex relationships and interactions between residues in the pocket. To address these challenges, DiffBP [Lin *et al.*, 2022] generates 3D molecular structures in a non-autoregressive manner while satisfying the physical properties of the molecule, based on the protein target as a constraint. Using diffusion models and SE(3)-equivariant networks, TargetDiff [Guan *et al.*, 2023] learns atomic types and

atomic coordinates to generate protein target molecules with satisfying geometric properties.

Fragment-based drug discovery is also a widely adopted paradigm in drug development, which can provide promising solutions for molecular docking by generating 3D molecules fragment-by-fragment and incorporating diffusion models. For instance, FragDiff [Peng *et al.*, 2023] generates 3D molecules fragment-by-fragment for pockets. In each generation step, FragDiff generates a molecular fragment around the pocket. The atom types, atom coordinates and bonds on this fragment are predicted. Then the fragments are gradually joined together to produce the complete molecule. Given some fragments, DiffLink [Igashov *et al.*, 2022] generates the rest of a molecule in 3D. The generator of DiffLink is an E(3) equivariant denoising diffusion model to generate fragments. It is conditioned on the positions of the fragment atoms, and optionally also on the protein pocket that the molecule should fit into. Finally, DiffLink splices these fragments into a complete drug candidate molecule. Similar to the transformation of sentiment classification tasks into generative tasks in the NLP field [Liu *et al.*, 2023], DiffDock [Corso *et al.*, 2022] uses diffusion model to form docking pose prediction problem as a generation problem and executes a reverse diffusion process using separate ligands and proteins as inputs by randomly selecting the initial states and ranking them.

4.2 Protein Modeling

Protein modeling is to generate and predict the structure of proteins. This task is instrumental in comprehending the function and interactions of proteins, and is widely used in the fields of drug discovery and the design of novel proteins with specific characteristics. Previously, proteins have been represented as sequences of amino acids, leading to successes in modeling proteins using language models [Ferruz *et al.*, 2022]. With the advent of diffusion models in image generation [Ramesh *et al.*, 2021], a growing number of applications using diffusion models in protein modeling have emerged.

Protein Generation

The objective of computational protein design is to automate the generation of proteins with specific structural and functional properties. This field has experienced significant advancements in recent decades, including the design of novel 3D folds [Jumper *et al.*, 2021], enzymes [Giessel *et al.*, 2022] and complexes [Réau *et al.*, 2023].

Pre-training protein representations on enormous unlabeled protein structures have drawn more and more interest from researchers. Siamese Diffusion Trajectory Prediction (SiamDiff) [Zhang *et al.*, 2022] obtains counterparts by adding random structural perturbations to natural proteins and diffuses them structurally and sequence-wise through the pre-training process. The diffusion process of the original protein and its counterpart is referred to as two related views. In this process, SiamDiff uses the noise of one view to predict the noise of the other view to better learn the mutual information between the two views. In contrast to pre-training techniques, ProteinSGM [Lee and Kim, 2022] applies conditional generation, paint credible backbones and functional sites into structures of predetermined length to generate proteins. ProSSDG [Anand and Achim, 2022] combines the protein's structure and sequence to generate proteins with the desired 3D structures and chemical properties. Based on a brief

description of the protein’s topology, ProSSDG generates a complete protein configuration.

Despite recent developments in protein structure prediction, it is still difficult to directly generate a variety of unique protein structures using DNNs techniques. DiffFold [Wu *et al.*, 2022a] generates protein backbone structures by imitating natural folding progress. DiffFold develops new structures by denoising from a chaotic, unfolded state towards a stable folded shape. In addition, DiffFold represents protein backbone structure as a sequence of sequential angles representing the relative orientation of the component amino acid residues.

A stable protein backbone that has the motif is referred to as a scaffold. It can greatly benefit from building a scaffold that supports a functional motif. SMCDiff [Trippe *et al.*, 2023] uses a particle filtering algorithm for conditional sampling of protein backbone structures, where priority is given to backbone structures more consistent with the motif.

Immune system proteins (called antibodies) attach to particular antigens like germs and viruses to defend the host. The complementarity-determining regions (CDRs) of the antibodies play a major role in regulating the interaction between antibodies and antigens. DiffAntigen [Luo *et al.*, 2022a] jointly generates the sequence and structure of the CDRs of an antibody, based on the framework region of antibody and the target antigen. DiffAntigen is able to regulate the generation at the antigen structure, not just in the framework region. Additionally, it can also predict the side-chain orientation. RFdiffusion [Watson *et al.*, 2022] combines a diffusion model with a protein prediction model, RoseTTAFold [Baek *et al.*, 2021b]. During the forward diffusion progress, RFdiffusion perturbs the 3D structural coordinates locally to enhance the model’s representational capacity.

