

λ -Diffusion Language Models: Controllable Generalization of Continuous and Discrete Diffusion Dynamics

DongNyeong Heo
Ulsan National Institute of
Science and Technology
Ulsan, South Korea
sjglsks@gmail.com

Hyunmin Song
Ulsan National Institute of
Science and Technology
Ulsan, South Korea
hyunminsong@unist.ac.kr

Taehwan Kim
Ulsan National Institute of
Science and Technology
Ulsan, South Korea
taehwankim@unist.ac.kr

Abstract

Diffusion language models (DLMs) have emerged as a promising alternative to autoregressive language models. While discrete DLMs achieve strong performance across language generation tasks, continuous DLMs remain relatively underexplored despite their smooth latent-space dynamics. In this work, we present a unified embedding-space interpretation of continuous and discrete DLMs. Within this framework, we analyze continuous diffusion dynamics as a retention channel that preserves intermediate denoising information during sampling. Motivated by this perspective, we propose λ -DLM, which smoothly interpolates between continuous and discrete diffusion dynamics through a controllable parameter λ . By training with randomly sampled λ , a single model generalizes across continuous, discrete, and intermediate diffusion dynamics, unifying diverse retention behaviors within one framework. Unlike prior methods that rely on fixed or self-confidence-guided diffusion dynamics, λ -DLM enables controllable exploitation–exploration trade-offs without re-training. Experiments on LM1B and OWT unconditional generation benchmarks show that λ -DLM consistently outperforms strong baselines. Moreover, varying λ enables semantically diverse generation without additional training. Compared with temperature sweeps, varying λ shows a more structured relationship between lexical diversity and sentence-embedding-based mode coverage.

1 Introduction

Diffusion language models (DLMs) have recently emerged as a promising alternative to autoregressive language models (Austin et al., 2021; Li et al., 2022; Sahoo et al., 2024; Nie et al., 2026). Unlike autoregressive models (Radford et al., 2019; Brown et al., 2020), DLMs iteratively refine noisy representations through denoising, enabling paral-

lel generation and self-correcting sampling dynamics. Existing DLMs can largely be categorized into discrete (He et al., 2023; Shi et al., 2024; Sahoo et al., 2024; Nie et al., 2026) and continuous (Li et al., 2022; Gong et al., 2022; Dieleman et al., 2022; Gao et al., 2024) variants. While discrete DLMs have shown strong empirical performance, continuous DLMs remain relatively underexplored despite their fine-grained latent-space dynamics.

To better understand both paradigms within a unified framework, we present an embedding-space interpretation of continuous and discrete diffusion dynamics (Section 3). Under this perspective, continuous diffusion dynamics transport a unimodal Gaussian from data embeddings toward terminal embeddings, whereas discrete diffusion dynamics form an anchored multimodal mixture of Gaussians (MoG) whose component means remain fixed while their mixture weights evolve over time. Figure 1 illustrates these dynamics through the top and bottom diffusion pathways corresponding to the continuous and discrete dynamics, respectively.

Importantly, during sampling, continuous diffusion dynamics progressively transport the current latent state using previously denoised predictions and terminal embeddings, thereby carrying information from earlier denoising steps forward. In contrast, discrete diffusion dynamics update latent states through fixed anchored modes, limiting the direct propagation of previous denoising outcomes. This distinction motivates our retention-based interpretation of diffusion dynamics. We use *retention* to denote the degree to which information from the model’s previous denoising prediction is carried into subsequent reverse-diffusion states. Under this view, stronger retention promotes refinement of previous predictions (exploitation), whereas weaker retention allows broader deviation from them (exploration). Controlling retention therefore provides a mechanism for modulating the balance between generation fidelity and diversity.

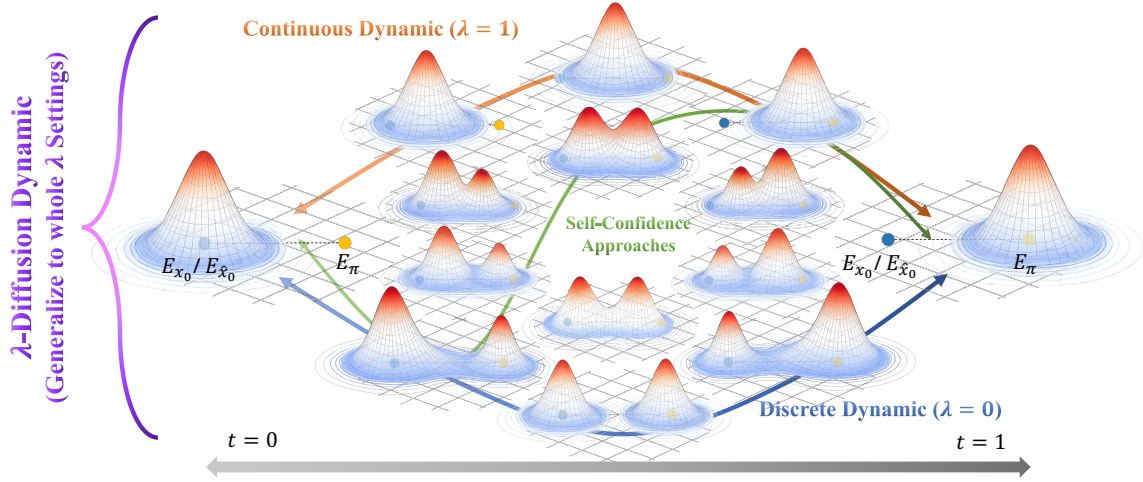


Figure 1: Illustration of continuous (top, $\lambda = 1$), discrete (bottom, $\lambda = 0$), and intermediate diffusion dynamics based on absorbing-state DLMs. Continuous and discrete DLMs follow fixed trajectories, while prior self-confidence-based approaches adaptively traverse intermediate trajectories between them. In contrast, λ -DLM generalizes across the full spectrum of diffusion dynamics, enabling controllable denoising behaviors through a single parameter λ .

Several recent approaches, including hybrid DLMs and soft-absorbing DLMs (Zheng et al., 2025; Pynadath et al., 2025; Hersche et al., 2025; Jo et al., 2025), can be interpreted as introducing such retention mechanisms into discrete DLMs. However, these methods determine retention strength using the model’s own confidence scores, implicitly assuming that the model can reliably evaluate its predictions. Our analysis in Section 3.3 shows that this assumption can be problematic: more reliable confidence estimates from larger DLMs consistently improve generation quality over self-confidence-based retention.

Motivated by these observations, we propose a λ -diffusion dynamic that smoothly interpolates between continuous and discrete diffusion dynamics through a controllable parameter λ (see the middle pathways in Figure 1). By simply varying λ , the diffusion trajectory can transition between continuous, discrete, and intermediate dynamics. Based on this dynamic, we further propose λ -DLM, which is trained with randomly sampled λ values and therefore generalizes across diverse diffusion dynamics within a single model. We hypothesize that this generalized capability enables flexible control over retention behavior during denoising, allowing modulation of the exploitation–exploration trade-off throughout the entire sampling process. In contrast, prior methods are trained only under fixed or self-confidence-guided diffusion dynamics, making it difficult to flexibly control retention behavior beyond their predefined denoising trajectories.

Experiments on LM1B (Chelba et al., 2013) and OpenWebText(OWT) (Gokaslan and Cohen, 2019) unconditional generation benchmarks show that λ -DLM achieves strong MAUVE performance across model scales and NFE settings, with λ -UDLM showing the most consistent gains, particularly in low-NFE regimes. Our diversity analysis further shows that, compared with temperature sweeps on λ -UDLM and SoftMask, varying λ yields a clearer relationship between lexical diversity and sentence-embedding-based mode coverage. Taken together, these results support λ as a useful inference-time control dimension for modulating retention and the exploitation–exploration balance in unconditional generation.

Our contributions are summarized as follows:

- **Unified Interpretation and Retention Perspective:** We present a unified embedding-space interpretation of continuous and discrete diffusion dynamics and reinterpret continuous dynamics as a retention mechanism.
- **λ -Diffusion Dynamic and λ -DLM:** We propose λ -diffusion dynamic and λ -DLM, which generalize across diverse diffusion trajectories through controllable retention strengths.
- **Exploitation–Exploration Controllability:** We show that λ -DLM effectively controls exploitation and exploration, improving both generation quality and semantic diversity.

2 Background

2.1 Notations

Throughout this paper, we handle both discrete tokens and continuous embeddings. A token is represented as a one-hot vector $x \in \{0, 1\}^{|\mathcal{V}|}$, and its embedding as $z = Ex \in \mathbb{R}^d$, where $E \in \mathbb{R}^{d \times |\mathcal{V}|}$ is the embedding matrix. DLMs (Li et al., 2022; Austin et al., 2021) define diffusion processes over either the discrete token space or the continuous embedding space. We denote the noised token and embedding at timestep t by x_t and z_t , respectively, where x_0 denotes the original clean token.

2.2 Continuous Diffusion Language Models

In continuous DLMs (Li et al., 2022; Gong et al., 2022; Dieleman et al., 2022; Gao et al., 2024), a token x_0 is mapped to an embedding $z_0 = Ex_0$, and the noising process is defined as

$$q^c(z_t|x_0) = \mathcal{N}(z_t; \sqrt{\alpha_t}z_0 + (1 - \sqrt{\alpha_t})E\pi, \sigma_t^2 I),$$

where α_t and σ_t are timestep-dependent noise schedules, $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$, and $E\pi$ denotes the terminal embedding, typically set to the zero vector. During denoising, the backbone model predicts the original embedding z_0 from z_t and is trained with mean squared error (MSE) loss.

2.3 Discrete Diffusion Language Models

In discrete DLMs (Austin et al., 2021; He et al., 2023; Sahoo et al., 2024; Nie et al., 2026), the diffusion process is defined over discrete token states. In absorbing-state DLMs such as MDLM (Sahoo et al., 2024), tokens are progressively replaced with [MASK] according to

$$q^d(x_t|x_0) = \begin{cases} \bar{\alpha}_t, & x_t = x_0 \\ 1 - \bar{\alpha}_t, & x_t = [\text{MASK}]. \end{cases} \quad (1)$$

where $\bar{\alpha}_t$ controls the masking probability over timesteps. During denoising, the backbone model predicts $p_\theta(x_0|x_t)$ and is trained with the negative evidence lower bound (NELBO) objective.

Uniform-state DLMs (UDLMs) (Austin et al., 2021; Sahoo et al., 2025) similarly define discrete noising processes, except that tokens converge toward uniformly sampled vocabulary tokens rather than a dedicated [MASK] token.

2.4 Hybrid and Soft-Absorbing DLMs

In this paper, we focus on retention mechanisms in DLM architectures. Under this perspective, hybrid DLMs and soft-absorbing DLMs are closely

related model families. Hybrid DLMs (Zheng et al., 2025; Pynadath et al., 2025) combine continuous and discrete diffusion dynamics, where masked tokens typically follow continuous dynamics and unmasked tokens follow discrete dynamics. During sampling, the continuous component is constructed using the model’s previous denoising outputs due to the absence of ground-truth tokens. Soft-absorbing DLMs (Hersche et al., 2025; Jo et al., 2025) relax the strict discrete-state assumption by introducing intermediate states between masked and unmasked tokens, typically constructed from previous denoising outputs during both training and sampling.

