Understanding Time-Dependent Stability Boundary in Guidance through Three-regime Dynamics

Yiyu Qiu[*]
Shanghai Jiao Tong University
3500063778@sjtu.edu.cn

Ruofeng Yang[*]
John Hopcroft Center for Computer Science
Shanghai Jiao Tong University
wanshuiyin@sjtu.edu.cn

Tong Yu
Adobe Research
tyu@adobe.com

Shuai Li[†]
The University of Hong Kong
Shanghai Innovation Institute
shuaili@hku.hk

Abstract

Classifier-Free Guidance (CFG) is widely used for conditional diffusion, yet a fixed guidance scale can become excessively aggressive at different stages of the denoising process, leading to artifacts such as over-exposure and color distortion. Existing empirical strategies indicate that the guidance effect is time-dependent, necessitating the application of varying guidance intensities at different stages; however, there remains a lack of understanding regarding when and why specific guidance intensities become unstable. In this work, we study the time-dependent stability of CFG through a Gaussian Mixture Model (GMM) framework under a general SDE formulation. We characterize the guided dynamics into three distinct regimes, and analyze their dependence jointly on the guidance strength ω and reverse time t. In a structured GMM setting, we derive a time-dependent sufficient condition for global non-expansiveness of the scalar update, expressed as $0 \leq \omega(t) - 1 \leq \omega_{\text{upper}}(t)$. Our analysis shows that the stability boundary varies throughout the denoising trajectory, so a fixed guidance scale can transition from a stable regime to a divergent regime as denoising proceeds. For VP and VE diffusion processes, the boundary exhibits exponential and polynomial scaling respectively. These results provide a theoretical perspective on the late stage instability that can arise under high guidance and motivate a time-dependent guidance schedule informed by the temporal evolution of the derived stability boundary. Empirically, we show that a simple schedule inspired by this boundary trend mitigates over-exposure under high guidance settings while maintaining comparable semantic alignment.

1 Introduction

As controllable generation becomes an essential capability of diffusion models, classifier-free guidance (CFG) [Ho and Salimans, 2022] has become a standard approach for improving conditional alignment across image, video, and audio generation [Feng et al., 2024, Hong et al., 2023, Papalampidi et al., 2025, Saharia et al., 2022]. CFG has also been used for data augmentation in offline-to-online reinforcement learning [Huang et al., 2025].Given conditional and unconditional score estimates, CFG modifies the sampling direction as

[*]Equal contribution.
[†]Corresponding author.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

$$s_{\mathrm{CFG}}(x_t, y) = s_{\mathrm{uncond}}(x_t) + \omega \left(s_{\mathrm{cond}}(x_t, y) - s_{\mathrm{uncond}}(x_t) \right),$$

where ω controls the guidance strength. Increasing ω generally strengthens the influence of the condition and can improve semantic alignment, which makes large guidance scales common in practical generation. However, excessively strong guidance can also distort the sampling trajectory and lead to artifacts and over-exposure [Chidambaram et al., 2024, Sadat et al., 2025], as illustrated in Figure 1. This suggests that the effect of guidance may not increase smoothly with its strength. High guidance strength may eventually move the sampling dynamics into a qualitatively different regime, raising the possibility of a critical boundary separating stable guidance from overly aggressive dynamics. A complication is that such a boundary need not remain fixed throughout denoising. Recent studies have shown that the effect of CFG varies substantially across denoising stages [Jin et al., 2026, Gao et al., 2026, Jiang and Ma, 2026, Rojas et al., 2026]. This suggests that the range of guidance strengths associated with stable dynamics may also evolve along the reverse trajectory.

