Can Edge Addition Be Safe and Effective? Adjacency-Centered Augmentation via Langevin and SDE Diffusion for Self-Supervised Graph Anomaly Detection

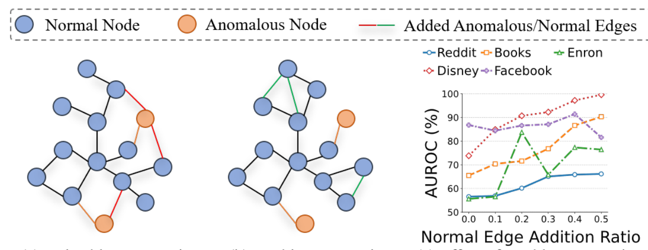
Haokai Gao[1], **Menghua Jiang**[1], **Xuantao Yang**[1], **Jiale Liu**[1], **Yubin Li**[1], **Yuncheng Jiang**[1,2,3] *
[1]School of Computer Science, South China Normal University, Guangzhou, China
[2]Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents of Ministry of Education (South China Normal University), Guangzhou, China
[3]School of Artificial Intelligence, South China Normal University, Foshan, China
{gaohaokai, jiangmenghua, 2024010274,liujiale,2024010263}@m.scnu.edu.cn, ycjiang@scnu.edu.cn

Abstract

Edge addition is commonly considered risky in Graph Anomaly Detection (GAD), as random edge addition may induce anomaly–normal connectivity. Consequently, most existing augmentation strategies focus on feature perturbation, edge removal, or subgraph sampling, leaving edge addition largely unexplored. In contrast, we empirically find that moderate and targeted structural completion among normal nodes consistently improves GAD performance, revealing guided edge addition as an overlooked yet effective augmentation dimension. Motivated by this observation, we introduce two adjacency-centered edge generation strategies with complementary mechanisms. One performs a training-free structural completion scheme via spectrum-aware Langevin dynamics, enriching graph connectivity while preserving node features. The other models the joint evolution of node features and graph structure through a stochastic differential equation–based diffusion process, producing structurally coherent and anomaly-aware complementary graphs. Extensive experiments on 12 benchmark datasets with 7 state-of-the-art GAD models demonstrate consistent and substantial improvements in both AUROC and AUPRC. Code and appendices are available at https://github.com/GaoHaokai222/LangGen-JointSDE

1 Introduction

Graph Anomaly Detection (GAD) aims to identify nodes in static graphs that deviate significantly from the majority in real-world information networks. Although such anomalies are typically rare, their detection is crucial due to the potential economic and social consequences. Representative examples include fraudulent users on online shopping platforms [Yu *et al.*, 2024; Zhang *et al.*, 2022], fake news on social media [Feng *et al.*, 2025; Sun *et al.*, 2025], spam on review-sharing platforms [Deng *et al.*, 2022], and intruders in communication networks [Westphal *et al.*, 2025]. With the increasing availability of graph-structured data and the rapid advancement of deep graph neural networks, GAD has attracted substantial research interest across multiple disciplines [Ma *et al.*, 2025; Ma *et al.*, 2021], and has become an important problem in both theoretical research and practical applications.



(a) Unhealthy Connections Involving Anomalies (b) Healthy Connections among Normal Nodes (c) Effect of Healthy Connection Ratio on BOURNE

Figure 1: A toy illustration of healthy vs. unhealthy connections, and the performance gains of BOURNE on five real-world anomaly datasets with varying normal–normal connection ratios.

Existing GAD methods are fundamentally challenged by the severe class imbalance between normal and anomalous nodes, which biases the learning process. Prior work addresses this issue through model design and data augmentation, with prevailing strategies relying on node feature masking, random edge dropping, and subgraph sampling, and recent extensions incorporating diffusion-based graph generators to enrich anomaly patterns. Despite these advances, edge addition or structural completion is rarely explored in GAD augmentation. The primary concern is that introducing new edges may contaminate anomaly-sensitive representations by creating spurious connections involving anomalous nodes, thereby degrading detection performance, as illustrated in Fig. 1(a).

However, as shown in Fig. 1(b) and (c), we empirically observe that carefully increasing edges among normal nodes consistently improves the performance of existing GAD models, with gains scaling approximately proportionally to the number of added edges. This suggests that guided edge addition preserving intra-community connectivity enhances feature aggregation among normal nodes while suppressing the influence of anomalous nodes. These results reveal an underexplored yet promising direction for graph augmentation and naturally raise the question of how to design a principled, label-free strategy for guided edge addition that reliably improves GAD.

To assess the feasibility of guided edge addition in real-world GAD scenarios, we revisit graph denoising and

*Corresponding author

diffusion-based generative models from a spectral energy perspective and examine their applicability to anomaly detection. Leveraging the relationship between graph signal frequencies and anomalous connectivity patterns, we propose two adjacency-matrix-centered diffusion-based augmentation methods, LangGen and JointSDE. LangGen is a feature-independent, training-free approach based on Langevin dynamics, which integrates the Boltzmann distribution with the Rayleigh quotient to increase edge density while preserving low-frequency spectral energy, thereby strengthening intra-community connectivity without model training. JointSDE, in contrast, formulates a joint feature–structure diffusion process via stochastic differential equations, where spectral graph principles guide the synchronized evolution of node features and the adjacency matrix. This joint diffusion gradually densifies graph structure while regularizing feature distributions, enabling the reconstruction of edge-augmented graphs whose structural and feature representations are jointly evolved and mutually consistent. Extensive experiments on multiple benchmarks with state-of-the-art GAD models demonstrate that both methods generate high-quality edge-augmented graphs with efficient computational and memory overhead, leading to consistent improvements in AUROC and AUPRC.

The contributions of this work are summarized as follows:

- To the best of our knowledge, this work is the first to explore edge addition as a graph augmentation strategy for graph anomaly detection, showing that carefully increasing edges among normal nodes can improve the performance of diverse GAD models.
- Two adjacency-matrix-centered, spectral-guided diffusion methods, LangGen and JointSDE, are proposed to generate high-quality edge-augmented graphs without label information.
- Extensive evaluations on seven state-of-the-art GAD models across twelve benchmark datasets validate the effectiveness, efficiency, and broad applicability of the proposed augmentation strategies.

2 Related Work

2.1 Graph Anomaly Detection

Graph anomaly detection aims to identify rare patterns that deviate from dominant graph structures or attributes [Ma *et al.*, 2021; Pazho *et al.*, 2023; Wang *et al.*, 2025; Jin *et al.*, 2024]. Early methods such as Radar [Li *et al.*, 2017] and ANOMALOUS [Peng *et al.*, 2018] detect anomalies via matrix factorization and residual analysis. Deep learning-based approaches, including DOMINANT [Ding *et al.*, 2019], AnomalyDAE [Fan *et al.*, 2020], and ComGA [Luo *et al.*, 2022], detect anomalies via reconstruction errors in graph autoencoders, under the assumption that anomalous nodes are harder to faithfully reconstruct from learned normal patterns. Contrastive learning methods, such as CoLA [Liu *et al.*, 2021], SL-GAD [Zheng *et al.*, 2021], HUGE [Pan *et al.*, 2025], and SAMCL [Hu *et al.*, 2025], enhance detection by contrasting node- and subgraph-level representations. Recently, inspired by the observation that anomalies manifest distinctive spectral energy patterns [Tang *et al.*, 2022], spectral-based GAD has emerged as an active research direction [Gao *et al.*, 2023; Dong *et al.*, 2024; Xu *et al.*, 2024; Zheng *et al.*, 2025]. Representative methods such as UniGAD [Lin *et al.*, 2024] leverage MRQ-based spectral sampling for multi-level detection, while SALAD [Ma *et al.*, 2025] and ADA-GAD [He *et al.*, 2024] mitigate anomaly interference in normal structure learning through loss reweighting or Rayleigh energy–based denoising. Despite these advances, existing GAD methods predominantly rely on limited and fixed augmentation strategies, such as node masking, feature perturbation, edge dropping, and subgraph sampling, while the potential of edge addition or structural completion remains unexplored.