In practice, these methods retain previous denoising results through weighted combinations of candidate token embeddings and terminal embeddings. The weights are usually determined by the model’s own confidence scores, assuming that the model can reliably estimate prediction correctness. We refer readers to SoftMask (Hersche et al., 2025) for representative implementations of such self-confidence-based retention mechanisms.

3 A Unified Embedding-Space View of Continuous and Discrete DLMs

Although discrete DLMs define diffusion processes over discrete token states, the noised tokens x_t are ultimately mapped into continuous embeddings through the backbone embedding table. This motivates us to analyze discrete diffusion dynamics within the same embedding-space framework as continuous DLMs. From this perspective, discrete DLMs can be interpreted as inducing diffusion dynamics over continuous representations through the pushforward distribution of discrete transition probabilities (Eq. 1).

3.1 Continuous Relaxation

To analyze discrete diffusion dynamics in embedding space, we adopt a continuous-relaxation approximation using embedding matrix $E \in \mathbb{R}^{d \times |\mathcal{V}|}$:

$$q_{\text{relax}}(z|x; E) = \mathcal{N}(z; Ex, \sigma^2 I) \triangleq \mathcal{N}_{Ex}^\sigma(z),$$

$$p_{\text{relax}}(x|z; E) \approx \text{softmax}(E^\top z / \sigma^2),$$

where σ controls the relaxation variance. The detailed derivation and assumptions are provided in Appendix A.1. Intuitively, the relaxed continuous variable z forms a Gaussian centered at the embedding of the discrete token x .

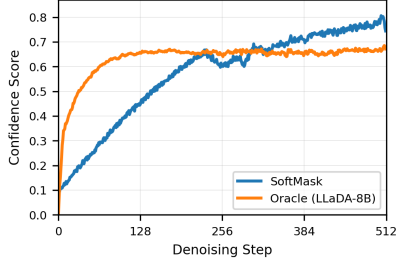


Figure 2: Confidence score trends of SoftMask and LLaDA-8B, averaged over 256 unconditional generations.

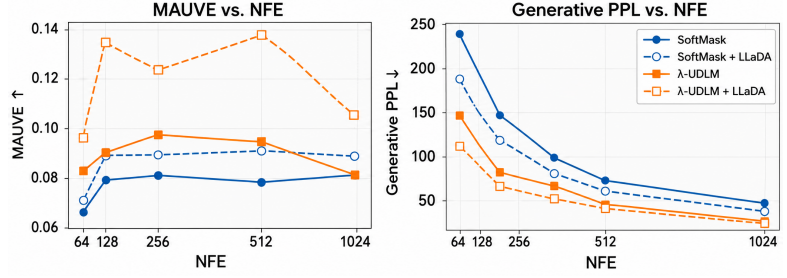


Figure 3: LLaDA-guided generation results across number of forward evaluation (NFE) settings. λ -UDLM benefits more from LLaDA-guidance than SoftMask due to its training across diverse denoising trajectories.

3.2 Diffusion Dynamics in Embedding Space

Using this relaxation, the embedding-space noising distribution induced by discrete diffusion can be derived by marginalizing over the intermediate discrete state x_t :

$$\begin{aligned} q^d(z_t|x_0) &= \sum_{x_t \sim q^d(x_t|x_0)} q_{\text{relax}}(z_t|x_t) q^d(x_t|x_0) \\ &= \bar{\alpha}_t \mathcal{N}_{E_{x_0}}^\sigma + \frac{(1 - \bar{\alpha}_t)}{|\Pi|} \sum_{\pi \in \Pi} \mathcal{N}_{E_\pi}^\sigma, \quad (2) \end{aligned}$$

where Π denotes the set of terminal states. Notably, the resulting pushforward distribution forms an MoG whose mixture weights evolve according to $\bar{\alpha}_t$. In absorbing-state DLMs, the second term simplifies to $(1 - \bar{\alpha}_t) \mathcal{N}_{E[\text{MASK}]}^\sigma$.

Based on derived formulation and continuous DLM’s noising distribution, we characterize continuous and discrete dynamics as *transporting a unimodal Gaussian* and *anchored multimodal MoG*, respectively. Figure 1 illustrates these dynamics as the top (continuous)¹ and bottom (discrete) diffusion trajectories. Both share the same initial and terminal distributions centered at the data and terminal embeddings, respectively, but differ in their intermediate dynamics. Continuous diffusion progressively transports a unimodal Gaussian toward the terminal embedding, whereas discrete diffusion keeps Gaussian modes anchored at the data and terminal embeddings while modulating only their mixture weights over time. In Appendix A.2, we further show that the same characterization also holds for reverse posterior denoising dynamics.

Importantly, during sampling ($t : 1 \rightarrow 0$), the reverse posterior distributions are approximated using the model’s own predictions, replacing the ground-truth embedding with previously denoised

outputs. From this perspective, continuous diffusion dynamics naturally retain previous denoising results through the evolving latent state, whereas discrete diffusion dynamics restrict such retention due to their fixed Gaussian modes anchored at the data and terminal embeddings. This suggests that continuous diffusion dynamics inherently provide a retention channel over previous denoising results.

3.3 Prior Retention Methods and Limitations

Under the retention-channel perspective, hybrid and soft-absorbing DLMs can be viewed as adaptive retention methods guided by self-confidence. The pure continuous and pure discrete DLMs can be viewed as stronger- and weaker-retention endpoints, respectively. As illustrated in Figure 1, self-confidence-based methods can therefore be interpreted as intermediate denoising trajectories between these two extremes.

Although self-confidence-based retention methods substantially improve over pure DLMs, they can also reinforce the model’s own biases and prediction errors. To investigate this issue, we compare self-confidence scores with those estimated by a larger pre-trained DLM, LLaDA-8B (Nie et al., 2026), which we treat as an external stronger confidence estimator. See Appendix B.1 for the detail of the external confidence estimation procedure. We conduct this analysis using SoftMask (Hersche et al., 2025) models trained on the tiny-scale LM1B setting (Section 5.2; Table 1). Figure 2 compares the confidence trends of SoftMask and LLaDA-8B on the same denoising samples generated using SoftMask’s likelihoods. SoftMask assigns lower confidence during early denoising stages but becomes more confident than LLaDA during later stages, suggesting that self-confidence estimates can both under- and overestimate generation quality. More importantly, Figure 3 shows

¹For visualization, we fix the variance across timesteps and use the same terminal state as the discrete DLM.

that replacing self-confidence-based retention with LLaDA’s confidence consistently improves both MAUVE and generative PPL. This suggests that self-confidence-guided denoising trajectories can constrain the model’s capability for better generation quality.

4 λ -Diffusion Language Model

We propose the λ -diffusion dynamic, which smoothly controls retention strength through a single parameter $\lambda \in [0, 1]$. Based on this dynamic, we develop λ -DLM by incorporating it into standard DLMs. By training with randomly sampled λ values, λ -DLM generalizes across diverse diffusion trajectories and retention behaviors.

4.1 λ -Diffusion Dynamic

Based on the relaxed discrete noising distribution in Eq. 2, we propose the λ -diffusion dynamic, which interpolates between continuous and discrete diffusion dynamics through

$$\begin{aligned} q^\lambda(z_t^\lambda|x_0) &= \sum_{x_t \sim q^d(x_t|x_0)} w_{x_t}(t) \mathcal{N}_{\mu_\lambda(x_0, x_t, t)}^{\sigma_\lambda(t)}, \\ \mu_\lambda(x_0, x_t, t) &= \lambda \mu^c(x_0, x_t, t) + (1 - \lambda) \mu^d(x_t), \\ w_{x_t}(t) &= \begin{cases} \bar{\alpha}_t, & x_t = x_0 \\ (1 - \bar{\alpha}_t)/|\Pi|, & x_t \in \Pi, \end{cases} \end{aligned}$$

where $\mu^c(x_0, x_t, t)$ and $\mu^d(x_t)$ denote the mean of continuous and continuously relaxed discrete dynamics at timestep t , respectively. The continuous component transports the mean between Ex_0 and Ex_t , while the discrete component remains anchored at Ex_t . The parameter λ interpolates between these two dynamics, recovering pure continuous and discrete dynamics when $\lambda = 1$ and $\lambda = 0$, respectively. Figure 4 illustrates this process. See Appendix A.3 for the derivation of variance term $\sigma_\lambda(t)$ in our λ -diffusion dynamic.

Under this formulation, the smoothly incorporated continuous component naturally acts as a retention channel during sampling, with retention strength controlled by λ . We therefore interpret λ as controlling the exploitation–exploration trade-off: large λ values strongly retain previous denoising results, whereas small λ values encourage broader exploration near terminal states.

For the denoising process, using Bayes’ rule, we derive the exact reverse posterior under the λ -

diffusion dynamic as follows:

$$q^\lambda(z_s^\lambda|z_t^\lambda, x_0) = \sum_{x_t} q(x_t|z_t^\lambda, x_0) \sum_{x_s} q^d(x_s|x_t, x_0) q^\lambda(z_s^\lambda|x_s, x_0),$$

where $s < t$. Because x_t lies in a finite discrete support set, the conditional distribution admits a closed-form solution for both MDLM and UDLM architectures. Detailed derivations are provided in Appendix A.4.

4.2 Training and Sampling of λ -DLM

Based on Section 4.1, we develop λ -DLM by integrating the proposed λ -diffusion dynamic into modern DLM frameworks. Except for the modified noising and denoising processes, we largely follow MDLM (Sahoo et al., 2024) and UDLM (Austin et al., 2021). We adopt DiT (Peebles and Xie, 2023) as the backbone architecture and incorporate the additional λ condition through a lightweight MLP with SiLU activation (Elfwing et al., 2018). The processed λ embedding is concatenated with the timestep embedding and injected into the DiT conditioning pathway.

Algorithm 1 summarizes the training procedure. To generalize across diverse denoising trajectories, we randomly sample $\lambda \sim \text{Uniform}(0, 1)$ at each iteration and conduct noising/denoising using the sampled value. For sampling, we follow ancestral sampling with the derived reverse posterior distribution, replacing the unavailable ground-truth token x_0 with the model prediction $\hat{x}_0$. Since the reverse posterior involves nested marginalizations over x_s and x_t , we employ Monte Carlo approximation using sampled discrete states. Empirically, a single-sample approximation was sufficient to preserve generation quality while substantially improving sampling efficiency. The overall procedure is summarized in Algorithm 2.

4.3 Additional Training Techniques

Using ground-truth information in the continuous component can introduce training–sampling discrepancy because ground-truth tokens are unavailable during inference. However, ground-truth supervision also provides stable learning signals during early training. To balance these trade-offs, we propose a two-pass training strategy with an annealed warmup schedule.