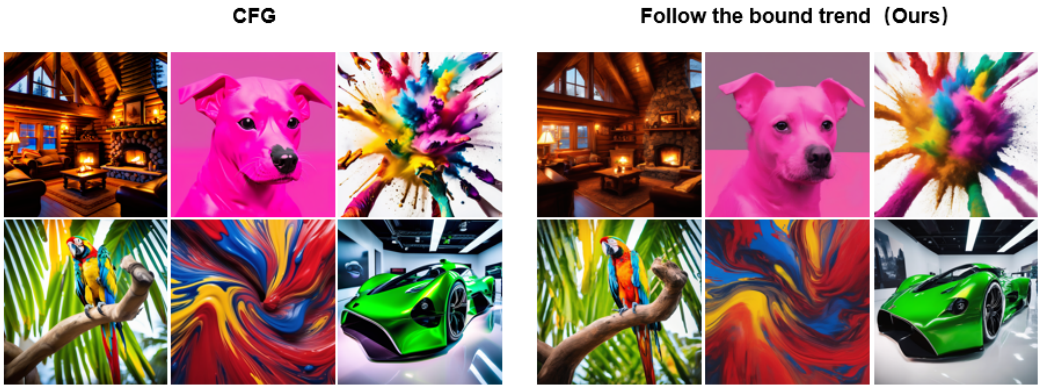

Figure 1: Comparison on SDXL (50 steps) at large guidance ($\omega = 20$). (Left) Constant CFG, Maintains alignment but suffers from discretization-induced "overexposure". (Right) Our schedule follows the temporal trend of the derived stability boundary and mitigates these artifacts.

Recent studies show CFG impacts denoising stages heterogeneously, motivating timestep-dependent strategies [Jin et al., 2026, Gao et al., 2026, Jiang and Ma, 2026, Rojas et al., 2026]. Related empirical methods explore guidance truncation and other time-dependent schedules [Kynkäänniemi et al., 2024, Liu et al., 2024, Balaji et al., 2022, Sadat et al., 2025]. These works establish the temporal dependence of guidance, but primarily focus on optimal schedule allocation. A complementary question remains largely unexplored. At any timestep, how large can the guidance scale grow before the dynamics undergoes a qualitative change? Instead of searching for specific schedules, we investigate how the admissible range of guidance strength evolves with denoising time. This shifts the focus from designing a particular schedule to characterizing the stability boundary encountered by high guidance trajectories, raising a fundamental question concerning the stability of guided sampling dynamics:

how does the stability of a given guidance strength evolve as the denoising process proceeds?

To study this question, we analyze CFG from the perspective of dynamical stability. We adopt shared-covariance collinear Gaussian mixture models as a tractable setting and use a Mahalanobis-based transformation to separate diffusion-time geometry from guidance induced mode dynamics. This allows us to characterize the guided dynamics as convergent, mixed, or divergent. We further derive a time dependent stability boundary $\omega_{\mathrm{upper}}(t)$ for global non-expansiveness. The boundary evolves throughout the reverse trajectory, so a fixed guidance scale can move from a stable to an unstable regime as denoising proceeds.

Importantly, $\omega_{\mathrm{upper}}(t)$ is a model-conditional stability boundary rather than an optimal guidance scale or a certified guarantee for practical neural diffusion models. Instead, it provides a theoretically grounded reference for identifying when guidance becomes excessively aggressive. We use its temporal trend to construct a simple schedule with a similar profile, aiming to reduce boundary crossing, and empirically examine whether it can mitigate the instability and associated artifacts caused by such crossings.

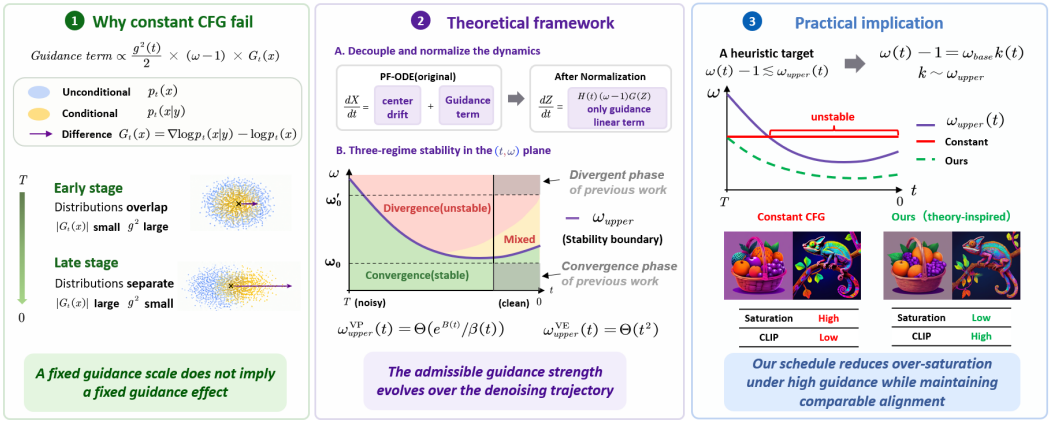

Figure 2: Overview of our framework. (Left) constant guidance does not imply a constant guidance effect, as the effective guidance term depends on both the guidance scale and diffusion dynamics. (Middle) we decouple diffusion-time geometry from guidance-induced mode dynamics and characterize three dynamical regimes separated by a time-dependent stability boundary $\omega_{\text{upper}}(t)$. (Right) we use the temporal trend of this boundary to construct a trend-following guidance schedule and evaluate its effect on high-guidance sampling.