2.2 Diffusion Model

Diffusion models [Ho *et al.*, 2020; Song *et al.*, 2021a; Song *et al.*, 2021b] have recently emerged as powerful generative frameworks, achieving strong performance across image synthesis [Huang *et al.*, 2025], time-series modeling [Yang *et al.*, 2024], and other domains [Cao *et al.*, 2024]. This success has motivated their extension to graph-structured data. DiGress [Vignac *et al.*, 2023] introduces a discrete diffusion process that perturbs graph structures by stochastically modifying edges and node attributes, and learns an inverse transformation to recover clean graphs via node- and edge-level prediction. GDSS [Jo *et al.*, 2022] formulates graph generation as a continuous-time diffusion process based on stochastic differential equations, demonstrating the effectiveness of whole-graph diffusion for molecular generation. Diffusion models have also been explored in graph anomaly detection. [Ma *et al.*, 2024] generates node features via probabilistic denoising diffusion while keeping the adjacency matrix fixed, aiming to simulate anomalous perturbations. AGDiff [Cai *et al.*, 2025] addresses graph-level anomaly detection by modeling the generation of abnormal graphs, and DiffGAD [Li *et al.*, 2025] further integrates diffusion into the anomaly scoring mechanism. Despite progress, diffusion-based GAD methods remain conservative in structural modeling: some diffuse only node features without modifying connectivity, while others rely on feature-similarity–based adjacency reconstruction, which is fragile under distribution shifts. Moreover, general graph diffusion models neither explicitly model anomalies nor examine the co-evolution of edge density and feature diffusion, leaving principled edge addition in GAD largely unexplored.

3 Methodology

3.1 Preliminary and Problem Formulation

Given an attributed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{X})$, where $\mathcal{V} = \{v_1, \ldots, v_N\}$ denotes the node set, $\mathcal{E}$ the edge set, and $\mathbf{X} \in \mathbb{R}^{N \times d}$ the node feature matrix. Let $\mathbf{A} \in \mathbb{R}^{N \times N}$ denote the adjacency matrix such that $A_{ij} = 1$ if $(v_i, v_j) \in \mathcal{E}$ and $A_{ij} = 0$ otherwise. The goal of self-supervised graph anomaly detection is to learn a function $\mathcal{F}_\theta : \mathcal{G} \rightarrow \mathbb{R}^N$ that assigns each node an anomaly score quantifying its deviation from normal patterns.

3.2 Training-free Generation Approach: LangGen

Langevin Dynamics. Langevin dynamics [Song and Ermon, 2019] enables sampling from high-probability regions of a

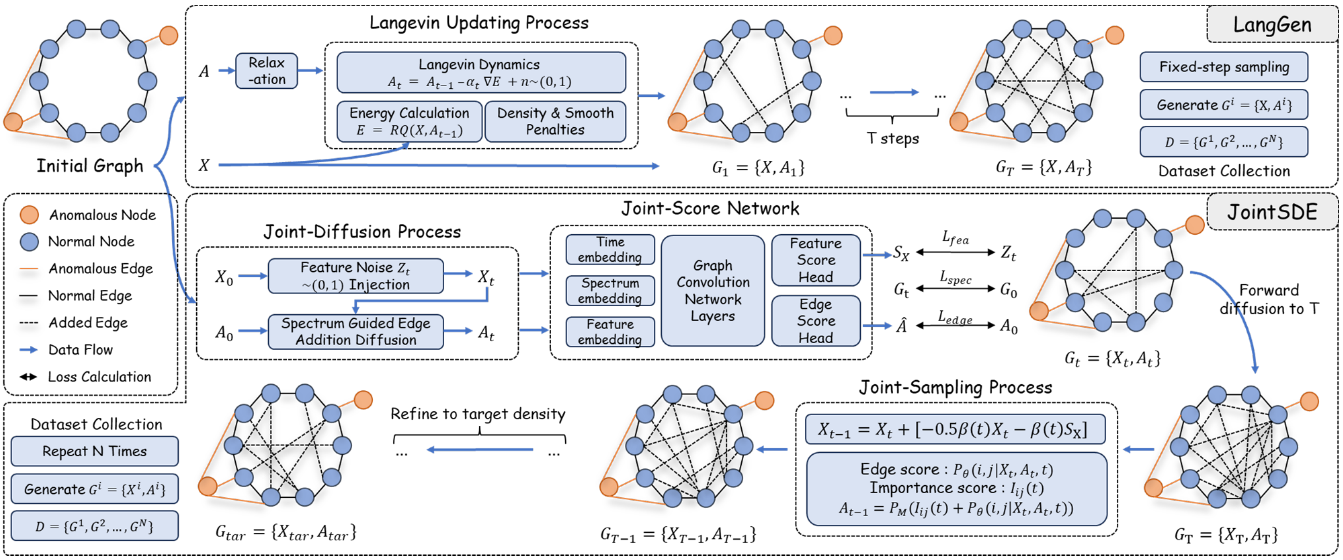


Figure 2: Illustration of the proposed LangGen and JointSDE frameworks. LangGen performs training-free, spectrum-guided structural completion by adding edges to the input graph based on an energy formulation, while keeping node features fixed. JointSDE jointly models the coupled evolution of node features and graph structure through a diffusion–reverse process, where spectral graph principles guide the synchronized generation of both the adjacency matrix and node representations.

target distribution $p_\theta(\mathbf{x})$ by iteratively combining gradient-driven updates with Gaussian noise:

$$\tilde{\mathbf{x}}_{t+1} = \tilde{\mathbf{x}}_t + \epsilon \nabla_{\mathbf{x}} \log p_\theta(\mathbf{x}) + \sqrt{2\epsilon}\,\mathbf{z}_t, \quad \mathbf{z}_t \sim \mathcal{N}(0, \mathbf{I}), \quad (1)$$

where ϵ denotes the step size. The gradient term drives samples toward regions of high probability, while the injected noise promotes exploration and prevents premature convergence to poor local minima. In the continuous-time limit, this process asymptotically samples from the target distribution.

Boltzmann Distribution and Energy-Based Formulation. The Boltzmann distribution establishes a principled connection between energy functions and probability densities:

$$p_\theta(\mathbf{x}) \propto \exp\big(-E_\theta(\mathbf{x})\big), \quad (2)$$

where $E_\theta(\mathbf{x})$ is a differentiable energy function. Under this formulation, the gradient of the log-density simplifies to

$$\nabla_{\mathbf{x}} \log p_\theta(\mathbf{x}) = -\nabla_{\mathbf{x}} E_\theta(\mathbf{x}), \quad (3)$$

enabling efficient sampling by directly optimizing the energy landscape without explicitly evaluating or normalizing the probability density $p_\theta(\mathbf{x})$.

Spectral Motivation for Graph Anomaly Detection. Recent studies [Tang *et al.*, 2022] have shown that anomalies exhibit elevated high-frequency energy in the graph spectral domain, and such accumulated energy can be captured by the Rayleigh Quotient (RQ):

$$RQ(\mathbf{X}, \mathbf{A}) = \frac{\mathrm{tr}(\mathbf{X}^\top L \mathbf{X})}{\mathrm{tr}(\mathbf{X}^\top \mathbf{X})} = \frac{\sum_{i,j} A_{ij} \|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sum_i \|\mathbf{x}_i\|^2}, \quad (4)$$

where, $L = D - A$ denotes the combinatorial graph Laplacian, D is the degree matrix. Minimizing RQ promotes spectrally smooth structures while suppressing anomalous high-frequency patterns. By incorporating RQ into energy-based Langevin dynamics, the sampling process is explicitly guided toward spectrally coherent graph configurations. Notably, when the node feature matrix $\mathbf{X}$ is fixed during generation, energy optimization with respect to the adjacency matrix A directly governs structural evolution, marking a shift from conventional feature-generation paradigms to explicit edge generation. Under this setting, minimizing RQ favors the addition of structurally consistent edges among normal nodes rather than indiscriminate edge removal.