In the first pass, λ -diffusion noising and denoising are performed using ground-truth tokens to

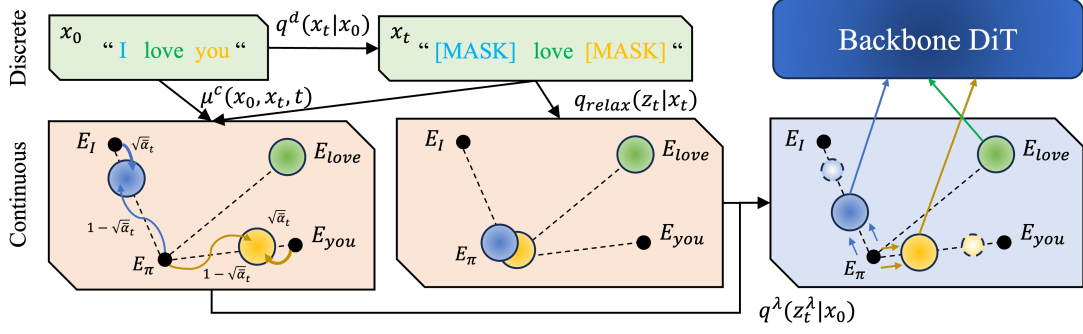


Figure 4: Pipeline of λ -DLM. A discrete noising process first samples the noised token state $x_t \sim q^d(x_t|x_0)$. The proposed λ -diffusion dynamic then constructs a continuous latent representation by interpolating between the transported continuous mean $\mu^c(x_0, x_t, t)$ and the anchored discrete mean E_{x_t} . The resulting latent variable $z_t^\lambda \sim q^\lambda(z_t^\lambda|x_0)$ is finally processed by the backbone DiT model for denoising prediction.

Algorithm 1: One Training Step of λ -DLM

Input: Clean x_0 , timestep t , $\lambda \sim \text{Uniform}(0, 1)$
Input: Iteration n , warmup W , skip schedule p_{sk}

- 1 **Sample** $z_t^\lambda \sim q^\lambda(z_t^\lambda|x_0)$ // λ -diffusion by GT
- 2 $p_0^{(1)} \leftarrow p_\theta(\cdot|z_t^\lambda, t, \lambda)$
- 3 **if** $n < W$ **with probability** $p_{\text{sk}}(n)$ **then**
- 4 $\theta \leftarrow \text{Update}(\theta, \nabla_\theta \mathcal{L}_{\text{DLM}}(p_0^{(1)}, x_0, t))$
- 5 **else**
- 6 **Sample** $\tilde{x}_0 \sim p_0^{(1)}$
- 7 **Sample** $\tilde{z}_t^\lambda \sim q^\lambda(\tilde{z}_t^\lambda|\tilde{x}_0)$ // λ -diffusion by $\hat{x}_0$
- 8 $p_0^{(2)} \leftarrow p_\theta(\cdot|\tilde{z}_t^\lambda, t, \lambda)$
- 9 $\theta \leftarrow \text{Update}(\theta, \nabla_\theta \mathcal{L}_{\text{DLM}}(p_0^{(2)}, x_0, t))$
- 10 **end**

Algorithm 2: Sampling of λ -DLM

Input: Trained model p_θ , selected λ , steps T

- 1 Initialize terminal token $x_T \in \Pi$
- 2 **Sample** $z_T^\lambda \sim q^\lambda(z_T^\lambda|x_T)$
- 3 **for** $t = T, \dots, 1$ **do**
- 4 $s \leftarrow t - 1$
- 5 **Sample** $\hat{x}_0 \sim p_\theta(\cdot|z_t^\lambda, t, \lambda)$
- 6 **Sample** $(x_t, x_s) \sim q(x_t|z_t^\lambda, \hat{x}_0) \& q^d(x_s|x_t, \hat{x}_0)$
// High $\lambda \rightarrow$ Strong Retention
// Low $\lambda \rightarrow$ Weak Retention
- 7 **Sample** $z_s^\lambda \sim q^\lambda(z_s^\lambda|x_s, \hat{x}_0)$
- 8 $z_t^\lambda \leftarrow z_s^\lambda$
- 9 **end**

Output: Generated sample $\hat{x}_0$

obtain initial denoising predictions. In the second pass, the same processes are repeated using the model’s gradient-detached predictions instead of ground-truth tokens, and the resulting outputs are used for backpropagation. During early warmup iterations, the second pass is probabilistically skipped, and the model is directly optimized using the ground-truth-conditioned process. The skip probability is monotonically annealed throughout the warmup stage. The detailed procedure is summarized in Algorithm 1.

The proposed strategy is related to self-conditioning (Strudel et al., 2022) and soft-absorbing DLM training frameworks. Unlike prior methods that directly mix $\hat{x}_0$ into the latent state, our method transforms denoised predictions through the continuous diffusion component to preserve consistency with the timestep-conditioned noise level.

5 Experiment Results

Our experiments consist of unconditional generation evaluations on LM1B (Chelba et al., 2013)

and OWT (Gokaslan and Cohen, 2019) and faithful diversity analyses. These experiments evaluate whether λ -DLM effectively controls the exploitation–exploration trade-off, leading to improved generation quality and semantic diversity.

5.1 Experimental Setup

5.1.1 General Setup

We conduct experiments on LM1B (Chelba et al., 2013) and OpenWebText (OWT). LM1B contains approximately 0.8B training tokens and around 6K sentences for validation. OWT consists of approximately 8M documents and 9B tokens. Following the preprocessing protocol of prior work², we reserve the last 100K OWT documents for validation and use the remaining documents for training. For both datasets, we adopt the corresponding preprocessing and tokenization pipelines from the same repository. We compare our proposed models against recent strong DLM baselines, including UDLM (Austin et al., 2021), MDLM (Sahoo

²<https://github.com/s-sahoo/duo>

et al., 2024), ReMDM (Wang et al., 2026), and SoftMask (Hersche et al., 2025).

5.1.2 Models and Training

For LM1B, we adopt tiny- and small-scale DiT backbones (Peebles and Xie, 2023) with 8 and 12 Transformer layers (Vaswani et al., 2017), resulting in 24M and 139M parameters, respectively. We largely follow the standard LM1B training setup of prior work (Zhang et al., 2026), using 200K training iterations. For OWT, we use a small DiT backbone with 169M parameters and a context length of 1024, and train each model for 100K updates. Additional architecture and training details are provided in Appendix B.

5.1.3 λ -DLM Implementation and Training

We instantiated the proposed framework on both ReMDM and UDLM architectures, resulting in λ -ReMDM and λ -UDLM, corresponding to absorbing-state and uniform-state variants, respectively. For the continuous component, we adopted the standard continuous DLM noise schedule (Li et al., 2022) independently from the discrete schedule. By setting $\sigma^d = 0$ in the continuous relaxation component, we preserved the original discrete diffusion dynamics within λ -DLM. For the annealing strategy in Section 4.3, we used a 10K-step warmup stage with a linearly decreasing second-pass skipping probability p_{sk} from 1 to 0.

After training, we select the inference-time λ schedule using held-out tuning seeds. We consider a constant- λ schedule and a timestep-dependent log-shaped schedule inspired by the confidence tendency of LLaDA in Figure 2:

$$\lambda_{\text{const}}(u) = \lambda_0, \quad (3)$$

$$\lambda_{\text{log}}(u; r) = \frac{\log(1 + ru)}{\log(1 + r)}, \quad u \in [0, 1], \quad (4)$$

where u denotes the normalized denoising progress and r controls the curvature (log-rate) of the log-shaped schedule. For each NFE setting, we perform a brute-force search over the corresponding schedule parameter, using $\lambda_0 \in [0, 1]$ with a step size of 0.1 for the constant schedule and $r \in [1, 10]$ with a step size of 1 for the log-shaped schedule. MAUVE (Pillutla et al., 2021) is used as the primary selection criterion. To avoid test leakage, schedule selection and final evaluation are conducted using separate random seeds. Empirically, the constant schedule performs robustly across most LM1B and λ -ReMDM settings, whereas

OWT λ -UDLM benefits from the log-shaped schedule. We therefore use the constant schedule as the default and report the log-shaped variant where it provides meaningful improvements. All selected parameter values are summarized in Table 6.

5.2 LM1B Unconditional Generation

Tables 1 and 2 summarize unconditional generation results on LM1B across different NFE settings and different seeds. See Tables 9 and 10 for standard deviation results. Each model generated 1024 unconditional samples for evaluation. We report MAUVE, generative PPL measured by GPT2-large (Radford et al., 2019), and sample entropy. Since generative PPL is known to be unstable in open-ended generation settings (Wang et al., 2026), we mainly focus on MAUVE. Overall, λ -DLM achieves strong MAUVE across most NFE settings and model scales, with λ -UDLM showing particularly consistent gains in low-NFE regimes. At the same time, the MAUVE-selected λ settings often exhibit lower sample entropy than the corresponding baselines, indicating that stronger retention can concentrate the token-level output distribution while improving distributional quality.

We additionally apply our λ tuning strategy to trained SoftMask models without retraining, denoted as “SoftMask+ λ ”. Compared with the original SoftMask, post-hoc λ control improves MAUVE in many settings, indicating that retention strength itself provides a useful inference-time control axis. However, λ -UDLM generally achieves stronger MAUVE than SoftMask+ λ , suggesting that its gains arise not only from post-hoc λ selection but also from training across diverse denoising trajectories. Additional external-confidence-guidance results in Figure 3 further support the benefit of more effective retention control.

To directly verify the contribution of our training strategy, Table 3 compares the standard random- λ , two-pass training with fixed- λ and one-pass variants. The standard strategy achieves the highest MAUVE for both λ -ReMDM and λ -UDLM, while fixed continuous ($\lambda = 1$), discrete ($\lambda = 0$), intermediate, and one-pass variants all degrade distributional quality. These results support the importance of both training across diverse diffusion trajectories and reducing the training-sampling discrepancy through the second pass. Notably, fixed $\lambda = 1$ for λ -UDLM attains very low generative PPL but much lower MAUVE, again showing that PPL alone can be misleading for open-ended generation.

Table 1: Unconditional generation results of **tiny DiT**-based models on LM1B. We report the mean MAUVE($\uparrow$) and generative PPL($\downarrow$) over three runs. Values in parentheses denote sample entropy. For λ -based methods, λ is selected with respect to MAUVE.

Model	NFE = 64		NFE = 128		NFE = 256		NFE = 512		NFE = 1024	
	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL
UDLM	0.061	400.9 (4.3)	0.059	391.8 (4.3)	0.061	381.8 (4.3)	0.062	377.0 (4.3)	0.062	376.5 (4.3)
MDLM	0.061	345.2 (4.3)	0.062	306.7 (4.2)	0.064	275.6 (4.2)	0.066	243.0 (4.2)	0.065	210.5 (4.2)
ReMDM	0.065	261.2 (4.2)	0.070	169.8 (4.1)	0.071	110.3 (4.0)	0.071	72.9 (3.8)	0.070	50.4 (3.6)
SoftMask	0.069	239.0 (4.2)	0.072	152.0 (4.1)	0.082	96.4 (4.0)	0.084	63.9 (3.9)	0.080	44.9 (3.7)
SoftMask+ λ	0.071	168.4 (4.1)	0.072	169.3 (4.1)	0.084	70.5 (3.9)	0.081	76.6 (3.9)	0.086	54.5 (3.8)
λ -ReMDM	0.067	210.2 (3.9)	0.074	108.4 (3.7)	0.076	90.4 (3.6)	0.079	75.4 (3.2)	0.072	62.3 (2.6)
λ -UDLM	0.099	93.9 (4.0)	0.095	78.9 (3.9)	0.099	47.8 (3.6)	0.096	30.5 (3.0)	0.081	17.5 (2.0)

Table 2: Unconditional generation results of **small DiT**-based models on LM1B. We report the mean MAUVE($\uparrow$) and generative PPL($\downarrow$) over three runs. Values in parentheses denote sample entropy. For λ -based methods, λ is selected with respect to MAUVE.