Protein-ligand Complex Structure Prediction

The prevalence of protein-ligand complexes makes predicting their 3D structure valuable for generating new enzymes and drug compounds. NeuralPLexer [Qiao *et al.*, 2022] predicts the structure of protein-ligand complexes by combining multi-scale induced bias in biomolecular complexes with diffusion models. It takes molecular graphs as ligand input and samples 3D structures from a learned statistical distribution. To overcome the difficulties of high-dimensional modeling and to extend the range of protein-ligand complexes, DiffEE [Nakata *et al.*, 2022] proposes an end-to-end diffusion generative model, which is based on pre-trained protein language model. DiffEE is able to generate a variety of structures for protein-ligand complexes with correct binding pose.

5 Future Challenges and Opportunities

There are increasing efforts to develop diffusion models on graphs. Next we discuss potential future research directions.

Conditional Generation for Graph Diffusion Models. Incorporating conditions into generative models is critical to guide desired generation. In graph domain, to generate molecules and proteins with specified properties, it is significant to set certain constraints on the design of graph generative models. Thus, introducing extra information as conditions into graph diffusion models has become an imperative research direction. One type of extra knowledge can be formed by a knowledge graph [Chen *et al.*, 2022], which can assist in controlling the generation process to obtain desired graphs, and enhancing the diversity of graph generation.

Trustworthiness for Graph Diffusion Models. Recent years have witnessed growing concerns about AI models’ trustworthiness [Liu *et al.*, 2022; Dai *et al.*, 2022; Chen *et al.*, 2023]. As one of the most representative AI-powered applications, graph generation might cause unintentional harm to users and society in various real-world tasks, especially those in safety-critical fields such as drug discovery. Therefore, how to build trustworthy graph diffusion models has become critical in both academia and industry

Evaluation Metrics. The evaluation of graph generation remains a challenge. Most existing metrics are usually based on graph statistics and properties (e.g., node’ degree and sparsity) [O’Bray *et al.*, 2022], which are not fully trustable. Meanwhile, validity and diversity for graph generation are important in different applications. Thus, efforts are desired to quantitatively measure the quality of generated graphs.

Graph Diffusion Applications. Most existing graph diffusion techniques are used for molecule and protein generation, while many applications on graphs are rarely explored, such as *Recommender Systems*, *Graph Anomaly Detection*, and *Causal Graph Generation*. For example, graph learning techniques have been successfully used for recommendations (i.e., distributions) [Fan *et al.*, 2019b; Fan *et al.*, 2022b]. Thus, diffusion models on graphs have the potential to better generate recommendation lists for users. Anomalies refer to graph objects such as nodes, edges, and sub-graphs. Recent works have shown that diffusion models can be leveraged to purify image data for better adversarial robustness [Xiao *et al.*, 2022]. Thus, graph diffusion models provide great opportunities to improve graph anomaly detection.

6 Conclusion

As one of the most advanced generative techniques, diffusion models have achieved great success in advancing various generative tasks, particularly in the image domain. Similarly, many efforts have been devoted to studying graph generation based on diffusion model techniques. However, it lacks a systematic overview and discussion of the state-of-the-art diffusion models on graphs. To bridge this gap, we provided a comprehensive overview of deep diffusion models on graphs including representative models and applications. We also discussed some promising future research directions.

Acknowledgments

The research described in this paper has been partly supported by NSFC (project no. 62102335), General Research Funds from the Hong Kong Research Grants Council (Project No.: PolyU 15200021 and 15207322), internal research funds from The Hong Kong Polytechnic University (project no. P0036200 and P0042693), Collaborative Project no. P0041282, and SHTM Interdisciplinary Large Grant (project no. P0043302). This research is also supported by the National Science Foundation (NSF) under grant numbers CNS1815636, IIS1845081, IIS1928278, IIS1955285, IIS2212032, IIS2212144, IOS2107215, and IOS2035472, the Army Research Office (ARO) under grant number W911NF-21-1-0198, the Home Depot, Cisco Systems Inc, Amazon Faculty Award, Johnson&Johnson, and SNAP.