Energy Function Design. To enable gradient-based optimization over graph structure, we treat the adjacency matrix A as a diffusion variable and apply a continuous relaxation. Given an initial binary adjacency matrix $\mathbf{A}_0$, we first clamp entries to $[\epsilon, 1-\epsilon]$ to avoid logit singularities ($\epsilon = 10^{-6}$), then compute

$$\tilde{\mathbf{A}}_0 = \mathrm{sigm}\left(\frac{1}{\tau_t} \log \frac{\mathbf{A}_0}{1 - \mathbf{A}_0}\right), \quad (5)$$

where $\mathrm{sigm}(\cdot)$ is the sigmoid function and τ_t controls the relaxation temperature, allowing differentiable updates without introducing learnable parameters.

The total energy function is defined as

$$E_\theta(\mathbf{X}, \mathbf{A}) = RQ(\mathbf{X}, \mathbf{A}) + \lambda_{dens} \left\|\mathrm{Dens}(\mathbf{A}) - \rho_{\mathrm{tar}}\right\|^2 + \lambda_{smo} \|\mathbf{A} - \mathbf{A}_0\|^2, \quad (6)$$

where $\mathrm{Dens}(\mathbf{A}) = 2|\mathcal{E}|/N^2$ denotes graph density, ρ_{tar} is the target density, and $\lambda.$ are trade-off coefficients. The density and smoothness terms prevent degenerate solutions and stabilize structural evolution.

Langevin Update and Annealing. The Langevin update for adjacency matrix generation is given by

$$\tilde{\mathbf{A}}_{t+1} = \tilde{\mathbf{A}}_t - \epsilon \nabla_A E_\theta(\mathbf{X}, \mathbf{A}_t) + \epsilon \sigma_t \mathbf{z}_t, \quad \mathbf{z}_t \sim \mathcal{N}(0, \mathbf{I}), \quad (7)$$

where ϵ and σ_t follow annealing schedules to allow coarse exploration in early iterations and fine-grained refinement later. This process produces spectrally coherent, edge-augmented graphs without requiring any model training.

3.3 Feature-Structure Joint Diffustion: JointSDE

Joint Noising Process. The joint noising process characterizes the coupled degradation of node features and graph structure under a unified diffusion time scale, allowing structural uncertainty to adapt to the evolving feature geometry. As the diffusion time t increases, node features progressively lose discriminative information, while structural constraints are gradually relaxed, driving the system toward a highly stochastic state and providing diverse initial conditions for subsequent reverse refinement.

For continuous node features, we adopt a variance-preserving diffusion formulation [Song *et al.*, 2021b] with the closed-form forward process

$$\mathbf{X}_t = \alpha(t)\mathbf{X}_0 + \sigma(t)\mathbf{Z}, \qquad \mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \tag{8}$$

where $\alpha(t) \in (0,1]$ controls signal preservation and $\sigma(t) = \sqrt{1-\alpha^2(t)}$ denotes the noise scale. As t increases, $\mathbf{X}_t$ converges to an isotropic Gaussian.

For graph structure, we adopt a diffusion-inspired probabilistic relaxation that maps the discrete adjacency into time-dependent edge probabilities $P_{ij}(t) \in (0,1)$:

$$P_{ij}(t) = \mathrm{sigm}\left(\frac{1}{\tau(t)} \log \frac{P_{ij}^{(0)}}{1 - P_{ij}^{(0)}} + S_{ij}(t)\right), \tag{9}$$

where $P_{ij}^{(0)}$ is obtained from the clamped $A_{0,ij}$ via logit, $\tau(t)$ controls temperature, and $\mathrm{sigm}(\cdot)$ is the sigmoid function. The final adjacency used for backbone training is binarized by thresholding at 0.5, with optional symmetry enforcement and self-loop handling.

The guidance term $S_{ij}(t)$ jointly encodes feature coherence and topological accessibility:

$$S_{ij}(t) = \frac{1}{2}\,\mathrm{Sim}_{ij}(t) + \frac{1}{2}\,\frac{R_{\mathrm{eff}}(i,j;\mathbf{A}_0)}{\max_{p,q} R_{\mathrm{eff}}(p,q;\mathbf{A}_0)}, \tag{10}$$

where $\mathrm{Sim}_{ij}(t) = \frac{1}{2}\big(\cos(\mathbf{x}_i(t), \mathbf{x}_j(t)) + 1\big)$ measures feature similarity under diffusion, and $R_{\mathrm{eff}}(i,j;\mathbf{A}_0) = L^+_{ii} + L^+_{jj} - 2L^+_{ij}$ denotes the effective resistance distance derived from the pseudoinverse of the original Laplacian, capturing global multi-path connectivity. This formulation enables a tightly coupled noising process in which structural uncertainty is continuously modulated by the evolving feature state.

Joint Score Network. The joint score network models feature scores and structural uncertainty on the noisy graph $\mathcal{G}_t = (\mathbf{X}_t, \mathbf{A}_t)$ under a shared diffusion time t, enabling coordinated reverse-time refinement of node attributes and graph topology. A unified encoder f_θ processes $(\mathbf{X}_t, \mathbf{A}_t, t)$ with explicit time conditioning, where the time embedding is defined as

$$\phi(t) = W_{t,2}\,\mathrm{SiLU}(W_{t,1}t + b_{t,1}) + b_{t,2}. \tag{11}$$

To capture global structural perturbations induced by structural noise, we construct the symmetric normalized Laplacian

$$L_t = I - D_t^{-1/2}\mathbf{A}_t D_t^{-1/2}, \tag{12}$$

and perform spectral decomposition $L_t = \Phi_t \Lambda_t \Phi_t^\top$. The first k low-frequency eigenvectors $\Phi_{t,k}$ provide spectral coordinates for each node. The initial node representation is given by

$$\mathbf{h}_i^{(0)} = \Big[\mathbf{X}_{t,i}W_e + b_e \;\|\; E_{\mathrm{spec}}(\Phi_{t,k}[i,:])\Big], \tag{13}$$

and updated through $\mathcal{L}$ layers of time-conditioned message passing:

$$\mathbf{h}^{(\ell)} = \mathrm{SiLU}\Big(\mathrm{LN}\Big(\tilde{D}_t^{-1/2}\tilde{A}_t\tilde{D}_t^{-1/2}\mathbf{h}^{(\ell-1)}W^{(\ell)} + \phi(t)\Big)\Big), \tag{14}$$

where $\tilde{A}_t = \mathbf{A}_t + I$ and $\tilde{D}_t$ is the corresponding degree matrix. The final node representations are denoted by $\mathbf{H} = \mathbf{h}^{(\mathcal{L})}$.

The feature score head predicts the denoising score $\mathbf{S}_X = g_{\mathrm{fea}}(\mathbf{H})$ and is trained via denoising score matching:

$$\mathcal{L}_{\mathrm{fea}} = \mathbb{E}_{\mathbf{X}_0,\mathbf{Z},t}\big[\|\mathbf{S}_X + \tfrac{\mathbf{Z}}{\sigma(t)}\|^2\big], \tag{15}$$

where $\mathbf{Z}$ denotes standard Gaussian noise.

For structural denoising, we estimate edge existence likelihoods conditioned on the noisy graph as a tractable surrogate for reverse-time refinement. An edge-level representation is constructed as $\mathbf{u}_{ij} = \big[\mathbf{h}_i^{(\mathcal{L})} \,\|\, \mathbf{h}_j^{(\mathcal{L})} \,\|\, \phi(t)\big]$, and the edge head outputs $\hat{A}_{ij} = g_{\mathrm{edge}}(\mathbf{u}_{ij})$, with the objective

$$\mathcal{L}_{\mathrm{edge}} = -\sum_{i,j}\big(A_{0,ij}\log\hat{A}_{ij} + (1 - A_{0,ij})\log(1-\hat{A}_{ij})\big). \tag{16}$$

To preserve global geometry during reverse refinement, we introduce a low-frequency spectral regularization

$$\mathcal{L}_{\mathrm{spec}} = \sum_{k=1}^{K}(\lambda_k^{\mathrm{gen}} - \lambda_k^{\mathrm{orig}})^2, \tag{17}$$

where λ_k^{gen} and $\lambda_k^{\mathrm{orig}}$ denote the k-th eigenvalues of the symmetric normalized Laplacian of the generated and original graphs, respectively. The joint score network is trained by minimizing

$$\mathcal{L}_{\mathrm{joint}} = \mathcal{L}_{\mathrm{fea}} + \mathcal{L}_{\mathrm{edge}} + \mathcal{L}_{\mathrm{spec}}, \tag{18}$$

enforcing consistency between feature-level score estimation and structure-level refinement under a shared latent representation.