Model	NFE = 64		NFE = 128		NFE = 256		NFE = 512		NFE = 1024	
	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL
UDLM	0.078	237.3 (4.3)	0.081	229.7 (4.2)	0.088	227.0 (4.3)	0.087	221.9 (4.3)	0.086	222.0 (4.3)
MDLM	0.078	224.9 (4.3)	0.085	189.8 (4.2)	0.091	162.9 (4.2)	0.094	142.5 (4.2)	0.094	124.5 (4.1)
ReMDM	0.077	205.1 (4.2)	0.087	128.0 (4.2)	0.094	83.4 (4.1)	0.101	59.2 (4.0)	0.098	43.2 (3.9)
SoftMask	0.080	172.5 (4.2)	0.090	110.8 (4.2)	0.101	73.1 (4.1)	0.101	51.8 (4.0)	0.099	38.3 (4.0)
SoftMask+ λ	0.088	148.2 (4.2)	0.099	83.6 (4.1)	0.102	84.3 (4.1)	0.104	59.7 (4.0)	0.102	43.8 (3.9)
λ -ReMDM	0.078	263.4 (4.3)	0.099	157.1 (4.2)	0.116	92.3 (4.0)	0.119	62.4 (3.9)	0.116	50.0 (3.7)
λ -UDLM	0.108	90.0 (4.0)	0.112	58.9 (4.0)	0.122	42.5 (3.9)	0.118	33.5 (3.6)	0.106	23.7 (3.2)

Table 3: Ablation study of λ -DLM training strategies on LM1B (tiny DiT, NFE= 256). For the intermediate fixed- λ variants, $\lambda = 0.46$ and $\lambda = 0.68$ are the average best-MAUVE λ values across NFE settings for λ -ReMDM and λ -UDLM, respectively (see Table 6). “Standard” denotes the proposed training strategy with $\lambda \sim \text{Uniform}(0, 1)$ and two-pass training.

Model	Variant	Gen. PPL $\downarrow$	MAUVE $\uparrow$
λ -ReMDM	Standard	80.7	0.077
	Fixed $\lambda = 0$	164.2	0.069
	Fixed $\lambda = 1$	119.2	0.067
	Fixed $\lambda = 0.46$	121.8	0.059
	One-pass	252.8	0.050
λ -UDLM	Standard	72.8	0.100
	Fixed $\lambda = 0$	253.3	0.055
	Fixed $\lambda = 1$	12.3	0.061
	Fixed $\lambda = 0.68$	29.7	0.063
	One-pass	234.8	0.064

5.3 OWT Unconditional Generation

Table 4 summarizes unconditional generation results on OWT across different NFE settings and random seeds; standard deviation results are reported in Table 11. Each model generates 256 samples with a context length of 1024, and we report MAUVE, generative PPL, and sample entropy. Overall, the proposed λ -DLM variants achieve

strong performance on OWT, extending the effectiveness of retention control beyond LM1B. In particular, λ -UDLM with the log-shaped schedule achieves strong MAUVE across NFE settings and increasingly favorable generative PPL at larger NFE, while the constant- λ variant tends to favor stronger exploitation with lower sample entropy. These results suggest that the preferred retention schedule can depend on the denoising budget.

Interestingly, λ -ReMDM behaves substantially better on OWT than on LM1B. A similar tendency is observed within LM1B, where λ -ReMDM becomes more competitive when moving from tiny DiT to the larger small DiT model. Together, these observations suggest that λ -ReMDM becomes more stable with increased data diversity and model capacity. We conjecture that its instability in lower-data and lower-capacity regimes is related to a directional degeneracy tendency induced by the fixed terminal embedding of absorbing-state diffusion dynamics (see Appendix B.3); this effect appears to be mitigated as the dataset scale or model capacity increases.

Table 4: Unconditional generation results of small DiT-based models on OWT. We report mean MAUVE($\uparrow$) and generative PPL($\downarrow$) over three runs. Values in parentheses denote sample entropy.

Model	NFE = 64		NFE = 128		NFE = 256		NFE = 512		NFE = 1024	
	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL
MDLM	0.244	72.6 (5.5)	0.211	56.6 (5.5)	0.374	45.3 (5.4)	0.577	38.0 (5.3)	0.242	30.7 (5.2)
ReMDM	0.125	70.9 (5.6)	0.300	43.5 (5.5)	0.568	29.3 (5.3)	0.664	21.1 (5.2)	0.625	15.4 (5.1)
SoftMask	0.132	63.1 (5.5)	0.398	38.6 (5.4)	0.753	25.8 (5.3)	0.712	18.4 (5.2)	0.635	13.8 (5.1)
λ -ReMDM (λ_{const})	0.270	48.5 (5.4)	0.605	29.8 (5.2)	0.724	22.4 (5.2)	0.866	20.0 (5.1)	0.782	17.2 (5.1)
λ -UDLM (λ_{const})	0.258	39.6 (5.0)	0.438	24.6 (5.0)	0.575	16.1 (4.8)	0.598	19.9 (5.0)	0.691	15.7 (4.9)
λ -UDLM (λ_{log})	0.424	82.8 (5.7)	0.678	34.7 (5.5)	0.729	21.1 (5.3)	0.792	13.9 (5.1)	0.803	9.9 (4.9)

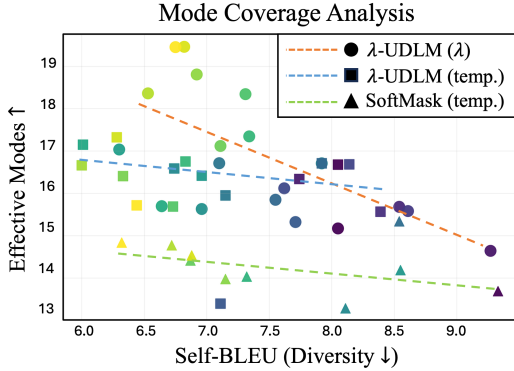


Figure 5: Mode coverage analysis of λ -UDLM and SoftMask under λ and temperature sweeps. Brighter points indicate settings that generally yield higher diversity, and dashed lines denote linear fits.

5.4 Faithful Diversity Control via λ

To examine whether λ controls semantically meaningful diversity beyond superficial lexical variation, we conduct a mode coverage analysis on generated samples. We use sentence-transformers/all-mpnet-base-v2 to encode reference sentences from the LM1B validation set and cluster them with K-means to define semantic modes. Generated samples are embedded with the same encoder and assigned to their nearest reference clusters. Based on the resulting cluster frequencies, we compute mode entropy and the effective number of modes via exponentiated entropy. In this analysis, we define $K = 25$ modes using 10,000 reference sentences and evaluate 256 generated samples at NFE= 256.

We compare three diversity-control sweeps: λ -UDLM with λ sweeping, λ -UDLM with temperature sweeping at fixed $\lambda = 0.8$, and SoftMask with temperature sweeping using the checkpoint presented in Table 1. The temperature ranges are chosen to cover self-BLEU (Zhu et al., 2018) levels comparable to those induced by λ variation, enabling comparison under similar lexical diver-

sity.

Figure 5 presents the mode coverage results. Among the three sweeps, varying λ in λ -UDLM shows the clearest relationship between lexical diversity and semantic mode coverage: lower self-BLEU generally corresponds to a larger effective number of modes. Temperature sweeping also changes mode coverage, but the relationship is less consistent for both λ -UDLM and SoftMask. These results suggest that λ provides a more structured control axis for semantic exploration than temperature-based modulation.

6 Conclusion

In this paper, we propose a novel diffusion dynamic that smoothly interpolates between continuous and discrete diffusion dynamics through a single parameter, λ . While previous methods focus on a fixed noising and denoising trajectory, our proposed λ -DLM generalizes across a broad range of diffusion trajectories. From the perspective of interpreting continuous diffusion dynamics as a retention mechanism for previous denoising results, we hypothesize that λ -DLM enables controllable exploitation-exploration trade-offs during sampling.

Across LM1B and OWT unconditional generation benchmarks, λ -DLM achieves strong performance against DLM baselines, with λ -UDLM showing the most consistent improvements. We further find that λ -sweeping produces a more structured relationship between lexical diversity and sentence-embedding-based mode coverage than temperature sweeping. These findings support λ as a useful inference-time mechanism for modulating retention and exploration without retraining.

Limitations

Most of our experiments use a constant sequence-level λ shared across denoising steps and token positions, while we additionally evaluate a log-shaped

timestep schedule for OWT λ -UDLM. Although the LLaDA-based confidence analysis and the log-shaped schedule suggest that time-varying retention can be beneficial, we do not learn λ schedules or consider token-wise adaptive control. Learning timestep-dependent or token-adaptive retention policies therefore remains an important direction for future work.

Our current inference-time λ selection also relies on a manual per-setting grid search. Across the evaluated sweeps, generative PPL, sample entropy, and diversity vary relatively systematically with λ , whereas MAUVE fluctuates more across nearby λ values and does not exhibit a comparably smooth trend. Consequently, the MAUVE-optimal setting can vary across benchmarks and NFE settings. In addition, the effectiveness of λ -DLM is architecture-sensitive: λ -UDLM exhibits relatively consistent improvements, whereas λ -ReMDM is less stable, particularly in lower-data and lower-capacity regimes. Developing more robust or adaptive selection mechanisms remains an important direction for future work.

Finally, our empirical evaluation is restricted to unconditional language generation on LM1B and OWT. We do not evaluate conditional generation or downstream language-generation tasks. Accordingly, our empirical conclusions should be interpreted within the scope of unconditional generation, and the applicability of λ -diffusion dynamics to broader settings remains to be established.

Acknowledgments

This work was supported by the InnoCORE program of the Ministry of Science and ICT (UNIST 1.260006.01), and Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. IITP-2026-RS-2024-00360227, Leading Generative AI Human Resources Development, No. RS-2025-25442824, AI Star Fellowship Program (Ulsan National Institute of Science and Technology), & No. RS-2020-II201336, Artificial Intelligence graduate school support (UNIST)).

We gratefully acknowledge Professor Heeyoul Choi at Handong Global University for providing the GPU computing resources used in our experiments.