References

- [Anand and Achim, 2022] Namrata Anand and Tudor Achim. Protein structure and sequence generation with equivariant denoising diffusion probabilistic models. *arXiv:2205.15019*, 2022.
- [Austin *et al.*, 2021] Jacob Austin, Daniel D Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg. Structured denoising diffusion models in discrete state-spaces. *NeurIPS*, 2021.
- [Baek *et al.*, 2021a] Jinheon Baek, Minki Kang, and Sung Ju Hwang. Accurate learning of graph representations with graph multiset pooling. In *ICLR*, 2021.
- [Baek *et al.*, 2021b] Minkyung Baek, Frank DiMaio, Ivan Anishchenko, Justas Dauparas, Sergey Ovchinnikov, Gyu Rie Lee, Jue Wang, Qian Cong, Lisa N Kinch, R Dustin Schaeffer, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 2021.
- [Bao *et al.*, 2023] Fan Bao, Min Zhao, Zhongkai Hao, Peiyao Li, Chongxuan Li, and Jun Zhu. Equivariant energy-guided sde for inverse molecular design. In *ICLR*, 2023.
- [Barabási, 2013] Albert-László Barabási. Network science. *Philos. Trans. Royal Soc. A*, 2013.
- [Bond-Taylor *et al.*, 2021] Sam Bond-Taylor, Adam Leach, Yang Long, and Chris G Willcocks. Deep generative modelling: A comparative review of vaes, gans, normalizing flows, energy-based and autoregressive models. *IEEE TPAMI*, 2021.
- [Brown *et al.*, 2020] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *NeurIPS*, 2020.
- [Cao *et al.*, 2022] Hanqun Cao, Cheng Tan, Zhangyang Gao, Guangyong Chen, Pheng-Ann Heng, and Stan Z Li. A survey on generative diffusion model. *arXiv:2209.02646*, 2022.
- [Chen *et al.*, 2022] Jingfan Chen, Wenqi Fan, Guanghui Zhu, Xiangyu Zhao, Chunfeng Yuan, Qing Li, and Yihua Huang. Knowledge-enhanced black-box attacks for recommendations. In *KDD*, 2022.
- [Chen *et al.*, 2023] Xiao Chen, Wenqi Fan, Jingfan Chen, Haochen Liu, Zitao Liu, Zhaoxiang Zhang, and Qing Li. Fairly adaptive negative sampling for recommendations. *arXiv:2302.08266*, 2023.
- [Cornish *et al.*, 2020] Rob Cornish, Anthony Caterini, George Deligiannidis, and Arnaud Doucet. Relaxing bijectivity constraints with continuously indexed normalising flows. In *ICML*, 2020.
- [Corso *et al.*, 2022] Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi S Jaakkola. Diffdock: Diffusion steps, twists, and turns for molecular docking. In *NeurIPS Workshop*, 2022.
- [Croitoru *et al.*, 2023] Florinel-Alin Croitoru, Vlad Hondru, Radu Tudor Ionescu, and Mubarak Shah. Diffusion models in vision: A survey. *IEEE TPAMI*, 2023.