1.1 Our Contribution

In this work, we make the first step toward answering this question by characterizing the fine-grained dynamics of the reverse process. Our contributions are as follows:

A unified stability framework. We introduce a Mahalanobis-based transformation that separates diffusion time geometry from guidance mode dynamics, reducing the stability analysis of shared-covariance GMM to a tractable representative configuration.

Three-regime characterization and time-dependent boundary. We characterize the Convergence, Mixed, and Divergence regimes and derive a time-dependent stability boundary $\omega_{upper}(t)$ that specifies the upper bound on guidance scale for global non-expansiveness under our model assumptions.

Practical implication for high guidance sampling. We show that the guidance upper bound evolves along the reverse trajectory, providing a stability-based explanation for late-stage artifacts under fixed high guidance. We further use this temporal trend to construct a simple theory-inspired schedule and examine its practical effect in diffusion sampling.

2 Related Work

For guidance-based diffusion models, several works analyze the mechanism of guidance from different perspectives. Li et al. [2025] decompose CFG into different terms and study their roles from a contrastive perspective, while Bradley and Nakkiran [2024] connect CFG with the predictor-corrector framework. Other theoretical studies investigate the effects of guidance on classification confidence, sample diversity, and extreme guidance. Under a nearly orthogonal GMM assumption, Wu et al. [2024] provide convergence guarantees for classification confidence under both the reverse SDE and probability flow ODE, while Li and Jiao [2025] consider more general data distributions and characterize the preferential enhancement of high-probability classes. For extreme guidance, Wu et al. [2024] and Chidambaram et al. [2024] show that sufficiently strong guidance can induce mode collapse or related instabilities in GMM-based settings. Similar trade-offs between alignment and task performance arise in cooperative AI [Ma et al., 2025]. More recently, Yang et al. [2026b] analyze the convergence rate with respect to guidance scale for DDPM-type and EDM-type samplers.

A related line of work studies the temporal dependence of guidance during denoising and uses it to motivate time-dependent guidance. Jin et al. [2026] provide a stage-wise dynamical analysis of CFG,

while Gao et al. [2026] derive time-dependent bounds from the discrepancy between conditional and unconditional score fields. Jiang and Ma [2026] further characterize the stage-dependent distribution induced by CFG and construct corresponding guidance schedules. Related studies in masked diffusion also suggest that guidance effects vary across denoising stages [Rojas et al., 2026]. Beyond theoretical analyses, empirical methods have explored timestep-dependent guidance through truncation, interval guidance, and other scheduling strategies [Kynkäänniemi et al., 2024, Liu et al., 2024, Balaji et al., 2022, Sadat et al., 2025, Malarz et al., 2025]. Together, these works establish the temporal dependence of guidance and demonstrate that varying guidance over the sampling trajectory can affect generation behavior.

The most relevant to our work are Wu et al. [2024] and Chidambaram et al. [2024], which analyze CFG using Gaussian Mixture Models (GMM) but typically assume orthogonal modes and constant scales. While these analyses provide valuable insights into the effect of guidance, they do not explicitly characterize how the stability transition evolves jointly with the guidance strength and the denoising time. Existing theories typically provide bounds only for very small or sufficiently large guidance, leaving the vast sensitive areas in the (ω, t) plane unexplored. Figure 2 explicitly maps these regions.

3 Preliminaries

General SDE/PFODE framework. Consider a data distribution p^* over (X, y) with $X \in \mathbb{R}^d$. The forward diffusion process $\{X_t\}_{t=0}^T$ is governed by the SDE

$$\mathrm{d}X_t = f(t)X_t\mathrm{d}t + g(t)\mathrm{d}B_t, \quad X_0 \sim p^*(\cdot|y), \quad p_t(X_t|X_0) = \mathcal{N}(X_t; \gamma(t)X_0, \sigma^2(t)\mathbf{I})$$

where $f(t), g(t)$ are scalar coefficients and B_t is a d-dimensional Wiener process.