Joint Sampling Process. Unlike conventional diffusion models that initialize sampling from pure Gaussian noise, JointSDE injects noise at a randomly sampled diffusion time t and performs partial reverse-time evolution. During sampling, node features are refined via a continuous-time reverse diffusion process, while graph structure evolves through a deterministic, feature-conditioned edge selection mechanism.

The reverse-time update of node features follows the VP–SDE and is discretized using an Euler scheme:

$$\mathbf{X}_{t-\Delta t} = \mathbf{X}_t + \big(-\tfrac{1}{2}\beta(t)\mathbf{X}_t - \beta(t)\mathbf{S}_X(\mathbf{X}_t, \mathbf{A}_t, t)\big)\,\Delta t, \tag{19}$$

where $\mathbf{S}_X$ denotes the learned feature score function.

Structural refinement is performed deterministically at each step by retaining or pruning edges according to a feature-conditioned priority score, avoiding the need to explicitly define a stochastic reverse process over discrete adjacency variables. For each edge (i, j), the structural importance score is defined as

$$I_{ij}(t) = \lambda_{\rm spec}\, \tilde{S}_{ij}^{(\rm spec)} + \lambda_{\rm sim}\, \tilde{S}_{ij}^{(x)} + \lambda_{\rm deg}\, \tilde{D}_{ij}, \qquad (20)$$

where $\tilde{S}_{ij}^{(\rm spec)} = \sum_{k=1}^{K'} \omega_k \big(\phi_k(i) - \phi_k(j)\big)^2$, $\tilde{S}_{ij}^{(x)} = \mathrm{sigm}\Big(\frac{\mathbf{x}_i^\top \mathbf{x}_j}{\|\mathbf{x}_i\|\,\|\mathbf{x}_j\|}\Big)$. Here, $\tilde{S}_{ij}^{(\rm spec)}$ captures the contribution of an edge to low-frequency Laplacian components, with frequency-decaying weights ω_k emphasizing global geometric stability, while $\tilde{S}_{ij}^{(x)}$ measures local feature homophily. The degree-based regularizer $\tilde{D}_{ij} = \frac{1}{d_i + d_j + 1}$ suppresses hub-dominated structural bias.

Learned structural preferences are incorporated via the model-predicted edge retention probability

$$P_\theta(i, j \mid \mathbf{X}_t, \mathbf{A}_t, t) = \mathrm{sigm}(e_\theta(i, j \mid \mathbf{X}_t, \mathbf{A}_t, t)), \qquad (21)$$

and the final edge priority is defined as

$$J_{ij}(t) = I_{ij}(t) + \lambda_{\rm score}\, P_\theta(i, j \mid \mathbf{X}_t, \mathbf{A}_t, t). \qquad (22)$$

At each reverse step, the top M_t edges with the highest priority scores are retained:

$$\mathbf{A}_{t-\Delta t} = \mathcal{P}_{M_t}(J(t)), \qquad (23)$$

where $\mathcal{P}_{M_t}(\cdot)$ selects the M_t highest-scoring edges and M_t decreases monotonically as $t \to 0$. Overall, the joint sampling process is summarized as

$$(\mathbf{X}_{t-\Delta t}, \mathbf{A}_{t-\Delta t}) = \Psi\big((\mathbf{X}_t, \mathbf{A}_t), t\big), \qquad (24)$$

where Ψ denotes the coupled reverse-time operator combining stochastic feature refinement with spectrally guided deterministic structural evolution.

3.4 GAD Training with Synthetic Data

To enhance a given anomaly detector $\mathcal{F}_\theta$, we augment the original graph $\mathcal{G}$ with a set of generated graphs $\mathcal{D} = \{\mathcal{G}^1, \ldots, \mathcal{G}^{|\mathcal{D}|}\}$. For LangGen, $\mathcal{G}^i = \{X, A^i\}$ uses original node features with generated edges, whereas for JointSDE, $\mathcal{G}^i = \{X^i, A^i\}$ generates both node features and edges. The detector is further optimized on both the original and generated graphs by minimizing

$$\mathcal{L}_{\rm GAD} = \mathcal{L}(\mathcal{F}_\theta(\mathcal{G}), \mathcal{G}) + \frac{1}{|\mathcal{D}|} \sum_{i=1}^{|\mathcal{D}|} \mathcal{L}\big(\mathcal{F}_\theta(\mathcal{G}^i), \mathcal{G}^i\big).$$

4 Experiments

In this section, we conduct extensive experiments to answer the following research questions: **Q1:** Can the proposed methods consistently improve state-of-the-art GAD models? **Q2:** How do individual components in the energy formulation and diffusion framework contribute to structure generation during sampling? **Q3:** How does the proposed approach compare with existing edge-based graph augmentation strategies? **Q4:** What is the computational efficiency of the proposed graph generation process? **Q5:** Does the proposed method preferentially promote edge formation among normal nodes?

Dataset	#Node	#Edge	Dim.	Avg.Deg.	#Ano.	Ratio (%)
Cora	2,708	11,060	1,433	4.1	70	2.5
Amazon	13,752	515,042	767	37.2	350	2.5
BlogCatalog	5,196	171,743	8,189	66.11	300	5.77
ACM	16,484	71,980	8,337	8.73	597	3.62
Citeseer	3,327	4,732	3,703	2.77	150	4.50
PubMed	19,717	44,338	500	4.50	600	3.04
Weibo	8,405	407,963	400	48.5	868	10.3
Reddit	10,984	168,016	64	15.3	366	3.3
Books	1,418	3,695	21	2.6	28	2.0
Enron	13,533	176,987	18	13.1	5	0.04
Disney	124	335	28	2.7	6	4.8
Facebook	1,081	55,104	576	50.97	25	2.31

Table 1: Statistics of datasets used in our experiments, where # denotes quantity, Dim. the feature dimension, #Ano. the number of anomalies, Avg. Deg. the average node degree, and Ratio the proportion of anomalous nodes.

4.1 Experimental Setup

Datasets. We evaluate our method on twelve widely used graph anomaly detection benchmarks, including six datasets with *organic anomalies* (Weibo, Reddit, Books, Enron, Disney, Facebook) and six with *injected anomalies* (Cora, Amazon, BlogCatalog, ACM, Citeseer, PubMed), with dataset statistics reported in Table 1 and originally collected from [Liu *et al.*, 2024b; Liu *et al.*, 2024c].

Baselines and Metrics. We compare against seven representative unsupervised or self-supervised GAD methods: DOMINANT[Ding *et al.*, 2019], DONE[Bandyopadhyay *et al.*, 2020], AnomalyDAE[Fan *et al.*, 2020], CoLA[Liu *et al.*, 2021], GADNR[Roy *et al.*, 2024], BOURNE[Liu *et al.*, 2024a], and CGADM[Wei *et al.*, 2025]. Performance is evaluated using AUROC and AUPRC.

Implementation Details. LangGen performs training-free edge generation via energy-guided Langevin sampling with 1000 iterations. JointSDE adopts a diffusion-based framework with 1000 forward noising steps and 500 reverse denoising steps. For each graph, ten augmented graphs are generated. The number of edges in each generated graph is fixed to 150% of the original. Baselines are implemented using PyGOD[Liu *et al.*, 2024b] or official code with default settings. Results are averaged over five runs. All experiments are conducted on a single NVIDIA RTX 4090 GPU.

4.2 Improvement on Anomaly Detection Performance (Q1)

Tables 3 and 2 summarize the performance changes of different detectors after incorporating additional generated graphs during training. Across six organic and six injected anomaly datasets, all evaluated baselines consistently achieve improvements in both AUROC and AUPRC, demonstrating that the proposed edge-augmentation-based strategy effectively enhances the discriminative capability of existing graph anomaly detection models. These gains are observed across diverse graph types, including citation, social, review, transaction, and email networks, indicating that the method generalizes well despite substantial differences in degree distributions, community structures, and edge semantics.