- [Dai *et al.*, 2022] Enyan Dai, Tianxiang Zhao, Huaisheng Zhu, Junjie Xu, Zhimeng Guo, Hui Liu, Jiliang Tang, and Suhang Wang. A comprehensive survey on trustworthy graph neural networks: Privacy, robustness, fairness, and explainability. *arXiv:2204.08570*, 2022.
- [De Cao and Kipf, 2018] Nicola De Cao and Thomas Kipf. MolGAN: An implicit generative model for small molecular graphs. *ICML workshop*, 2018.
- [Derr *et al.*, 2020] Tyler Derr, Yao Ma, Wenqi Fan, Xiaorui Liu, Charu Aggarwal, and Jiliang Tang. Epidemic graph convolutional network. In *WSDM*, 2020.
- [Dosovitskiy *et al.*, 2020] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. In *ICLR*, 2020.
- [Dwivedi and Bresson, 2021] Vijay Prakash Dwivedi and Xavier Bresson. A generalization of transformer networks to graphs. *AAAI Workshop*, 2021.
- [Fan *et al.*, 2019a] Wenqi Fan, Tyler Derr, Yao Ma, Jianping Wang, Jiliang Tang, and Qing Li. Deep adversarial social recommendation. In *IJCAI*, 2019.
- [Fan *et al.*, 2019b] Wenqi Fan, Yao Ma, Qing Li, Yuan He, Eric Zhao, Jiliang Tang, and Dawei Yin. Graph neural networks for social recommendation. In *WWW*, 2019.
- [Fan *et al.*, 2020] Wenqi Fan, Yao Ma, Qing Li, Jianping Wang, Guoyong Cai, Jiliang Tang, and Dawei Yin. A graph neural network framework for social recommendations. *IEEE TKDE*, 2020.
- [Fan *et al.*, 2022a] Wenqi Fan, Xiaorui Liu, Wei Jin, Xiangyu Zhao, Jiliang Tang, and Qing Li. Graph trend filtering networks for recommendation. In *ACM SIGIR*, 2022.
- [Fan *et al.*, 2022b] Wenqi Fan, Xiangyu Zhao, Xiao Chen, Jingran Su, Jingtong Gao, Lin Wang, Qidong Liu, Yiqi Wang, Han Xu, Lei Chen, et al. A comprehensive survey on trustworthy recommender systems. *arXiv:2209.10117*, 2022.
- [Ferruz *et al.*, 2022] Noelia Ferruz, Steffen Schmidt, and Birte Höcker. Protgpt2 is a deep unsupervised language model for protein design. *Nature communications*, 2022.
- [Giessel *et al.*, 2022] Andrew Giessel, Athanasios Dousis, Kanchana Ravichandran, Kevin Smith, Sreyoshi Sur, Iain McFadyen, Wei Zheng, and Stuart Licht. Therapeutic enzyme engineering using a generative neural network. *Scientific Reports*, 2022.
- [Guan *et al.*, 2023] Jiaqi Guan, Wesley Wei Qian, Xingang Peng, Yufeng Su, Jian Peng, and Jianzhu Ma. 3d equivariant diffusion for target-aware molecule generation and affinity prediction. In *ICLR*, 2023.
- [Guo and Zhao, 2022] Xiaojie Guo and Liang Zhao. A systematic survey on deep generative models for graph generation. *IEEE TPAMI*, 2022.
- [Haefeli *et al.*, 2022] Kilian Konstantin Haefeli, Karolis Martinkus, Nathanaël Perraudin, and Roger Wattenhofer. Diffusion models for graphs benefit from discrete state spaces. In *LoG*, 2022.