For reverse-time sampling, we parameterize the Probability Flow ODE using the same diffusion-time variable t, which evolves from T to 0. Thus, $\mathrm{d}t < 0$ along the reverse trajectory.

$$\frac{\mathrm{d}X}{\mathrm{d}t} = f(t)X - \frac{1}{2}g^2(t)\nabla_X \log p_t(X). \quad dt < 0$$

Classifier-Free Guidance (CFG). CFG modifies the score estimate by extrapolating the conditional score $\nabla_X \log p_t(X|y)$ away from the unconditional score $\nabla_X \log p_t(X)$ with a guidance scale ω. Let $G(X,t) = \nabla_X \log p_t(y|X) = \nabla_X \log p_t(X|y) - \nabla_X \log p_t(X)$ denote the guidance direction.

$$\begin{aligned}\frac{\mathrm{d}X}{\mathrm{d}t} &= f(t)X - \frac{1}{2}g^2(t)(\nabla_X \log p_t(X) + \omega\nabla_X \log p_t(y|X)) \\ &= \underbrace{f(t)X - \frac{1}{2}g^2(t)\nabla_X \log p_t(X|y)}_{\text{Conditional Drift}} \underbrace{- \frac{1}{2}g^2(t)(\omega-1)G(X,t)}_{\text{Guidance Term}}.\end{aligned}$$

Analytical Setting. We study the stability of classifier-free guidance under a shared-covariance Gaussian mixture model. The following assumptions define the analytical setting considered throughout our theoretical analysis.

Assumption 3.1. The data distribution satisfies the following conditions:

1. **Shared covariance.** All components share a common positive-definite covariance Σ:

$$p^*(X) = \sum_{i=1}^{N} \pi_i \mathcal{N}(X; \mu_i, \Sigma).$$

2. **Collinear component means.** There exists a unit vector $\hat{\mu}$ for all components i such that

$$\mu_i = l_i\hat{\mu}, \qquad \|\hat{\mu}\| = 1.$$

3. **Alignment of the covariance.** The covariance preserves the direction spanned by the component means:

$$\Sigma\hat{\mu} = \lambda\hat{\mu}, \qquad \lambda > 0.$$

These conditions enable a scalar reduction of the guided dynamics, providing a tractable setting for isolating the interaction between guidance strength and discrete-time stability. Our theoretical guarantees are established under these assumptions, while additional experiments examine the qualitative behavior beyond this setting.

4 Three-regime Stability Analysis of CFG

4.1 From High-Dimensional Dynamics to a Scalar Representation

To separate the diffusion-induced geometry from the relative configuration of the mixture components, we introduce the time-dependent Mahalanobis transformation

$$Z_t = \Sigma_t^{-1/2}\left(X_t - \mu_{t,y}\right), \tag{1}$$

where $\mu_{t,y}$ and Σ_t denote the time-dependent mean and covariance of the target component.

Lemma 4.1. *(Proof in App. A.1) Under Assumption 3.1, the transformation in Eq. 1 maps the target component to an isotropic Gaussian and separates the diffusion-time geometry from the mode drift of the mixture components. In the transformed coordinates, the probability flow ODE can be written as*

$$\frac{\mathrm{d}Z}{\mathrm{d}t} = -\underbrace{\frac{1}{2}g^2\Sigma_t^{-1}}_{\mathcal{H}(t)}(\omega - 1)\ \underbrace{G(Z)}_{\textit{Mode Drift}}, \tag{2}$$

where $\mathcal{H}(t)$ collects the diffusion-time and covariance-dependent factors, while $G(Z)$ characterizes the interaction between mixture modes.

Under the assumption, the transformed component means are collinear, i.e., $\mu'_{t,i} = l'_{t,i}\hat{\mu}$, where $\|\hat{\mu}\| = 1$. Consequently, the guidance-induced vector field is aligned with $\hat{\mu}$, and the posterior mixture weights depend on z_t only through its projection onto this direction. This motivates the scalar coordinate $v_t = \langle z_t, \hat{\mu}\rangle$, which reduces the mode-competition dynamics to a one-dimensional representation (see App. A.2).

4.2 Time-Dependent Stability Boundary through three-regime dynamics

The scalar dynamics exhibit a restorative structure in the continuous reverse-time limit. In the sign-indefinite case, the scalar guidance field $G_t(v)$ is monotone and admits a unique zero $v_c(t)$, toward which the reverse-time drift is directed (see App. A.2). In discrete sampling, the step size becomes an additional factor in sampling guarantees [Yang et al., 2025]. In our setting, a sufficiently large update may overshoot the instantaneous equilibrium and violate non-expansiveness, motivating a time-dependent stability bound.