Across most experimental settings, the absolute improvements in AUROC are larger than those in AUPRC. This trend

Methods	Weibo		Reddit		Books		Enron		Disney		Facebook	
	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC
DOMINANT	55.37±18.00	11.55±5.68	56.12±0.06	3.73±0.02	39.39±2.66	1.56±0.03	51.92±0.66	0.08±0.00	51.55±3.04	5.56±0.57	11.04±2.03	1.31±0.02
+ *LanGen*	89.58±0.10	23.37±0.35	57.24±0.11	3.74±0.02	41.50±1.69	1.61±0.02	53.72±0.73	0.08±0.00	53.01±1.24	5.82±0.18	14.73±0.09	1.36±0.00
+ *JointSDE*	89.51±0.26	23.41±0.58	57.20±0.04	3.74±0.00	42.79±1.27	1.65±0.04	55.43±0.11	0.08±0.00	66.61±1.76	8.78±0.84	14.22±0.18	1.36±0.00
DONE	57.96±0.97	5.63±0.23	51.40±2.26	4.09±0.12	37.91±1.64	1.51±0.05	41.10±3.15	0.14±0.03	42.80±4.23	4.80±0.04	17.14±1.94	1.39±0.04
+ *LanGen*	84.59±0.57	30.40±0.24	57.51±1.85	4.14±0.07	40.97±3.15	1.73±0.16	42.72±0.93	0.16±0.08	43.22±4.50	4.89±0.17	23.65±2.33	3.38±2.65
+ *JointSDE*	84.61±0.16	30.45±0.21	57.18±0.55	4.09±0.08	42.01±2.49	1.82±0.25	41.68±0.60	0.11±0.03	43.56±0.26	4.88±0.03	23.45±0.66	2.11±0.46
AnomalyDAE	76.55±1.31	14.80±0.43	50.06±2.03	3.22±0.22	48.87±4.82	2.45±1.32	45.85±13.16	0.09±0.02	66.67±4.17	11.02±0.67	46.31±2.03	2.48±0.02
+ *LanGen*	78.99±0.95	15.36±0.85	55.69±3.47	3.91±0.41	61.86±2.01	3.08±0.18	70.50±2.22	0.09±0.00	72.36±0.24	12.55±0.23	51.69±0.52	2.57±0.06
+ *JointSDE*	85.30±0.92	19.52±1.47	56.41±2.48	4.03±0.26	59.22±6.50	2.88±0.87	69.38±2.62	0.09±0.01	71.16±2.22	11.95±0.39	53.00±0.50	2.63±0.02
CoLA	27.54±1.22	3.65±0.57	50.01±1.10	3.82±0.28	49.69±4.20	2.02±0.18	50.30±2.33	0.05±0.02	42.45±4.25	5.64±0.38	71.98±1.82	7.34±1.03
+ *LanGen*	30.60±1.00	5.78±1.18	52.01±0.64	3.99±0.27	52.95±4.43	4.17±1.27	62.19±10.27	0.12±0.04	54.46±3.72	6.72±0.55	78.66±1.66	6.24±0.56
+ *JointSDE*	30.20±2.25	4.28±0.46	54.04±2.65	4.07±0.33	57.68±0.98	2.48±0.11	63.09±2.54	0.06±0.01	59.49±2.74	10.07±2.36	81.36±0.71	8.35±1.68
GADNR	38.36±4.19	3.15±0.18	47.24±7.60	3.31±0.58	44.15±0.61	1.75±0.02	45.18±3.66	0.07±0.00	27.97±0.50	3.72±0.03	14.22±3.67	1.38±0.05
+ *LanGen*	47.42±1.24	3.82±0.18	55.96±0.36	3.74±0.01	50.59±1.96	1.96±0.04	53.46±0.46	0.08±0.00	50.19±2.82	7.47±1.26	27.40±17.53	1.65±0.45
+ *JointSDE*	62.55±1.70	5.54±0.43	55.94±0.32	3.75±0.01	45.61±0.45	1.75±0.03	72.24±1.56	0.11±0.00	49.77±6.61	5.49±0.66	47.77±8.28	2.17±0.42
BOURNE	89.00±1.08	8.37±1.02	56.54±0.81	3.71±0.08	65.49±4.08	2.49±0.00	55.59±5.40	0.05±0.01	73.81±5.84	8.20±2.01	86.77±1.20	3.57±0.26
+ *LanGen*	91.26±0.10	10.11±0.72	64.81±0.51	4.34±0.05	70.71±0.72	2.50±0.07	84.44±4.43	0.06±0.02	77.82±0.99	9.88±1.73	92.29±0.98	5.73±0.17
+ *JointSDE*	90.09±1.25	8.25±0.98	65.77±0.79	3.77±0.20	69.09±1.82	2.49±0.00	84.30±1.18	0.06±0.01	82.46±1.34	9.44±1.12	91.11±0.99	5.86±1.26
CGADM	92.54±3.02	33.13±2.04	65.68±1.01	7.28±0.64	51.40±10.82	8.59±3.28	59.17±8.66	0.08±0.10	46.49±7.34	6.00±1.35	96.06±0.28	62.19±0.36
+ *LanGen*	95.24±0.11	34.62±0.57	66.50±0.21	7.72±0.28	59.53±5.77	11.95±5.65	60.11±5.91	0.09±0.04	62.80±4.46	15.85±10.76	96.33±0.03	62.40±0.26
+ *JointSDE*	95.12±0.07	33.27±0.54	66.32±0.18	7.32±0.40	59.04±4.42	12.73±7.17	65.11±3.51	0.10±0.02	62.92±4.17	12.92±8.90	96.35±0.06	62.21±0.27

Table 2: Node anomaly detection performance on six real-world benchmark datasets with organic anomalies. Results are reported as mean ± standard deviation (in %). Green and red values indicate improvements and drops over the corresponding backbone model, respectively.