- [Han *et al.*, 2022] Peng Han, Peilin Zhao, Chan Lu, Junzhou Huang, Jiayang Wu, Shuo Shang, Bin Yao, and Xiangliang Zhang. Gnn-retro: Retrosynthetic planning with graph neural networks. In *AAAI*, 2022.
- [Ho *et al.*, 2020] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *NeurIPS*, 2020.
- [Hoogeboom *et al.*, 2022] Emiel Hoogeboom, Victor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3d. In *ICML*, 2022.
- [Huang *et al.*, 2022a] Han Huang, Leilei Sun, Bowen Du, Yanjie Fu, and Weifeng Lv. Graphgdp: Generative diffusion processes for permutation invariant graph generation. In *IEEE ICDM*, pages 201–210, 2022.
- [Huang *et al.*, 2022b] Lei Huang, Hengtong Zhang, Tingyang Xu, and Ka-Chun Wong. Mdm: Molecular diffusion model for 3d molecule generation. *arXiv:2209.05710*, 2022.
- [Igashov *et al.*, 2022] Iliia Igashov, Hannes Stärk, Clément Vignac, Victor Garcia Satorras, Pascal Frossard, Max Welling, Michael M Bronstein, and Bruno Correia. Equivariant 3d-conditional diffusion models for molecular linker design. In *NeurIPS*, 2022.
- [Jing *et al.*, 2022] Bowen Jing, Gabriele Corso, Regina Barzilay, and Tommi S. Jaakkola. Torsional diffusion for molecular conformer generation. In *ICLR Workshop*, 2022.
- [Jo *et al.*, 2022] Jaehyeong Jo, Seul Lee, and Sung Ju Hwang. Score-based generative modeling of graphs via the system of stochastic differential equations. In *ICML*, 2022.
- [Jumper *et al.*, 2021] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Židek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 2021.
- [Kenton and Toutanova, 2019] Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *NAACL-HLT*, 2019.
- [Kong *et al.*, 2023] Ling kai Kong, Jiaming Cui, Haotian Sun, Yuchen Zhuang, B. Aditya Prakash, and Chao Zhang. Autoregressive diffusion model for graph generation. <https://openreview.net/forum?id=98J48HZXxd5>, 2023. Accessed: 2023-06-05.
- [Lee and Kim, 2022] Jin Sub Lee and Philip M Kim. Proteinsgm: Score-based generative modeling for de novo protein design. *bioRxiv*, 2022.
- [Lee *et al.*, 2022] Seul Lee, Jaehyeong Jo, and Sung Ju Hwang. Exploring chemical space with score-based out-of-distribution generation. *arXiv:2206.07632*, 2022.
- [Liao *et al.*, 2019] Renjie Liao, Yujia Li, Yang Song, Shenglong Wang, Will Hamilton, David K Duvenaud, Raquel Urtasun, and Richard Zemel. Efficient graph generation with graph recurrent attention networks. *NeurIPS*, 2019.
- [Lin *et al.*, 2022] Haitao Lin, Yufei Huang, Meng Liu, Xuanjing Li, Shuiwang Ji, and Stan Z Li. Diffbp: Generative diffusion of 3d molecules for target protein binding. *arXiv:2211.11214*, 2022.
- [Liu *et al.*, 2019] Jenny Liu, Aviral Kumar, Jimmy Ba, Jamie Kiros, and Kevin Swersky. Graph normalizing flows. *NeurIPS*, 2019.
- [Liu *et al.*, 2022] Haochen Liu, Yiqi Wang, Wenqi Fan, Xiaorui Liu, Yaxin Li, Shailei Jain, Yunhao Liu, Anil Jain, and Jiliang Tang. Trustworthy ai: A computational perspective. *ACM Trans Intell Syst Technol*, 2022.
- [Liu *et al.*, 2023] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM CSUR*, 2023.
- [Luo *et al.*, 2021a] Shitong Luo, Chence Shi, Minkai Xu, and Jian Tang. Predicting molecular conformation via dynamic graph score matching. In *NeurIPS*, 2021.
- [Luo *et al.*, 2021b] Youzhi Luo, Keqiang Yan, and Shuiwang Ji. Graphdf: A discrete flow model for molecular graph generation. In *ICML*, 2021.
- [Luo *et al.*, 2022a] Shitong Luo, Yufeng Su, Xingang Peng, Sheng Wang, Jian Peng, and Jianzhu Ma. Antigen-specific antibody design and optimization with diffusion-based generative models for protein structures. In *NeurIPS*, 2022.
- [Luo *et al.*, 2022b] Tianze Luo, Zhanfeng Mo, and Sinno Jialin Pan. Fast graph generative model via spectral diffusion. *arXiv:2211.08892*, 2022.
- [Ma and Zhang, 2021] Changsheng Ma and Xiangliang Zhang. Gf-vae: a flow-based variational autoencoder for molecule generation. In *CIKM*, 2021.
- [Nakata *et al.*, 2022] Shuya Nakata, Yoshiharu Mori, and Shigenori Tanaka. End-to-end protein-ligand complex structure generation with diffusion-based generative models. *bioRxiv*, 2022.
- [Niu *et al.*, 2020] Chenhao Niu, Yang Song, Jiaming Song, Shengjia Zhao, Aditya Grover, and Stefano Ermon. Permutation invariant graph generation via score-based generative modeling. In *AISTATS*, 2020.
- [O’Bray *et al.*, 2022] Leslie O’Bray, Max Horn, Bastian Rieck, and Karsten Borgwardt. Evaluation metrics for graph generative models: Problems, pitfalls, and practical solutions. In *ICLR*, 2022.
- [Peng *et al.*, 2023] Xingang Peng, Jiaqi Guan, Jian Peng, and Jianzhu Ma. Pocket-specific 3d molecule generation by fragment-based autoregressive diffusion models. <https://openreview.net/forum?id=HGsoe1wmRW5>, 2023. Accessed: 2023-06-05.
- [Qiao *et al.*, 2022] Zhuoran Qiao, Weili Nie, Arash Vahdat, Thomas F Miller III, and Anima Anandkumar. Dynamic-backbone protein-ligand structure prediction with multiscale generative diffusion models. *arXiv:2209.15171*, 2022.
- [Ramesh *et al.*, 2021] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. Zero-shot text-to-image generation. In *ICML*, 2021.