References

- Jacob Austin, Daniel D Johnson, Jonathan Ho, Daniel Tarlow, and Rianne Van Den Berg. 2021. Structured denoising diffusion models in discrete state-spaces. *Advances in neural information processing systems*, 34:17981–17993.
- Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, and 1 others. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.
- Ciprian Chelba, Tomas Mikolov, Mike Schuster, Qi Ge, Thorsten Brants, Phillipp Koehn, and Tony Robinson. 2013. One billion word benchmark for measuring progress in statistical language modeling. *arXiv preprint arXiv:1312.3005*.
- Sander Dieleman, Laurent Sartran, Arman Roshanai, Nikolay Savinov, Yaroslav Ganin, Pierre H Richemond, Arnaud Doucet, Robin Strudel, Chris Dyer, Conor Durkan, and 1 others. 2022. Continuous diffusion for categorical data. *arXiv preprint arXiv:2211.15089*.
- Stefan Elfving, Eiji Uchibe, and Kenji Doya. 2018. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural networks*, 107:3–11.
- Zhujin Gao, Junliang Guo, Xu Tan, Yongxin Zhu, Fang Zhang, Jiang Bian, and Linli Xu. 2024. Empowering diffusion models on the embedding space for text generation. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 4664–4683.
- Aaron Gokaslan and Vanya Cohen. 2019. Open-webtext corpus. <http://Skylion007.github.io/OpenWebTextCorpus>.
- Shansan Gong, Mukai Li, Jiangtao Feng, Zhiyong Wu, and LingPeng Kong. 2022. Diffuseq: Sequence to sequence text generation with diffusion models. *arXiv preprint arXiv:2210.08933*.
- Zhengfu He, Tianxiang Sun, Qiong Tang, Kuanning Wang, Xuan-Jing Huang, and Xipeng Qiu. 2023. Diffusionbert: Improving generative masked language models with diffusion models. In *Proceedings of the 61st annual meeting of the association for computational linguistics (volume 1: Long papers)*, pages 4521–4534.
- Michael Hersche, Samuel Moor-Smith, Thomas Hofmann, and Abbas Rahimi. 2025. Soft-masked diffusion language models. *arXiv preprint arXiv:2510.17206*.
- Jonathan Ho, Ajay Jain, and Pieter Abbeel. 2020. Denoising diffusion probabilistic models. *Advances*

- in neural information processing systems, 33:6840–6851.
- Mingyu Jo, Jaesik Yoon, Justin Deschenaux, Caglar Gulcehre, and Sungjin Ahn. 2025. Loopholing discrete diffusion: Deterministic bypass of the sampling wall. *arXiv preprint arXiv:2510.19304*.
- Xiang Li, John Thickstun, Ishaan Gulrajani, Percy S Liang, and Tatsunori B Hashimoto. 2022. Diffusion-lm improves controllable text generation. *Advances in neural information processing systems*, 35:4328–4343.
- Shen Nie, Fengqi Zhu, Zebin You, Xiaolu Zhang, Jingyang Ou, Jun Hu, Jun Zhou, Yankai Lin, Ji-Rong Wen, and Chongxuan Li. 2026. Large language diffusion models. *Advances in Neural Information Processing Systems*, 38:50608–50646.
- William Peebles and Saining Xie. 2023. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4195–4205.
- Krishna Pillutla, Swabha Swayamdipta, Rowan Zellers, John Thickstun, Sean Welleck, Yejin Choi, and Zaid Harchaoui. 2021. Mauve: Measuring the gap between neural text and human text using divergence frontiers. *Advances in Neural Information Processing Systems*, 34:4816–4828.
- Patrick Pynadath, Jiaxin Shi, and Ruqi Zhang. 2025. Candi: Hybrid discrete-continuous diffusion models. *arXiv preprint arXiv:2510.22510*.
- Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, and 1 others. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.
- Subham S Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Marroquin, Justin T Chiu, Alexander Rush, and Volodymyr Kuleshov. 2024. Simple and effective masked diffusion language models. *Advances in Neural Information Processing Systems*, 37:130136–130184.
- Subham Sekhar Sahoo, Justin Deschenaux, Aaron Gokaslan, Guanghan Wang, Justin Chiu, and Volodymyr Kuleshov. 2025. The diffusion duality. *arXiv preprint arXiv:2506.10892*.
- Jiaxin Shi, Kehang Han, Zhe Wang, Arnaud Doucet, and Michalis Titsias. 2024. Simplified and generalized masked diffusion for discrete data. *Advances in neural information processing systems*, 37:103131–103167.
- Robin Strudel, Corentin Tallec, Florent Alth  , Yilun Du, Yaroslav Ganin, Arthur Mensch, Will Grathwohl, Nikolay Savinov, Sander Dieleman, Laurent Sifre, and 1 others. 2022. Self-conditioned embedding diffusion for text generation. *arXiv preprint arXiv:2211.04236*.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.
- Guanghan Wang, Yair Schiff, Subham Sahoo, and Volodymyr Kuleshov. 2026. Remasking discrete diffusion models with inference-time scaling. *Advances in Neural Information Processing Systems*, 38:147282–147339.
- Chi Zhang, Xixi Hu, Bo Liu, and Qiang Liu. 2026. Gumbel distillation for parallel text generation. *arXiv preprint arXiv:2603.22216*.
- Huangjie Zheng, Shansan Gong, Ruixiang Zhang, Tianrong Chen, Jiatao Gu, Mingyuan Zhou, Navdeep Jaitly, and Yizhe Zhang. 2025. Continuously augmented discrete diffusion model for categorical generative modeling. *arXiv preprint arXiv:2510.01329*.
- Yaoming Zhu, Sidi Lu, Lei Zheng, Jiaxian Guo, Weinan Zhang, Jun Wang, and Yong Yu. 2018. Texygen: A benchmarking platform for text generation models. In *The 41st international ACM SIGIR conference on research & development in information retrieval*, pages 1097–1100.

A Mathematical Details

A.1 Details of Continuous Relaxation Approximation

As described in Section 3.1, we define the continuous-relaxation distribution as

$$\begin{aligned} q_{\text{relax}}(z|x; E) &= \mathcal{N}_{Ex}^{\sigma}(z) \\ &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\|z-Ex\|^2}{2\sigma^2}}. \end{aligned}$$

Based on this formulation, the likelihood distribution over discrete tokens can be derived using Bayes’ rule:

$$\begin{aligned} p(x|z) &= \frac{p(z|x)p(x)}{\sum_{j \in V} p(z|x=j)p(x=j)} \\ &= \frac{p(x)e^{-\frac{1}{2\sigma^2}\|z-Ex\|^2}}{\sum_{j \in V} p(x=j)e^{-\frac{1}{2\sigma^2}\|z-Ej\|^2}}, \\ &\approx \frac{e^{Ex^{\top}z/\sigma^2}}{\sum_{j \in V} e^{Ej^{\top}z/\sigma^2}}, \\ &= \text{softmax}(E^{\top}z/\sigma^2), \end{aligned}$$

where the approximation in the third line assumes unit-norm embeddings, i.e., $|E_i| = 1$ for all $i \in V$, and a uniform data prior $p(x)$. Notably, the resulting likelihood distribution corresponds exactly

to a softmax distribution over embedding similarities $E^\top z$ with temperature σ^2 . This establishes a structural connection between continuous-relaxation likelihood estimation and the standard logit-softmax formulation commonly used in neural language models.

A.2 Characterized Dynamics in Reverse Posterior Perspectives

In continuous DLMs, following DDPM (Ho et al., 2020), the exact reverse posterior is generally defined as

$$\begin{aligned} q^c(z_{t-1}|z_t, x_0) &= \mathcal{N}(z_{t-1}; \tilde{\mu}_t(z_t, z_0), \tilde{\sigma}_t^2 I), \\ \tilde{\mu}_t(z_t, z_0) &= \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1 - \bar{\alpha}_t} z_0 + \frac{\bar{\alpha}_t(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} z_t, \\ \tilde{\sigma}_t^2 &= \frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t} \beta_t, \end{aligned}$$

where $\beta_t = 1 - \alpha_t$. Because the sum of the two coefficients multiplied by z_0 and z_t is not strictly equal to 1, the reverse posterior dynamics do not exactly follow the interpretation of *transporting a unimodal Gaussian* between the input and terminal embeddings, which would require a strict interpolation whose coefficients sum to 1. However, when the number of denoising steps is sufficiently large, the coefficient sum approaches 1 according to the following derivation:

$$\begin{aligned} S_t &= \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t + \sqrt{\bar{\alpha}_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} \\ &= \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t + \sqrt{\bar{\alpha}_t}(1 - \bar{\alpha}_{t-1})}{(1 - \bar{\alpha}_{t-1}) + \beta_t\bar{\alpha}_{t-1}} \\ &\approx \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t + (1 - \bar{\alpha}_{t-1}) - \frac{\beta_t}{2}(1 - \bar{\alpha}_{t-1})}{(1 - \bar{\alpha}_{t-1}) + \beta_t\bar{\alpha}_{t-1}} \\ &\approx \frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_{t-1}} = 1, \end{aligned}$$

where the first approximation is based on the Taylor expansion

$$\sqrt{\alpha_t} = \sqrt{1 - \beta_t} \approx 1 - \frac{\beta_t}{2} + O(\beta_t^2),$$

and the second approximation assumes $\beta_t \rightarrow 0$, which typically holds under large denoising-step settings.

Importantly, during sampling, the exact reverse posterior distribution is approximated using the model’s previous denoising output $\hat{x}_0$ because the ground-truth x_0 is unavailable. Consequently, the

transportation dynamic can be interpreted as evolving between the previously denoised token embedding and the terminal embedding.

The reverse posterior of discrete DLMs is generally formulated as follows:

$$q^d(x_s|x_t, x_0) = \begin{cases} 1, & x_t = x_0, x_s = x_0 \\ \frac{\bar{\alpha}_s - \bar{\alpha}_t}{1 - \bar{\alpha}_t}, & x_t \sim p(\pi), x_s = x_0 \\ \frac{1 - \bar{\alpha}_s}{1 - \bar{\alpha}_t} p(\pi), & x_t \sim p(\pi) \\ 0 & \text{otherwise.} \end{cases}$$

where $p(\pi)$ corresponds to either $\delta_{[\text{MASK}]}$ (Dirac-delta) or the uniform distribution over $\mathcal{V}$ for MDLM and UDLM architectures, respectively.

Based on the continuous relaxation scheme, each case of the reverse posterior can be derived in the continuous embedding space as follows:

$$\begin{aligned} q^d(z_s|x_t, x_0) &= \sum_{x_s} q_{\text{relax}}(z_s|x_s) q^d(x_s|x_t, x_0) \\ &= \begin{cases} \mathcal{N}_{E x_0}^\sigma, & x_t = x_0 \\ \frac{\mathcal{N}_{E \pi}^\sigma + \bar{\alpha}_s(\mathcal{N}_{E x_0}^\sigma - \mathcal{N}_{E \pi}^\sigma) - \bar{\alpha}_t \mathcal{N}_{E x_0}^\sigma}{1 - \bar{\alpha}_t}, & x_t \sim p(\pi) \\ 0 & \text{otherwise,} \end{cases} \end{aligned}$$

where the second case, $q^d(z_s|x_t = \pi, x_0)$, can be further expressed as

$$q^d(z_s|x_t = \pi, x_0) = \begin{cases} \bar{\alpha}_s \mathcal{N}_{E x_0}^\sigma + (1 - \bar{\alpha}_s) \mathcal{N}_{E \pi}^\sigma, & t = 1, s < t \\ \frac{\bar{\alpha}_s - \bar{\alpha}_t}{1 - \bar{\alpha}_t} \mathcal{N}_{E x_0}^\sigma + \frac{1 - \bar{\alpha}_s}{1 - \bar{\alpha}_t} \mathcal{N}_{E \pi}^\sigma & 0 < s < t < 1. \end{cases}$$

In summary, the pushforward reverse posterior distribution of discrete diffusion dynamics can also be characterized as an *anchored multimodal MoG*. Analogous to continuous DLMs, discrete DLMs typically parameterize their backbone models to directly predict x_0 . Therefore, during sampling, the exact reverse posterior is approximated using the model prediction $\hat{x}_0$, replacing $\mathcal{N}_{E x_0}^\sigma$ with $\mathcal{N}_{E \hat{x}_0}^\sigma$.