Methods	Cora		Amazon		BlogCatalog		ACM		Citeseer		PubMed	
	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC
DOMINANT	79.10±4.37	32.75±1.98	50.33±0.01	2.50±0.00	60.81±4.23	8.80±1.34	74.39±2.96	10.61±0.39	82.89±7.17	35.44±0.88	80.01±0.57	9.88±0.13
+ *LanGen*	87.36±0.15	36.35±0.10	50.35±0.01	2.51±0.00	68.71±2.03	11.27±0.82	79.99±0.00	11.38±0.01	91.12±2.15	36.36±0.12	81.03±0.66	10.08±0.11
+ *JointSDE*	87.84±0.03	36.33±0.13	50.42±0.01	2.51±0.00	68.68±0.71	11.22±0.32	80.02±0.72	11.35±0.00	91.42±0.02	36.40±0.04	80.80±0.11	9.96±0.10
DONE	70.10±2.40	23.71±0.42	60.84±1.29	5.36±0.26	44.94±0.64	5.01±0.08	64.59±0.34	7.42±0.30	64.83±0.84	20.69±1.41	87.70±0.22	20.97±0.28
+ *LanGen*	84.05±1.12	53.77±0.36	75.40±2.33	6.01±0.42	76.46±0.14	32.59±0.20	93.59±0.11	29.71±0.81	90.53±0.59	48.12±1.47	91.63±0.20	23.26±0.30
+ *JointSDE*	88.91±0.24	56.54±2.12	68.14±1.31	6.72±1.20	76.28±0.06	31.98±0.53	93.62±0.12	29.92±0.44	90.85±0.17	45.89±3.30	91.57±0.37	22.18±1.53
AnomalyDAE	51.05±12.71	13.20±5.94	77.39±0.01	5.48±0.06	48.87±13.78	18.89±9.56	52.47±12.00	13.69±7.09	60.57±12.46	17.83±7.66	73.87±0.27	23.83±0.88
+ *LanGen*	76.95±0.23	25.88±0.89	98.49±0.06	54.4±0.00	74.75±0.02	36.87±0.10	74.86±3.43	26.22±1.66	76.12±0.06	27.80±0.24	75.06±0.33	25.73±0.13
+ *JointSDE*	75.74±0.19	24.51±1.00	98.51±0.06	54.42±0.00	74.96±0.19	36.97±0.09	74.80±0.21	26.30±0.43	75.91±0.10	26.40±1.01	75.33±0.50	24.98±0.89
CoLA	55.39±1.53	6.38±0.42	61.00±1.09	9.60±0.35	53.28±1.19	8.30±0.28	50.16±0.85	4.62±0.38	54.05±2.06	5.41±0.51	70.08±0.24	13.99±0.43
+ *LanGen*	58.94±1.59	7.32±0.43	63.6±0.81	9.81±0.23	53.78±1.60	8.86±1.06	53.79±0.13	8.87±1.27	55.14±1.03	5.65±0.29	70.87±0.38	14.29±1.48
+ *JointSDE*	67.42±1.13	12.42±0.97	75.98±1.46	20.12±1.48	55.54±0.86	13.53±2.43	51.84±0.97	4.79±0.60	59.73±4.13	7.78±1.46	75.17±4.44	14.97±3.00
GADNR	72.57±1.32	26.96±2.52	49.57±0.46	2.49±0.03	53.98±0.76	6.81±0.09	71.56±0.20	10.88±0.44	74.07±0.63	32.54±1.24	73.95±0.18	9.40±0.09
+ *LanGen*	75.34±0.19	31.40±0.20	49.58±0.69	2.49±0.06	55.01±0.22	6.82±0.06	71.93±0.32	11.10±0.04	75.71±0.67	34.76±1.15	74.84±0.76	10.73±1.90
+ *JointSDE*	75.37±0.29	30.68±0.78	51.50±0.85	2.64±0.15	54.86±0.42	6.87±0.00	71.47±0.12	10.12±0.01	76.07±0.57	33.76±1.38	74.12±0.62	9.79±0.63
BOURNE	78.95±0.36	8.59±0.20	93.50±0.51	12.49±0.78	71.22±0.11	10.06±0.13	82.41±1.09	11.35±1.10	92.23±0.25	11.94±0.10	89.87±0.12	13.64±0.41
+ *LanGen*	80.76±0.98	8.94±0.20	94.23±0.55	13.82±0.83	72.35±0.35	10.93±0.59	83.11±0.39	11.79±0.92	94.20±0.18	14.15±0.37	90.77±0.35	13.98±1.06
+ *JointSDE*	79.11±0.36	7.94±0.42	94.42±0.16	12.73±0.39	72.66±0.23	10.29±0.44	82.94±0.39	12.44±1.01	93.87±0.61	13.73±0.48	91.08±0.19	13.99±0.46
CGADM	81.10±3.96	38.85±3.41	95.24±1.47	44.94±1.75	72.72±2.52	30.89±3.01	70.31±1.33	36.26±5.62	83.70±0.42	51.49±7.00	90.06±2.71	60.49±3.70
+ *LanGen*	84.18±0.39	42.04±1.77	96.31±0.00	45.78±0.98	75.30±0.00	33.02±0.00	71.26±0.81	41.44±0.72	84.45±0.57	58.76±0.81	91.50±1.02	62.46±1.57
+ *JointSDE*	84.28±0.02	42.87±0.97	96.31±0.40	46.49±0.41	75.29±2.14	33.01±0.68	71.55±0.78	43.20±0.92	84.77±0.15	57.63±3.36	91.24±3.31	62.33±0.17

Table 3: Node anomaly detection performance on six real-world benchmark datasets with injected anomalies. Results are reported as mean ± standard deviation (in %). Green and red values indicate improvements and drops over the corresponding backbone model, respectively.

is consistent with the strengthened modeling of normal nodes induced by edge augmentation: by promoting more coherent connections among normal nodes, the detector more accurately captures their feature and structural distributions, reducing false positives and improving the true negative rate, which is more directly reflected by the ROC metric. In contrast, AUPRC emphasizes anomaly-focused precision–recall behavior under severe class imbalance, leading to comparatively smaller gains.

Notably, even for datasets with relatively dense original structures, such as Weibo, Facebook, and BlogCatalog, the proposed strategy yields substantial improvements, with AUROC gains of up to 30% and AUPRC gains of up to 10% in certain settings. This suggests that the benefits extend beyond trivial edge completion and are largely attributable to the spectrum-energy-guided edge diffusion mechanism, which reinforces plausible intra-community connections while suppressing noisy inter-community links, thereby facilitating clearer structural decision boundaries.

4.3 Ablation Study (Q2)

We conduct ablation studies to examine whether the energy formulation $E_\theta(X, A)$ in Eq. 6 effectively guides edge generation in LangGen and whether JointSDE captures the joint evolution of node features and graph structure. For LangGen, we remove or replace individual energy components, including density regularization (*w/o dens*), smoothness regularization (*w/o smo*), both terms (*w/o d&s*), and the Rayleigh Quotient, where the RQ-based energy is replaced by a readout-based alternative (*w/o RQ*). For JointSDE, we ablate generated features (*w/o fea*) or structure (*w/o stru*), edge priority components in Eq. 22 (*w/o score*, *w/o spe*), and the diffusion-based sampling strategy by directly sampling from Gaussian noise (*w/o diff*).

We adopt CoLA as the backbone and report results in Ta-

Methods	Cora	Citeseer	Books	Facebook
CoLA	55.39±1.53	54.05±2.06	49.69±4.20	71.98±1.82
+LangGen	**58.94±1.59**	**55.14±1.03**	**52.95±4.43**	**78.66±1.66**
w/o den	57.11±2.34	52.41±2.20	50.66±3.90	76.63±3.26
w/o smo	58.83±1.57	52.37±2.41	49.0±7.67	74.50±3.21
w/o d&s	57.75±2.19	51.30±3.00	51.42±7.37	73.62±3.17
w/o RQ	56.70±2.06	49.08±7.67	46.57±4.73	74.68±3.51
+JointSDE	**67.42±1.13**	**59.73±4.13**	**57.68±0.98**	**81.36±0.71**
w/o fea	60.91±3.18	58.43±1.71	49.67±5.33	76.94±2.48
w/o stru	48.59±6.90	42.59±7.81	45.84±6.07	78.22±2.13
w/o score	64.09±3.09	48.65±9.39	52.99±6.06	77.29±3.29
w/o spe	63.73±5.16	45.98±10.04	47.89±7.22	77.68±3.19
w/o diff	47.21±4.31	48.85±3.07	50.72±7.23	78.03±2.31

Table 4: Ablation Result on LangGen and JointSDE.

Methods	Cora	Citeseer	Books	Facebook
CoLA	55.39±1.53	54.05±2.06	49.69±4.20	71.98±1.82
+ edge drop	52.50±0.71	50.48±1.44	50.85±1.05	69.22±0.88
+ edge add	51.63±0.72	49.61±0.72	50.80±1.12	72.32±0.62
+LangGen	58.94±1.59	55.14±1.03	52.95±4.43	78.66±1.66
+JointSDE	67.42±1.13	59.73±4.13	57.68±0.98	81.36±0.71

Table 5: Comparison with Other Edge Augmentations.

ble 4. For LangGen, removing any component consistently degrades performance, with *w/o RQ* yielding the weakest or even negative gains, underscoring the critical role of spectral energy in guiding meaningful edge addition, while density and smoothness terms further stabilize generation and improve robustness. For JointSDE, ablating either features or structure causes clear performance drops, confirming the necessity of joint feature–structure evolution. Although *w/o fea* underperforms the full model, it remains superior to CoLA due to the intermediate noising–denoising process. In contrast, *w/o diff* exhibits dataset-dependent behavior, and the comparison between *w/o spe* and *w/o score* shows that spectral guidance is essential for edge priority estimation, as learned scores alone fail to capture global structural evolution.

4.4 Edge Augmentation Comparison (Q3)

Table 5 compares our methods with random edge deletion and random edge addition, where both perturbation ratios are fixed at 10% following common practice; performance gains and degradations are highlighted in green and red, respectively.

Across datasets, both random edge deletion and random edge addition exhibit highly inconsistent behavior, improving performance on some datasets while degrading it on others, indicating that such heuristic perturbations fail to provide a stable inductive bias for graph anomaly detection and typically require dataset- and method-specific tuning.