- [Réau *et al.*, 2023] Manon Réau, Nicolas Renaud, Li C Xue, and Alexandre MJJ Bonvin. Deeprank-gnn: a graph neural network framework to learn patterns in protein-protein interfaces. *Bioinformatics*, 2023.
- [Shi *et al.*, 2021] Chence Shi, Shitong Luo, Minkai Xu, and Jian Tang. Learning gradient fields for molecular conformation generation. In *ICML*, 2021.
- [Shin *et al.*, 2021] Jung-Eun Shin, Adam J Riesselman, Aaron W Kollasch, Conor McMahon, Elana Simon, Chris Sander, Aashish Manglik, Andrew C Kruse, and Debora S Marks. Protein design and variant prediction using autoregressive generative models. *Nature communications*, 2021.
- [Simonovsky and Komodakis, 2018] Martin Simonovsky and Nikos Komodakis. Graphvae: Towards generation of small graphs using variational autoencoders. In *ICANN*, 2018.
- [Song and Ermon, 2019] Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. *NeurIPS*, 2019.
- [Song *et al.*, 2021] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *ICLR*, 2021.
- [Trippe *et al.*, 2023] Brian L Trippe, Jason Yim, Doug Tischer, Tamara Broderick, David Baker, Regina Barzilay, and Tommi Jaakkola. Diffusion probabilistic modeling of protein backbones in 3d for the motif-scaffolding problem. In *ICLR*, 2023.
- [Vignac *et al.*, 2023] Clement Vignac, Igor Krawczuk, Antoine Siraudin, Bohan Wang, Volkan Cevher, and Pascal Frossard. Digress: Discrete denoising diffusion for graph generation. In *ICLR*, 2023.
- [Wang *et al.*, 2018] Hongwei Wang, Jia Wang, Jialin Wang, Miao Zhao, Weinan Zhang, Fuzheng Zhang, Xing Xie, and Minyi Guo. Graphgan: Graph representation learning with generative adversarial nets. In *AAAI*, 2018.
- [Watson *et al.*, 2022] Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eisenach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, et al. Broadly applicable and accurate protein design by integrating structure prediction networks and diffusion generative models. *bioRxiv*, 2022.
- [Wu *et al.*, 2022a] Kevin E Wu, Kevin K Yang, Rianne van den Berg, James Y Zou, Alex X Lu, and Ava P Amini. Protein structure generation via folding diffusion. *arXiv:2209.15611*, 2022.
- [Wu *et al.*, 2022b] Lemeng Wu, Chengyue Gong, Xingchao Liu, Mao Ye, and qiang liu. Diffusion-based molecule generation with informative prior bridges. In *NeurIPS*, 2022.
- [Xia *et al.*, 2021] Feng Xia, Ke Sun, Shuo Yu, Abdul Aziz, Liangtian Wan, Shirui Pan, and Huan Liu. Graph learning: A survey. *IEEE TAI*, 2021.
- [Xiao *et al.*, 2022] Chaowei Xiao, Zhongzhu Chen, Kun Jin, Jiong Xiao Wang, Weili Nie, Mingyan Liu, Anima Anandkumar, Bo Li, and Dawn Song. Densepure: Understanding diffusion models towards adversarial robustness. *arXiv:2211.00322*, 2022.
- [Xu *et al.*, 2019] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *ICLR*, 2019.
- [Xu *et al.*, 2022] Minkai Xu, Lantao Yu, Yang Song, Chence Shi, Stefano Ermon, and Jian Tang. Geodiff: A geometric diffusion model for molecular conformation generation. In *ICLR*, 2022.
- [Yang *et al.*, 2022] Ling Yang, Zhilong Zhang, Yang Song, Shenda Hong, Runsheng Xu, Yue Zhao, Yingxia Shao, Wentao Zhang, Bin Cui, and Ming-Hsuan Yang. Diffusion models: A comprehensive survey of methods and applications. *arXiv:2209.00796*, 2022.
- [Yang *et al.*, 2023] Ling Yang, Zhilong Zhang, Wentao Zhang, and Shenda Hong. Score-based graph generative modeling with self-guided latent diffusion. <https://openreview.net/forum?id=AykEgQNPJEK>, 2023. Accessed: 2023-06-05.
- [Zhang *et al.*, 2020] Ziwei Zhang, Peng Cui, and Wenwu Zhu. Deep learning on graphs: A survey. *IEEE Trans Knowl Data Eng*, 2020.
- [Zhang *et al.*, 2022] Zuobai Zhang, Minghao Xu, Arian Jamasb, Vijil Chenthamarakshan, Aurelie Lozano, Payel Das, and Jian Tang. Protein representation learning by geometric structure pretraining. *arXiv:2203.06125*, 2022.
- [Zhou *et al.*, 2019] Zhenpeng Zhou, Steven Kearnes, Li Li, Richard N Zare, and Patrick Riley. Optimization of molecules via deep reinforcement learning. *Scientific reports*, 2019.