Proposition 4.2. *The guidance-induced dynamics depend on the state z_t only through the scalar variable $v_t = \langle z_t, \hat{\mu}\rangle$. Therefore, we express the discretized update in a one-dimensional form.*

$$v_{t+dt} = \Phi_t(v_t),$$

where Φ_t denotes the one-step reverse update induced by the guided PF-ODE in Eq. (2)*.*

Definition 4.3. We therefore characterize the discrete update $v_{t-\delta_t} = \Phi_t(v_t)$ as two dynamical cases through its contraction property:

- *Squeezing (Contractive):* The update is contractive, driving the sample toward a stable fixed point.
- *Pushing (Repulsive):* The guidance-induced update is over-amplified by the discretization step $\delta_t(-dt)$, causing local expansion or overshooting away from the target.

Definition 4.4. We define the Pushing set as

$$\mathcal{S}_t^p = \left\{ v \in \mathbb{R} : \left|\frac{\partial\Phi_t(v)}{\partial v}\right| > 1 \right\} \tag{3}$$

denotes the states exhibiting non-contractive dynamics at sampling time t.

Definition 4.5 (Three Stability Regimes)**.** We distinguish three qualitatively different regimes according to the structure of $\mathcal{S}_t^p$:

- **Convergence Regime:** $\mathcal{S}^p = \emptyset$. The update is globally non-expansive; in the sign-indefinite case, v_c is the unique stable equilibrium.
- **Mixed Regime:** $\mathcal{S}^p = \bigcup_i [a_{t,i}, b_{t,i}]$ consists of multiple disjoint intervals. While v_c may remain locally stable, the state space is fragmented by multiple repulsive barriers, leading to local mode-trapping.
- **Divergence Regime:** $\mathcal{S}^p = [a_t, b_t]$ collapses into a **single contiguous interval** that contains the target (i.e., $v_c \in [a_t, b_t]$). The central fixed point becomes locally unstable, and sufficiently small nonzero perturbations are amplified by the discrete update.

While the exact boundaries for these regimes in GMM involve solving a series of transcendental equations for multiple $[a_{t,i}, b_{t,i}]$ intervals, the condition to remain in the Convergence regime (where no Pushing regions exist) obeys a universal scaling law:

Theorem 4.6. *(Proof in App. A.3) Under Assumptions 3.1 and an explicit Euler discretization with reverse-time step size $\delta_t > 0$, we derive a time-dependent sufficient upper bound $\omega_{\text{upper}}(t)$. Whenever guidance scales satisfying $0 \le \omega(t) - 1 \le \omega_{\text{upper}}(t)$, the corresponding scalar update Φ_t is non-expansive. The upper bound scale obeys*

$$\omega_{\text{upper}}(t) = \Theta\left(\frac{[\gamma^2(t)\lambda + \sigma^2(t)]^2}{g^2(t)\gamma^2(t)\delta_t}\right). \tag{4}$$

up to constants independent of t. $g(t)$, $\gamma(t)$, and $\sigma(t)$ denote the corresponding diffusion coefficients under the transformed coordinate system(Section 3).

Proof Sketch. A sufficient condition for global non-expansiveness is $|\Phi_t'(v)| \le 1$ for all $v \in \mathbb{R}$. Differentiating the scalar update yields a posterior-variance term:

$$\Phi_t'(v) = 1 - \frac{g^2(t)(\omega(t) - 1)\delta_t}{2(\gamma^2(t)\lambda + \sigma^2(t))} \operatorname{Var}_{p_t(v)}(l'_{t,i}).$$

For $\omega(t) \ge 1$, we have $\Phi_t'(v) \le 1$, so it suffices to control the lower bound $\Phi_t'(v) \ge -1$. Applying the uniform variance bound from App. A.3 gives

$$(\omega(t) - 1)\frac{g^2(t)\gamma^2(t)\delta_t(\max_i l_i - \min_i l_i)^2}{16(\gamma^2(t)\lambda + \sigma^2(t))^2} \le 1,$$

which yields the stated sufficient upper bound.