In contrast, LangGen and JointSDE operate at a higher structural level and leverage spectral graph principles to guide edge evolution, yielding consistent and robust performance gains across datasets without extensive hyperparameter tuning, and thereby outperforming conventional edge-level augmentation strategies.

4.5 Computational Efficiency (Q4)

In LangGen, the dominant cost lies in computing the Rayleigh Quotient within the energy function, which can be optimized

Methods	Cora	Amazon	BlogCatalog	ACM	Citeseer	PubMed	Weibo	Reddit	Books	Enron	Disney	Facebook
LangGen	0.16	5.36	0.87	20.49	0.27	10.64	1.24	2.98	0.12	4.47	0.10	0.17
JointSDE	44.26	1929.28	160.50	372	67.70	390.62	438.22	961.51	16.54	1485.03	3.50	11.65

Table 6: Sampling Time on All Datasets (in seconds).

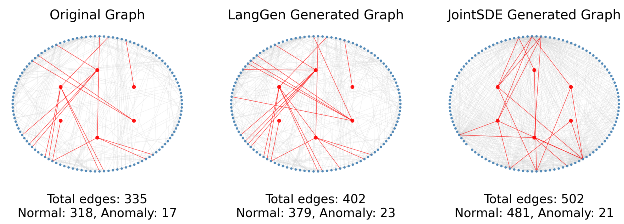


Figure 3: Visualization of graph structures before and after augmentation on the Disney dataset.

to $O(|\mathcal{E}|d)$ for a sparse graph with $|\mathcal{E}|$ edges and feature dimension d. In JointSDE, generation involves feature updates $O(Nd)$, edge importance scoring, and spectral decomposition, where computing the top k eigenvectors contributes $O(k|\mathcal{E}|)$ and score network inference over sampled edges $\mathcal{E}_{\text{sampled}}$ costs $O(|\mathcal{E}_{\text{sampled}}|d)$. Table 6 reports the average generation time per graph across all datasets.

4.6 Qualitative Analysis (Q5)

To examine whether our proposed method effectively increases connectivity among normal nodes, we visualize results on the Disney dataset, which was selected for its small size to facilitate clear and intuitive analysis. As shown in Figure 3, both LangGen and JointSDE generate substantially more edges compared to the original graph, whereas the number of edges involving anomalous nodes remains minimal. This visualization qualitatively confirms that our approach preferentially reinforces connections between normal nodes, highlighting its targeted enhancement capability.

4.7 More Experiments

Detailed descriptions of the method and experiments, comparisons with other augmentation approaches, qualitative analyses, validation on million-scale graphs, and sensitivity studies of hyperparameters and initialization appear in the appendix.

5 Conclusion and Future Work

This work demonstrates that moderate, principled edge addition can substantially improve graph anomaly detection, challenging the long-standing practice of relying on conservative edge dropping while deliberately avoiding edge addition. We introduce two self-supervised, spectrally guided augmentation approaches, in which graph structures are refined in a training-free manner via Langevin dynamics, or jointly evolved with node features through stochastic differential equations. Experiments on 12 benchmark datasets with 7 state-of-the-art GAD models show that the generated graphs consistently enrich connections among normal nodes while largely preserving anomalous structures. Future work will focus on reducing the number of hyperparameters and extending the framework to broader graph settings, including dynamic and structure-only graphs.

Acknowledgments

The works described in this paper are supported by The National Natural Science Foundation of China under Grant No. 62477015; Key Research and Development Program of Guangdong of China under Grant Nos. 2025B0101120004 and 2023B0303010004; The Innovation Team Project for Universities in Guangdong Province in China under Grant No. 2023KCXTD011.

References

[Bandyopadhyay *et al.*, 2020] S. Bandyopadhyay, L. N, S. V. Vivek, and M. N. Murty. Outlier resistant unsupervised deep architectures for attributed network embedding. In *Proc. Int. Conf. Web Search Data Min. (WSDM)*, pages 25–33, 2020.

[Cai *et al.*, 2025] Jinyu Cai, Yunhe Zhang, Fusheng Liu, and See-Kiong Ng. Leveraging diffusion model as pseudo-anomalous graph generator for graph-level anomaly detection. In *Proc. Int. Conf. Mach. Learn. (ICML)*, pages 6409–6425, 2025.

[Cao *et al.*, 2024] Hanqun Cao, Cheng Tan, Zhangyang Gao, Yilun Xu, Guangyong Chen, Pheng-Ann Heng, and Stan Z. Li. A survey on generative diffusion models. *IEEE Trans. Knowl. Data Eng. (TKDE)*, pages 2814–2830, 2024.

[Deng *et al.*, 2022] L. Deng, C. Wu, D. Lian, Y. Wu, and E. Chen. Markov-driven graph convolutional networks for social spammer detection. *IEEE Trans. Knowl. Data Eng. (TKDE)*, pages 12310–12322, 2022.

[Ding *et al.*, 2019] K. Ding, J. Li, R. Bhanushali, and H. Liu. Deep anomaly detection on attributed networks. In *Proc. SIAM Int. Conf. Data Min. (SDM)*, pages 594–602, 2019.

[Dong *et al.*, 2024] X. Dong, X. Zhang, and S. Wang. Rayleigh quotient graph neural networks for graph-level anomaly detection. In *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2024.

[Fan *et al.*, 2020] H. Fan, F. Zhang, and Z. Li. Anomalydae: Dual autoencoder for anomaly detection on attributed networks. In *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, pages 5685–5689, 2020.

[Feng *et al.*, 2025] S. Feng, G. Yu, D. Liu, H. Hu, Y. Luo, H. Lin, and Y.-S. Ong. Mhr: A multi-modal hyperbolic representation framework for fake news detection. *IEEE Trans. Knowl. Data Eng. (TKDE)*, pages 2015–2028, 2025.

[Gao *et al.*, 2023] Y. Gao, X. Wang, X. He, Z. Liu, H. Feng, and Y. Zhang. Addressing heterophily in graph anomaly detection: A perspective of graph spectrum. In *Proc. ACM Web Conf. (WWW)*, pages 1528–1538, 2023.

[He *et al.*, 2024] J. He, Q. Xu, Y. Jiang, Z. Wang, and Q. Huang. Ada-gad: Anomaly-denoised autoencoders for graph anomaly detection. In *Proc. AAAI Conf. Artif. Intell.*, pages 8481–8489, 2024.

[Ho *et al.*, 2020] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Adv. Neural Inf. Process. Syst. (NeurIPS)*, pages 6840–6851, 2020.

[Hu *et al.*, 2025] J. Hu, B. Xiao, H. Jin, J. Duan, S. Wang, Z. Lv, S. Wang, X. Liu, and E. Zhu. Samcl: Subgraph-aligned multiview contrastive learning for graph anomaly detection. *IEEE Trans. Neural Netw. Learn. Syst. (TNNLS)*, pages 1664–1676, 2025.

[Huang *et al.*, 2025] Yi Huang, Jiancheng Huang, Yifan Liu, Mingfu Yan, Jiaxi Lv, Jianzhuang Liu, Wei Xiong, He Zhang, Liangliang Cao, and Shifeng Chen. Diffusion model-based image editing: A survey. *IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI)*, pages 4409–4437, 2025.

[Jin *et al.*, 2024] M. Jin, H. Y. Koh, Q. Wen, D. Zambon, C. Alippi, G. I. Webb, I. King, and S. Pan. A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection. *IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI)*, pages 10466–10485, 2024.

[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 *Proc. Int. Conf. Mach. Learn. (ICML)*, pages 10362–10383, 2022.

[Li *et al.*, 2017] J. Li, H. Dani, X. Hu, and H. Liu. Radar: Residual analysis for anomaly detection in attributed networks. In *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, pages 2152–2158, 2017.

[Li *et al.*, 2025] Jinghan Li, Yuan Gao, Jinda Lu, Junfeng Fang, Congcong Wen, Hui Lin, and Xiang Wang. Diffgad: A diffusion-based unsupervised graph anomaly detector. In *Proc. Int. Conf. Learn. Represent. (ICLR)*, pages 18507–18529, 2025.