A.3 Noising Variance of λ -Diffusion Dynamic

Regarding $z^\lambda = \lambda z^c + (1 - \lambda)z^d$, the variance of z^λ could be simplified as follows:

$$\begin{aligned}
\sigma_\lambda^2(t) &= \text{Var}[z_t^\lambda | x_t, x_0], \\
&= \text{Var}[\lambda z_t^c + (1 - \lambda)z_t^d | x_t, x_0], \\
&= \text{Var}[\lambda(\mu^c(x_0, x_t, t) + \sigma_t^c \epsilon^c) \\
&\quad + (1 - \lambda)(\mu^d(x_t) + \sigma_t^d \epsilon^d)], \\
&= \text{Var}[\underbrace{\lambda\mu^c(x_0, x_t, t) + (1 - \lambda)\mu^d(x_t)}_{\text{mean}} + \lambda\sigma_t^c \epsilon^c \\
&\quad + (1 - \lambda)\sigma_t^d \epsilon^d], \\
&= \text{Var}[\lambda\sigma_t^c \epsilon^c + (1 - \lambda)\sigma_t^d \epsilon^d], \\
&= \lambda^2(\sigma_t^c)^2 + (1 - \lambda)^2(\sigma_t^d)^2,
\end{aligned}$$

where ϵ^c and ϵ^d are independently sampled from $\mathcal{N}(\epsilon; 0, I)$.

A.4 Reverse Posterior of λ -Diffusion Dynamic

The detailed λ -diffusion dynamic's reverse posterior derivation is done as follows:

$$\begin{aligned}
q^\lambda(z_s^\lambda | z_t^\lambda, x_0) &= \frac{q(z_s^\lambda, z_t^\lambda | x_0)}{q(z_t^\lambda | x_0)} \\
&= \sum_{x_s, x_t} \frac{q(z_s^\lambda, z_t^\lambda, x_s, x_t | x_0)}{q(z_t^\lambda | x_0)} \\
&= \sum_{x_s, x_t} \frac{q(z_s^\lambda | \cancel{x_t}, x_s, \cancel{x_t}, x_0) q(z_t^\lambda | \cancel{x_s}, x_t, x_0)}{q(z_t^\lambda | x_0)} \\
&\quad \frac{q(x_s | x_t, x_0) q(x_t | x_0)}{q(z_t^\lambda | x_0)} \\
&= \sum_{x_s, x_t} q(z_s^\lambda | x_s, x_0) q(x_t | z_t^\lambda, x_0) q(x_s | x_t, x_0) \\
&= \sum_{x_t} q(x_t | z_t^\lambda, x_0) \sum_{x_s} q^d(x_s | x_t, x_0) q^\lambda(z_s^\lambda | x_s, x_0),
\end{aligned}$$

where the fourth equation is derived by $q(z_t^\lambda | x_t, x_0) = \frac{q(x_t | z_t^\lambda, x_0) q(z_t^\lambda | x_0)}{q(x_t | x_0)}$. The cancellations at third equation is based on conditional independence relationships.

While $q(x_s | x_t, x_0)$ and $q(z_s^\lambda | x_s, x_0)$ are defined by the reverse posterior of discrete DLM and continuous relaxation scheme, respectively, $q(x_t | z_t^\lambda, x_0)$ should be newly defined. In MDLM and UDLM architectures, because x_t has two possible states, x_0 (data state) or π (a terminal state), $q(x_t | z_t^\lambda, x_0)$ can be derived as follows:

$$\begin{aligned}
q(x_t | z_t^\lambda, x_0) &\propto q(x_t | x_0) \mathcal{N}(z_t^\lambda; \mu_\lambda(x_0, x_t, t), \sigma^2 I), \\
&= \begin{cases} \bar{\alpha}_t \mathcal{N}_{\mu_\lambda(x_0, x_0, t)}^\sigma, & x_t = x_0, \\ (1 - \bar{\alpha}_t) \mathcal{N}_{\mu_\lambda(x_0, \pi, t)}^\sigma, & x_t = \pi \sim p(\pi). \end{cases}
\end{aligned}$$

B Experimental Setting Details

B.1 LLaDA-based Confidence Analysis Details

We use a frozen GSAI-ML/LLaDA-8B-Base model to obtain an external bidirectional confidence signal. At each ancestral denoising step, λ -DLM first samples a clean-token proposal $\hat{x}_0$. Because the generator and LLaDA use different tokenizers, the proposal is decoded with the generator tokenizer and re-tokenized with the LLaDA tokenizer. LLaDA runs in evaluation mode, and a single bidirectional forward pass produces position-wise logits. For each non-padding token, we compute normalized-entropy confidence

$$\begin{aligned}
c_i &= 1 - \frac{H(p_i)}{\log |\mathcal{V}_{\text{LLaDA}}|}, \\
H(p_i) &= - \sum_v p_i(v) \log p_i(v),
\end{aligned}$$

and average c_i across positions to obtain a sequence-level confidence score. For LLaDA-guided retention results (Figure 3), the resulting sequence-level confidence is recomputed at every denoising step and broadcast across token positions as the retention strength for that step.

B.2 Main Unconditional Generation Details

In this section, we provide additional details of the model architectures, training configurations, and inference-time λ selection protocol. Table 5 summarizes the configurations used for both LM1B and OWT experiments. For LM1B, we evaluate tiny- and small-DiT backbones with 24M and 139M parameters and a context length of 128, whereas the OWT experiments use a 169M small-DiT model with a context length of 1024. The larger parameter count of the OWT model mainly results from the larger GPT-2 vocabulary while retaining the same 12-layer, 768-dimensional Small-DiT backbone. The additional λ embedding consists of only a lightweight two-layer MLP and therefore introduces negligible parameter overhead.

For evaluating λ -DLM, we select the inference-time λ schedule separately for each model, benchmark, and NFE setting, following the protocol described in Section 5.1.3. For the constant schedule, we search $\lambda_0 \in [0, 1]$ with a step size of 0.1, while for the log-shaped schedule we search the log-rate $r \in [1, 10]$ with a step size of 1. Schedule selection is based primarily on mean MAUVE over

Table 5: Model architectures and principal training configurations of λ -DLM. The Small DiT architecture is shared between LM1B and OWT, while the different vocabulary sizes result in different total parameter counts.

Configuration	Tiny DiT (LM1B)	Small DiT (LM1B)	Small DiT (OWT)
Transformer layers	8	12	12
Hidden dimension	256	768	768
Attention heads	8	12	12
MLP expansion ratio	4	4	4
Total parameters	24M	139M	169M
Context length	128	128	1024
Vocabulary	BERT-base-uncased	BERT-base-uncased	GPT-2
Time embedding	Sinusoidal + MLP	Sinusoidal + MLP	Sinusoidal + MLP
λ embedding	2-layer MLP + SiLU	2-layer MLP + SiLU	2-layer MLP + SiLU
Backbone architecture	DiT	DiT	DiT
Training Setup			
Dataset	LM1B	LM1B	OWT
Training iterations	200K	200K	100K
Global batch size	512	512	512
Optimizer	AdamW	AdamW	AdamW
Learning rate	3×10^{-4}	3×10^{-4}	3×10^{-4}
Weight decay	0	0	0
Gradient clipping	1.0	1.0	1.0
Learning-rate schedule	Constant + warmup	Constant + warmup	Constant + warmup
LR warmup iterations	2.5K	2.5K	2.5K
Continuous noise	VP-style	VP-style	VP-style
Two-pass warmup iterations	10K	10K	10K
Second-pass skip schedule	Linear annealing	Linear annealing	Linear annealing

Table 6: Selected λ -schedule parameters for each model, benchmark, and NFE setting. λ_0 denotes the constant- λ parameter in Eq. 3, while r denotes the log-rate parameter in Eq. 4.

Setting	Model	Schedule parameter	NFE=64	NFE=128	NFE=256	NFE=512	NFE=1024
Tiny DiT – LM1B	SoftMask+ λ	Constant λ_0 (Eq. 3)	0.6	0.4	0.6	0.1	0.1
	λ -ReMDM	Constant λ_0 (Eq. 3)	0.5	0.9	0.4	0.3	0.2
	λ -UDLM	Constant λ_0 (Eq. 3)	0.8	0.7	0.7	0.6	0.6
Small DiT – LM1B	SoftMask+ λ	Constant λ_0 (Eq. 3)	0.6	0.7	0.4	0.3	0.3
	λ -ReMDM	Constant λ_0 (Eq. 3)	0.3	0.2	0.3	0.3	0.1
	λ -UDLM	Constant λ_0 (Eq. 3)	0.9	0.9	0.9	0.7	0.7
Small DiT – OWT	λ -ReMDM	Constant λ_0 (Eq. 3)	0.2	0.2	0.1	0.2	0.1
	λ -UDLM	Constant λ_0 (Eq. 3)	0.6	0.6	0.5	0.3	0.2
	λ -UDLM	Log-rate r (Eq. 4)	2	5	6	8	7

held-out tuning seeds, and the selected configuration is subsequently evaluated using separate test seeds to avoid test leakage. Specifically, we use seeds 3–5 for tuning and seeds 0–2 for final evaluation. Table 6 summarizes the selected schedule parameters across all evaluated settings, while the underlying constant- λ sweeps are reported in Tables 7 and 8. The constant schedule is used as our default, whereas the log-shaped schedule is additionally evaluated for OWT λ -UDLM, for which it provides meaningful improvements.

For training, the LM1B models are optimized for 200K updates, while the reported OWT models use 100K-update checkpoints with a global batch size of 512. For final multi-seed evaluation, LM1B uses 1024 generated samples per seed with context length 128, whereas the reported OWT comparison

uses 256 samples per seed with context length 1024. Additional evaluation-specific settings, including the MAUVE configuration, are kept fixed across methods within each benchmark.

For the LM1B experiments, we utilized a single RTX 3090 for Tiny DiT and a single A100 GPU for Small DiT. On average, training the tiny- and small-DiT LM1B models required approximately 3 and 2 days, respectively. For OWT, we used $4 \times$ A100 GPUs, and training the 169M Small DiT models for 100K updates required approximately 5 days.

B.3 Discussion of λ -ReMDM Instability

In this section, we discuss our interpretation of why λ -ReMDM (and similarly λ -MDLM) generally underperforms compared to λ -UDLM. We hypothesize that the fixed terminal embedding in MDLM

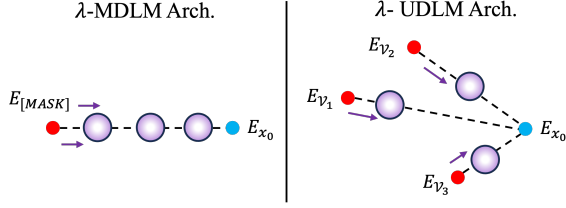


Figure 6: Comparison of λ -MDLM and λ -UDLM dynamics in the embedding space. Red dots indicate terminal embeddings at each training iteration. Blue dot indicates the ground-truth embedding. Larger purple circles indicate z_t^λ latent variable produced by different t and λ .

architectures, namely the [MASK] token embedding, induces a directionally degenerate training solution.