Corollary 4.7. *(Proof in App. A.5) For a constant discretization step $\delta_t = \delta$, the sufficient stability boundary exhibits distinct asymptotic scalings under different diffusion parameterizations in the large-forward-time regime:*

- ***VP (constant β, Gao et al. [2025]):** Exponential scaling $\Theta(e^{At})$.*
- ***VP (linear β, Song et al. [2021]):** Exponential-quadratic scaling $\Theta(e^{B(t)}/\beta(t))$, where $B(t)$ is quadratic and $\beta(t)$ is linear in t.*
- ***VE ($\sigma(t) = \sqrt{t}$, Song et al. [2021], Karras et al. [2022]):** Polynomial scaling $\Theta(t^2)$.*
- ***EDM ($\sigma(t) = t$,Karras et al. [2022]):** Polynomial scaling $\Theta(t^3)$.*

4.3 Exact Regime Characterization in a Symmetric 3GMM

The analysis above establishes the three dynamical regimes and characterizes the time-dependent stability boundary through its asymptotic scaling. To explicitly resolve the regime transitions, we further consider a symmetric three-component GMM under the VPSDE with $\beta(t) = 2$:

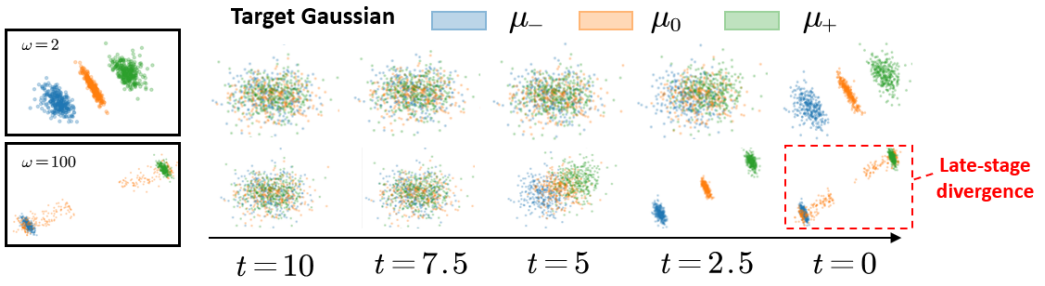

Figure 4: Mechanistic illustration of guidance-induced divergence in the collinear 3-component GMM. Moderate guidance ($\omega = 2$) preserves the mixture structure, whereas strong guidance ($\omega = 100$) induces outward splitting of trajectories from the central component.

$$p_{\text{data}} = \frac{1}{3} \sum_{y \in \{-,0,+\}} \mathcal{N}(\mu_y, \mathbf{I}), \qquad \mu_+ = -\mu_- = l\hat{\mu}, \qquad \mu_0 = 0. \tag{5}$$

This additional symmetry makes the discrete reverse dynamics analytically easy while retaining the three regimes identified above.

Closed-Form Stability Thresholds. (Proof in App. A.6) The boundaries of the regimes are partitioned by two analytical threshold functions j and g, where j is bounded by an auxiliary function f. let $l_t = e^{-t}l$, $M_t = e^{l_t^2/2}$, δ_t represents the discrete step size.

$$f(l, \delta_t, t) \le j(l, \delta_t, t) < g(l, \delta_t, t), \quad g(l, \delta_t, t) = \frac{M_t + 2}{\delta_t l_t^2}, \quad f(l, \delta_t, t) = 8\frac{(1 - 4/M_t^2)}{\delta_t l_t^2} \tag{6}$$

Convergence regime: $\omega - 1 \le j$ ($M_t > 4$) or $\omega - 1 \le g$ ($M_t \le 4$).

Mixed regime: $j < \omega - 1 \le g$. Only appearing when $M_t > 4$.

Divergence regime: $\omega - 1 > g$.

The regime transitions are summarized in Fig. 3.

Consistency with the General Boundary. As shown in App. A.6.5, the asymptotic behavior of the exact thresholds with respect to time above satisfies

$$\omega_{\text{upper}}(t) = \Theta\left(\frac{e^{2t}}{\delta_t}\right), \qquad t : T \to 0^+, \tag{7}$$

which is consistent with the VPSDE specialization of Corollary 4.7.

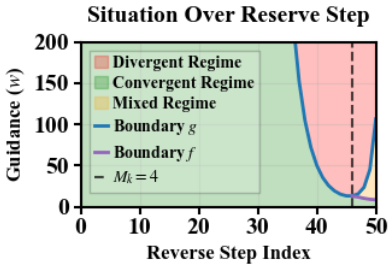

Figure 3: Characterized thresholds of regime transitions under the setting 5. Hyperparameter selection in Appendix B.1

.