[Lin *et al.*, 2024] Y. Lin, J. Tang, C. Zi, H. V. Zhao, Y. Yao, and J. Li. Unigad: Unifying multi-level graph anomaly detection. In *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, pages 136120–136148, 2024.

[Liu *et al.*, 2021] Y. Liu, Z. Li, S. Pan, C. Gong, C. Zhou, and G. Karypis. Anomaly detection on attributed networks via contrastive self-supervised learning. *IEEE Trans. Neural Netw. Learn. Syst. (TNNLS)*, pages 2378–2392, 2021.

[Liu *et al.*, 2024a] J. Liu, M. He, X. Shang, J. Shi, B. Cui, and H. Yin. Bourne: Bootstrapped self-supervised learning framework for unified graph anomaly detection. In *Proc. IEEE Int. Conf. Data Eng. (ICDE)*, pages 2820–2833, 2024.

[Liu *et al.*, 2024b] K. Liu, Y. Dou, X. Ding, X. Hu, R. Zhang, H. Peng, L. Sun, and P. S. Yu. Pygod: A python library for graph outlier detection. *J. Mach. Learn. Res. (JMLR)*, pages 1–9, 2024.

[Liu *et al.*, 2024c] Yixin Liu, Shiyuan Li, Yu Zheng, Qingfeng Chen, Chengqi Zhang, and Shirui Pan. Arc: A generalist graph anomaly detector with in-context learning. *Adv. Neural Inf. Process. Syst. (NeurIPS)*, pages 50772–50804, 2024.

[Luo *et al.*, 2022] X. Luo, J. Wu, A. Beheshti, J. Yang, X. Zhang, Y. Wang, and S. Xue. Comga: Community-aware attributed graph anomaly detection. In *Proc. ACM Int. Conf. Web Search Data Min. (WSDM)*, pages 657–665, 2022.

[Ma *et al.*, 2021] X. Ma, J. Wu, S. Xue, J. Yang, C. Zhou, Q. Z. Sheng, H. Xiong, and L. Akoglu. A comprehensive survey on graph anomaly detection with deep learning. *IEEE Trans. Knowl. Data Eng. (TKDE)*, pages 12012–12038, 2021.

[Ma *et al.*, 2024] X. Ma, R. Li, F. Liu, K. Ding, J. Yang, and J. Wu. Graph anomaly detection with few labels: A data-centric approach. In *Proc. ACM SIGKDD Conf. Knowl. Discov. Data Min. (KDD)*, pages 2153–2164, 2024.

[Ma *et al.*, 2025] X. Ma, F. Liu, J. Wu, J. Yang, S. Xue, and Q. Z. Sheng. Rethinking unsupervised graph anomaly detection with deep learning: Residuals and objectives. *IEEE Trans. Knowl. Data Eng. (TKDE)*, pages 881–895, 2025.

[Pan *et al.*, 2025] J. Pan, Y. Liu, X. Zheng, Y. Zheng, A. W.-C. Liew, F. Li, and S. Pan. A label-free heterophily-guided approach for unsupervised graph fraud detection. In *Proc. AAAI Conf. Artif. Intell.*, pages 12443–12451, 2025.

[Pazho *et al.*, 2023] A. D. Pazho, G. A. Noghre, A. A. Purkayastha, J. Vempati, O. Martin, and H. Tabkhi. A survey of graph-based deep learning for anomaly detection in distributed systems. *IEEE Trans. Knowl. Data Eng. (TKDE)*, pages 1–20, 2023.

[Peng *et al.*, 2018] Z. Peng, M. Luo, J. Li, H. Liu, Q. Zheng, et al. Anomalous: A joint modeling approach for anomaly detection on attributed networks. In *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, pages 3513–3519, 2018.

[Roy *et al.*, 2024] A. Roy, J. Shu, J. Li, C. Yang, O. Elshocht, J. Smeets, and P. Li. Gad-nr: Graph anomaly detection via neighborhood reconstruction. In *Proc. ACM Int. Conf. Web Search Data Min. (WSDM)*, pages 576–585, 2024.

[Song and Ermon, 2019] Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. *Adv. Neural Inf. Process. Syst. (NeurIPS)*, pages 11918–11930, 2019.

[Song *et al.*, 2021a] Yang Song, Conor Durkan, Iain Murray, and Stefano Ermon. Maximum likelihood training of score-based diffusion models. *Adv. Neural Inf. Process. Syst. (NeurIPS)*, pages 1415–1428, 2021.

[Song *et al.*, 2021b] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2021.

[Sun *et al.*, 2025] Z. Sun, M. Xu, G. Liang, and Q. Liu. Unifying adversarial multi-deconfounded learning paradigm for fake news detection. In *Proc. ACM SIGKDD Conf. Knowl. Discov. Data Min. (KDD)*, pages 2572–2583, 2025.

[Tang *et al.*, 2022] J. Tang, J. Li, Z. Gao, and J. Li. Rethinking graph neural networks for anomaly detection. In *Proc. Int. Conf. Mach. Learn. (ICML)*, pages 21076–21089, 2022.

[Vignac *et al.*, 2023] Clément Vignac, Igor Krawczuk, Antoine Siraudin, Bohan Wang, Volkan Cevher, and Pascal Frossard. Digress: Discrete denoising diffusion for graph generation. In *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2023.

[Wang *et al.*, 2025] Y. Wang, Y. Liu, X. Shen, C. Li, R. Miao, K. Ding, Y. Wang, S. Pan, and X. Wang. Unifying unsupervised graph-level anomaly detection and out-of-distribution detection: A benchmark. In *Proc. Int. Conf. Learn. Represent. (ICLR)*, pages 38164–38190, 2025.

[Wei *et al.*, 2025] Chunyu Wei, Haozhe Lin, Yueguo Chen, and Yunhai Wang. Conditional diffusion anomaly modeling on graphs. In *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2025.

[Westphal *et al.*, 2025] C. Westphal, S. Hailes, and M. Musolesi. Feature selection for network intrusion detection. In *Proc. ACM SIGKDD Conf. Knowl. Discov. Data Min. (KDD)*, pages 1599–1610, 2025.

[Xu *et al.*, 2024] F. Xu, N. Wang, H. Wu, X. Wen, X. Zhao, and H. Wan. Revisiting graph-based fraud detection in sight of heterophily and spectrum. In *Proc. AAAI Conf. Artif. Intell.*, pages 9214–9222, 2024.

[Yang *et al.*, 2024] Yiyuan Yang, Ming Jin, Haomin Wen, Chaoli Zhang, Yuxuan Liang, Lintao Ma, Yi Wang, Chenghao Liu, Bin Yang, and Zenglin Xu. A survey on diffusion models for time series and spatio-temporal data. *ACM Comput. Surv. (CSUR)*, pages 1–39, 2024.

[Yu *et al.*, 2024] J. Yu, H. Wang, X. Wang, Z. Li, L. Qin, W. Zhang, J. Liao, Y. Zhang, and B. Yang. Temporal insights for group-based fraud detection on e-commerce platforms. *IEEE Trans. Knowl. Data Eng. (TKDE)*, pages 951–965, 2024.

[Zhang *et al.*, 2022] G. Zhang, Z. Li, J. Huang, J. Wu, C. Zhou, J. Yang, and J. Gao. efraudcom: An e-commerce fraud detection system via competitive graph neural networks. *ACM Trans. Inf. Syst. (TOIS)*, pages 1–29, 2022.

[Zheng *et al.*, 2021] Y. Zheng, M. Jin, Y. Liu, L. Chi, K. T. Phan, and Y. P. P. Chen. Generative and contrastive self-supervised learning for graph anomaly detection. *IEEE Trans. Knowl. Data Eng. (TKDE)*, pages 12220–12233, 2021.

[Zheng *et al.*, 2025] J. Zheng, C. Yang, T. Zhang, L. Cao, B. Jiang, X. Fan, X. Wu, and X. Zhu. Dynamic spectral graph anomaly detection. In *Proc. AAAI Conf. Artif. Intell.*, pages 13410–13418, 2025.