Unlike UDLMs, where the terminal embedding varies across training iterations due to randomly sampled terminal tokens, MDLMs always utilize the same [MASK] embedding as the terminal state. As a result, information about the terminal state can become overly encoded within the backbone model itself. Consequently, regardless of the timestep t or the selected λ value, the generated latent variable z_t^λ always lies on the line segment between the [MASK] embedding and the ground-truth embedding as illustrated in Figure 6.

This geometric structure may implicitly encourage the model to regard the direction

$$z_t^\lambda - E_{[\text{MASK}]} \quad (5)$$

as a reliable signal for recovering the ground-truth embedding. We hypothesize that the model gradually overfits to this shortcut-like directional cue during training. However, during sampling, the latent variables are recursively constructed from the model’s own denoising outputs rather than the ground-truth embedding, making this directional signal unreliable. Consequently, the learned denoising behavior can become unstable and less generalizable during inference.

C Additional Experimental Information

Standard deviation report. To assess the robustness of the main unconditional generation results, we additionally report mean and standard deviation over three runs for both LM1B and OWT. Tables 9 and 10 provide the corresponding statistics for the LM1B results in Section 5.2, while Table 11 reports those for the OWT experiments. Each entry reports the mean and standard deviation of MAUVE and

generative PPL across three evaluation seeds. Overall, the results confirm that the main performance trends are stable across random seeds, while the magnitude of the gains varies across model architectures, datasets, and NFE settings.

λ -sweeping results. We further report the complete λ -sweeping results used for inference-time schedule selection. Tables 7 and 8 present the constant- λ sweeps for Tiny- and Small-DiT models on LM1B, respectively. These sweeps illustrate how varying retention strength changes MAUVE and generative PPL across different denoising budgets, and form the basis for selecting the schedule parameters summarized in Table 6.

Table 7: Full constant- λ coarse sweep on LM1B with tiny DiT models. Each entry reports MAUVE and generative PPL, with sample entropy shown in parentheses, for the corresponding λ and NFE setting.

Model	λ	NFE = 64		NFE = 128		NFE = 256		NFE = 512		NFE = 1024	
		MAUVE	PPL	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL
SoftMask+ λ	0.0	0.082	266.19 (4.23)	0.077	171.71 (4.14)	0.096	111.51 (4.06)	0.087	79.93 (3.95)	0.096	55.25 (3.79)
	0.1	0.069	265.51 (4.23)	0.087	171.06 (4.15)	0.087	109.39 (4.04)	0.103	78.02 (3.95)	0.110	54.91 (3.77)
	0.2	0.077	276.66 (4.23)	0.089	179.86 (4.16)	0.089	112.96 (4.05)	0.098	79.05 (3.92)	0.096	54.31 (3.76)
	0.3	0.079	283.13 (4.23)	0.086	184.53 (4.15)	0.090	113.01 (4.05)	0.092	78.86 (3.90)	0.088	52.13 (3.73)
	0.4	0.069	263.20 (4.21)	0.093	165.70 (4.13)	0.082	103.58 (4.03)	0.092	72.33 (3.88)	0.084	49.57 (3.71)
	0.5	0.081	216.21 (4.18)	0.076	137.95 (4.09)	0.093	88.07 (3.99)	0.090	61.56 (3.85)	0.090	42.73 (3.64)
	0.6	0.084	168.62 (4.10)	0.084	107.41 (4.00)	0.097	71.48 (3.87)	0.097	48.10 (3.69)	0.087	34.66 (3.49)
	0.7	0.069	142.17 (3.64)	0.084	81.41 (3.50)	0.083	52.62 (3.34)	0.080	35.22 (3.02)	0.078	23.80 (2.67)
	0.8	0.066	139.46 (2.20)	0.066	66.02 (1.96)	0.063	35.53 (1.75)	0.065	17.26 (1.38)	0.062	7.77 (0.96)
	0.9	0.055	127.65 (1.25)	0.057	58.30 (1.16)	0.060	23.17 (0.94)	0.057	9.38 (0.74)	0.057	5.93 (0.60)
	1.0	0.054	111.12 (1.48)	0.057	56.39 (1.18)	0.057	22.53 (0.83)	0.056	10.64 (0.73)	0.057	6.87 (0.58)
λ -ReMDM	0.0	0.061	418.36 (4.35)	0.068	392.33 (4.34)	0.068	358.05 (4.34)	0.065	353.61 (4.34)	0.062	316.33 (4.29)
	0.1	0.075	254.64 (3.95)	0.079	167.48 (3.79)	0.085	111.68 (3.66)	0.096	90.16 (3.30)	0.089	73.30 (2.64)
	0.2	0.077	238.44 (3.92)	0.075	154.96 (3.77)	0.090	103.58 (3.60)	0.090	80.42 (3.28)	0.093	65.90 (2.58)
	0.3	0.074	225.81 (3.91)	0.080	141.00 (3.75)	0.084	97.64 (3.59)	0.100	75.83 (3.23)	0.085	60.38 (2.53)
	0.4	0.079	214.35 (3.91)	0.077	134.31 (3.76)	0.096	89.94 (3.57)	0.094	68.14 (3.27)	0.078	57.03 (2.53)
	0.5	0.080	205.62 (3.91)	0.089	130.83 (3.75)	0.095	87.69 (3.58)	0.083	65.27 (3.23)	0.084	55.98 (2.52)
	0.6	0.074	197.94 (3.91)	0.081	122.33 (3.75)	0.075	83.38 (3.55)	0.091	64.53 (3.21)	0.085	54.64 (2.49)
	0.7	0.070	189.90 (3.91)	0.083	117.36 (3.74)	0.084	80.94 (3.54)	0.089	60.52 (3.20)	0.079	53.35 (2.45)
	0.8	0.069	182.97 (3.90)	0.089	110.58 (3.75)	0.090	75.05 (3.54)	0.083	58.99 (3.19)	0.075	51.71 (2.42)
	0.9	0.074	177.40 (3.90)	0.089	108.25 (3.73)	0.081	74.01 (3.53)	0.085	56.65 (3.18)	0.086	51.20 (2.48)
	1.0	0.069	172.25 (3.90)	0.083	103.83 (3.71)	0.080	69.06 (3.52)	0.084	54.47 (3.12)	0.082	46.39 (2.41)
λ -UDLM	0.0	0.056	235.24 (3.28)	0.061	222.44 (3.34)	0.060	216.70 (3.38)	0.060	213.48 (3.42)	0.062	194.27 (3.42)
	0.1	0.063	256.81 (3.45)	0.063	207.94 (3.51)	0.065	169.69 (3.52)	0.064	146.36 (3.55)	0.072	118.55 (3.51)
	0.2	0.067	269.86 (3.61)	0.071	203.84 (3.65)	0.072	163.08 (3.68)	0.078	130.71 (3.67)	0.077	92.69 (3.59)
	0.3	0.076	267.68 (3.76)	0.084	206.86 (3.78)	0.078	147.88 (3.80)	0.093	108.78 (3.72)	0.081	75.69 (3.54)
	0.4	0.073	234.71 (3.87)	0.087	195.56 (3.85)	0.088	118.79 (3.79)	0.102	81.07 (3.54)	0.093	43.53 (2.59)
	0.5	0.088	181.47 (3.91)	0.101	154.67 (3.90)	0.105	94.92 (3.77)	0.118	57.75 (3.31)	0.089	27.90 (2.15)
	0.6	0.124	133.51 (3.94)	0.124	103.77 (3.89)	0.117	57.13 (3.65)	0.125	29.14 (3.01)	0.094	15.41 (2.02)
	0.7	0.119	104.84 (3.95)	0.124	81.97 (3.92)	0.122	45.33 (3.59)	0.122	26.55 (3.15)	0.092	12.16 (2.12)
	0.8	0.132	91.79 (3.95)	0.101	71.08 (3.91)	0.113	38.91 (3.53)	0.109	24.07 (3.12)	0.089	11.26 (2.26)
	0.9	0.119	84.17 (3.91)	0.111	63.56 (3.85)	0.102	39.63 (3.54)	0.089	23.02 (3.04)	0.079	8.65 (2.07)
	1.0	0.108	86.39 (3.92)	0.094	64.06 (3.83)	0.092	41.10 (3.46)	0.084	23.39 (2.98)	0.070	12.49 (2.37)

Table 8: Full constant- λ coarse sweep on LM1B with small DiT models. Each entry reports MAUVE and generative PPL, with sample entropy shown in parentheses, for the corresponding λ and NFE setting.

Model	λ	NFE = 64		NFE = 128		NFE = 256		NFE = 512		NFE = 1024	
		MAUVE	PPL	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL	MAUVE	PPL
SoftMask+ λ	0.0	0.090	198.77 (4.25)	0.101	127.56 (4.18)	0.102	86.24 (4.12)	0.108	61.19 (4.05)	0.111	44.54 (3.99)
	0.1	0.096	194.92 (4.24)	0.103	126.56 (4.18)	0.111	83.95 (4.11)	0.102	61.38 (4.05)	0.116	43.07 (3.97)
	0.2	0.092	196.41 (4.24)	0.101	127.64 (4.17)	0.113	84.20 (4.11)	0.112	59.71 (4.03)	0.116	43.83 (3.97)
	0.3	0.082	209.72 (4.25)	0.095	130.21 (4.17)	0.108	82.93 (4.10)	0.118	61.11 (4.03)	0.119	43.69 (3.96)
	0.4	0.079	200.01 (4.24)	0.105	124.96 (4.17)	0.119	81.17 (4.09)	0.112	59.10 (4.02)	0.109	41.75 (3.93)
	0.5	0.083	180.10 (4.24)	0.096	112.77 (4.16)	0.102	73.68 (4.09)	0.111	52.61 (4.01)	0.112	37.85 (3.93)
	0.6	0.096	148.67 (4.21)	0.098	94.19 (4.14)	0.109	61.99 (4.04)	0.108	45.50 (3.98)	0.112	33.95 (3.91)
	0.7	0.092	126.18 (4.17)	0.117	84.15 (4.10)	0.113	57.42 (4.03)	0.111	41.37 (3.97)	0.111	31.33 (3.87)
	0.8	0.088	115.55 (3.80)	0.104	73.92 (3.78)	0.111	51.14 (3.73)	0.116	36.73 (3.62)	0.111	29.29 (3.54)
	0.9	0.084	108.10 (2.79)	0.082	69.50 (2.72)	0.097	46.53 (2.55)	0.098	32.68 (2.36)	0.098	24.93 (2.22)
	1.0	0.067	106.46 (0.90)	0.077	61.69 (0.77)	0.080	25.79 (0.62)	0.078	16.49 (0.55)	0.078	9.97 (0.39)
λ -ReMDM	0.0	0.083	450.42 (4.33)	0.092	410.13 (4.34)	0.088	377.87 (4.33)	0.091	353.76 (4.34)	0.102	311.84 (4.32)
	0.1	0.084	280.97 (4.30)	0.109	176.10 (4.20)	0.115	109.30 (4.06)	0.137	69.26 (3.88)	0.164	51.00 (3.64)
	0.2	0.079	266.19 (4.29)	0.109	165.34 (4.19)	0.134	101.90 (4.05)	0.143	64.02 (3.87)	0.162	48.51 (3.63)
	0.3	0.080	253.13 (4.28)	0.098	154.97 (4.16)	0.118	95.21 (4.01)	0.147	61.77 (3.83)	0.130	46.18 (3.56)
	0.4	0.082	236.33 (4.25)	0.100	143.75 (4.13)	0.134	86.84 (3.98)	0.144	59.69 (3.77)	0.133	45.20 (3.48)
	0.5	0.077	228.28 (4.23)	0.099	135.95 (4.10)	0.133	82.60 (3.92)	0.113	58.42 (3.68)	0.134	46.17 (3.33)
	0.6	0.083	221.44 (4.21)	0.094	131.92 (4.08)	0.120	82.36 (3.88)	0.132	58.56 (3.61)	0.122	47.66 (3.22)
	0.7	0.076	215.91 (4.20)	0.086	131.93 (4.06)	0.109	81.81 (3.85)	0.118	58.94 (3.56)	0.122	47.16 (3.14)
	0.8	0.078	219.83 (4.20)	0.088	130.61 (4.06)	0.115	82.55 (3.84)	0.115	59.04 (3.53)	0.120	47.90 (3.06)
	0.9	0.080	222.61 (4.20)	0.094	130.88 (4.05)	0.115	83.05 (3.83)	0.112	60.22 (3.51)	0.112	48.74 (3.04)
	1.0	0.076	228.64 (4.20)	0.113	132.97 (4.05)	0.115	83.99 (3.83)	0.113	61.18 (3.50)	0.133	49.03 (3.02)
λ -UDLM	0.0	0.066	282.42 (3.44)	0.072	282.46 (3.50)	0.083	275.52 (3.56)	0.078	263.14 (3.59)	0.080	262.95 (3.63)
	0.1	0.085	323.00 (3.95)	0.086	225.04 (4.00)	0.118	159.25 (4.00)	0.120	113.14 (3.96)	0.108	80.03 (3.87)
	0.2	0.090	274.29 (3.94)	0.109	185.19 (4.03)	0.129	119.99 (4.04)	0.123	79.05 (4.01)	0.137	55.51 (3.90)
	0.3	0.082	253.46 (3.89)	0.108	156.48 (4.03)	0.143	101.54 (4.03)	0.120	72.51 (3.97)	0.125	60.09 (3.86)
	0.4	0.090	243.71 (3.85)	0.104	178.66 (3.91)	0.113	167.72 (3.71)	0.083	113.39 (2.97)	0.077	99.31 (2.40)
	0.5	0.086	174.10 (3.77)	0.108	118.75 (3.78)	0.111	89.56 (3.45)	0.091	67.32 (2.78)	0.087	42.31 (2.55)
	0.6	0.096	125.43 (3.82)	0.125	80.57 (3.75)	0.134	56.53 (3.64)	0.130	39.95 (3.27)	0.095	30.47 (2.91)
	0.7	0.101	99.76 (3.85)	0.113	68.35 (3.86)	0.136	47.75 (3.79)	0.123	33.77 (3.56)	0.112	23.66 (3.18)
	0.8	0.109	93.76 (3.94)	0.127	62.32 (3.92)	0.151	43.12 (3.84)	0.132	31.20 (3.63)	0.118	22.44 (3.32)
	0.9	0.098	87.62 (3.98)	0.126	60.18 (3.95)	0.141	41.90 (3.85)	0.123	31.53 (3.68)	0.113	22.72 (3.34)
	1.0	0.107	86.28 (4.02)	0.138	58.17 (3.99)	0.143	42.86 (3.85)	0.119	31.96 (3.65)	0.118	23.85 (3.35)

Table 9: Multi-seed unconditional generation results of **tiny** DiT-based models on **LM1B**. Each entry reports mean $\pm$ standard deviation over three runs. For λ -based methods, λ is selected with respect to MAUVE.

Model	NFE = 64	NFE = 128	NFE = 256	NFE = 512	NFE = 1024
MAUVE					
UDLM	0.061 $\pm$ 0.004	0.059 $\pm$ 0.002	0.061 $\pm$ 0.004	0.062 $\pm$ 0.002	0.062 $\pm$ 0.003
MDLM	0.061 $\pm$ 0.004	0.062 $\pm$ 0.002	0.064 $\pm$ 0.004	0.066 $\pm$ 0.003	0.065 $\pm$ 0.002
ReMDM	0.065 $\pm$ 0.003	0.070 $\pm$ 0.005	0.071 $\pm$ 0.003	0.071 $\pm$ 0.002	0.070 $\pm$ 0.001
SoftMask	0.069 $\pm$ 0.002	0.072 $\pm$ 0.004	0.082 $\pm$ 0.006	0.084 $\pm$ 0.010	0.080 $\pm$ 0.005
SoftMask+ λ	0.071 $\pm$ 0.003	0.072 $\pm$ 0.002	0.084 $\pm$ 0.002	0.081 $\pm$ 0.008	0.086 $\pm$ 0.003
λ -ReMDM	0.067 $\pm$ 0.004	0.074 $\pm$ 0.003	0.076 $\pm$ 0.006	0.079 $\pm$ 0.007	0.072 $\pm$ 0.006
λ -UDLM	0.099 $\pm$ 0.005	0.095 $\pm$ 0.003	0.099 $\pm$ 0.003	0.096 $\pm$ 0.005	0.081 $\pm$ 0.010
Gen. PPL					
UDLM	400.88 $\pm$ 3.76	391.82 $\pm$ 0.97	381.82 $\pm$ 6.23	377.01 $\pm$ 5.13	376.51 $\pm$ 4.49
MDLM	345.20 $\pm$ 2.00	306.74 $\pm$ 4.84	275.55 $\pm$ 2.58	242.99 $\pm$ 0.79	210.49 $\pm$ 4.70
ReMDM	261.20 $\pm$ 4.31	169.75 $\pm$ 3.02	110.26 $\pm$ 1.22	72.86 $\pm$ 1.32	50.39 $\pm$ 0.69
SoftMask	239.00 $\pm$ 1.44	151.95 $\pm$ 0.63	96.41 $\pm$ 1.66	63.85 $\pm$ 0.27	44.85 $\pm$ 0.46
SoftMask+ λ	168.43 $\pm$ 0.32	169.29 $\pm$ 1.32	70.52 $\pm$ 0.29	76.63 $\pm$ 0.91	54.52 $\pm$ 0.64
λ -ReMDM	210.16 $\pm$ 3.80	108.35 $\pm$ 1.67	90.39 $\pm$ 1.17	75.43 $\pm$ 0.67	62.26 $\pm$ 1.62
λ -UDLM	93.86 $\pm$ 1.05	78.94 $\pm$ 1.88	47.82 $\pm$ 1.08	30.46 $\pm$ 1.47	17.48 $\pm$ 0.24

Table 10: Multi-seed unconditional generation results of **small DiT**-based models on **LM1B**. Each entry reports mean $\pm$ standard deviation over three runs. For λ -based methods, λ is selected with respect to MAUVE.

Model	NFE = 64	NFE = 128	NFE = 256	NFE = 512	NFE = 1024
MAUVE					
UDLM	0.078 ± 0.001	0.081 ± 0.003	0.088 ± 0.001	0.087 ± 0.003	0.086 ± 0.001
MDLM	0.078 ± 0.002	0.085 ± 0.004	0.091 ± 0.003	0.094 ± 0.006	0.094 ± 0.004
ReMDM	0.077 ± 0.000	0.087 ± 0.007	0.094 ± 0.002	0.101 ± 0.001	0.098 ± 0.003
SoftMask	0.080 ± 0.002	0.090 ± 0.002	0.101 ± 0.001	0.101 ± 0.002	0.099 ± 0.002
SoftMask+ λ	0.088 ± 0.003	0.099 ± 0.004	0.102 ± 0.002	0.104 ± 0.003	0.102 ± 0.001
λ -ReMDM	0.078 ± 0.001	0.099 ± 0.004	0.116 ± 0.002	0.119 ± 0.008	0.116 ± 0.004
λ -UDLM	0.108 ± 0.005	0.112 ± 0.002	0.122 ± 0.008	0.118 ± 0.002	0.106 ± 0.003
Gen. PPL					
UDLM	237.33 ± 3.34	229.70 ± 5.54	226.97 ± 5.06	221.85 ± 4.55	221.97 ± 2.14
MDLM	224.86 ± 4.69	189.78 ± 2.35	162.91 ± 6.08	142.51 ± 1.42	124.45 ± 2.62
ReMDM	205.09 ± 5.17	128.04 ± 2.04	83.41 ± 1.48	59.18 ± 0.95	43.21 ± 0.82
SoftMask	172.49 ± 2.09	110.78 ± 1.08	73.13 ± 0.15	51.81 ± 0.77	38.27 ± 0.82
SoftMask+ λ	148.19 ± 1.99	83.64 ± 2.04	84.28 ± 1.50	59.69 ± 1.60	43.78 ± 0.51
λ -ReMDM	263.44 ± 1.73	157.06 ± 3.73	92.34 ± 1.56	62.36 ± 0.37	49.98 ± 0.41
λ -UDLM	90.02 ± 0.49	58.89 ± 0.44	42.54 ± 0.42	33.53 ± 0.67	23.70 ± 0.37

Table 11: Multi-seed unconditional generation results of small DiT-based models on **OWT**. Each entry reports mean $\pm$ standard deviation over three runs.

Model	NFE = 64	NFE = 128	NFE = 256	NFE = 512	NFE = 1024
MAUVE					
MDLM	0.244 ± 0.089	0.211 ± 0.017	0.374 ± 0.105	0.577 ± 0.167	0.242 ± 0.108
ReMDM	0.125 ± 0.060	0.300 ± 0.058	0.568 ± 0.072	0.664 ± 0.061	0.625 ± 0.138
SoftMask	0.132 ± 0.046	0.398 ± 0.165	0.753 ± 0.042	0.712 ± 0.064	0.635 ± 0.081
λ -ReMDM (const.)	0.270 ± 0.070	0.605 ± 0.065	0.724 ± 0.165	0.866 ± 0.030	0.782 ± 0.022
λ -UDLM (const.)	0.258 ± 0.039	0.438 ± 0.082	0.575 ± 0.112	0.598 ± 0.105	0.691 ± 0.084
λ -UDLM (log)	0.424 ± 0.044	0.678 ± 0.035	0.729 ± 0.096	0.792 ± 0.089	0.803 ± 0.092
Gen. PPL					
MDLM	72.60 ± 1.24	56.59 ± 0.85	45.26 ± 0.97	37.99 ± 0.40	30.65 ± 0.34
ReMDM	70.88 ± 0.52	43.50 ± 0.59	29.29 ± 0.26	21.12 ± 0.25	15.39 ± 0.30
SoftMask	63.05 ± 0.54	38.60 ± 0.33	25.84 ± 0.17	18.42 ± 0.14	13.82 ± 0.07
λ -ReMDM (const.)	48.50 ± 0.49	29.75 ± 0.14	22.35 ± 0.39	19.97 ± 0.11	17.23 ± 0.02
λ -UDLM (const.)	39.58 ± 0.68	24.64 ± 0.61	16.07 ± 1.07	19.91 ± 0.65	15.65 ± 0.22
λ -UDLM (log)	82.76 ± 1.79	34.69 ± 0.24	21.09 ± 0.31	13.90 ± 0.26	9.88 ± 0.21