4.4 From the Stability Boundary to a Practical Instantiation

The derived stability boundary characterizes how the sufficiently stable range of guidance scales varies with diffusion time. This provides a simple interpretation of why a fixed high guidance scale can become unstable at later stages. When ω remains constant, the evolving boundary may move below it, so the non-expansiveness guarantee no longer applies.

To investigate how this temporal dependence can inform practical sampling, we consider a simple parameterization motivated by the boundary under specific settings (Selection of specific parameters for this method and other strategies are shown in Appendix B.3).

$$\omega(t) - 1 = \omega_{\text{base}} k(t), \tag{8}$$

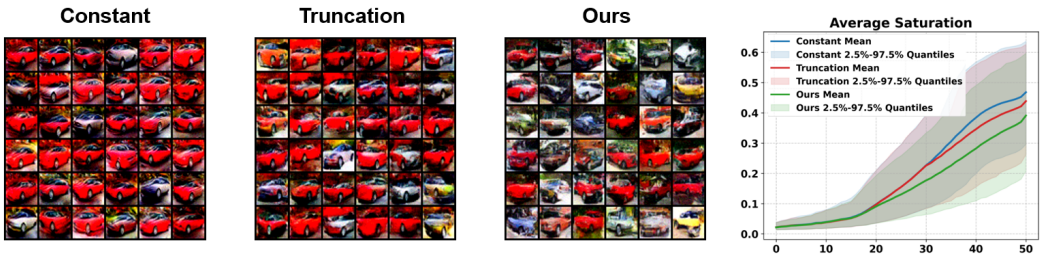

Figure 5: Controlled comparison on CIFAR-10 under high guidance ($\omega = 20$). Constant guidance exhibits pronounced late-stage saturation, while the theory-inspired schedule suppresses this increase more effectively than truncation.

where ω_{base} denotes the maximum excess guidance scale and $k(t)$ is a normalized temporal profile motivated by $\omega_{\text{upper}}(t)$.

The resulting schedule serves as a simple, theory-inspired instantiation rather than an optimality claim. The derived boundary provides a sufficient condition for non-expansiveness, but does not characterize sample quality within the bounded region; consequently, following the boundary's temporal trend does not guarantee improved generation quality. We evaluate this schedule against constant guidance, truncation, and alternative schedules across different initial guidance scales in Section 5.3.

Scope under approximate scores. Our stability analysis assumes exact score fields in the structured GMM setting, whereas practical neural diffusion models use estimated scores. The resulting approximation errors perturb the guided dynamics and can propagate along the sampling trajectory. Complementary work by He et al. [2026] characterizes score-error accumulation through stepwise Wasserstein bounds for VP and VE diffusions. This motivates considering both discretization stability and error propagation when designing practical guidance schedules. In particular, a permissive early-time stability bound alone does not justify large early guidance, since the present analysis does not account for the errors in the estimated conditional and unconditional scores. Quantifying their joint effect with guidance and developing error-aware schedules are left to future work.

5 Experimental Validation

5.1 Mechanistic Validation on GMM

We first examine the mechanism identified by our theoretical analysis in a controlled two-dimensional setting. We substitute $l\hat{\mu} = (2, 1)^T$ and $T = 10$ into Equation (5). Since the score function of the GMM admits a closed-form expression, we directly evaluate the exact score during PF-ODE sampling without introducing neural network approximation error.

Guidance-Induced Divergence. We compare moderate guidance ($\omega = 2$) with strong guidance ($\omega = 100$). As shown in Figure 4, moderate guidance preserves the three mixture modes, whereas strong guidance causes samples initialized around the middle component to split outward toward the neighboring modes. The corresponding trajectories directly visualize the transition from stable contraction to outward pushing predicted by our regime analysis.

5.2 Controlled Analysis on CIFAR-10

To examine whether the stability trends identified in our analysis are observable in image generation, we conduct a controlled experiment on CIFAR-10 under high guidance ($\omega = 20$). We compare constant guidance, guidance truncation, and our theory-inspired schedule.

As shown in Figure 5, constant guidance produces pronounced saturation and color distortion, while truncation partially mitigates these artifacts. Our theory-inspired schedule suppresses the saturation more effectively. The temporal saturation curve shows the same trend: constant guidance exhibits a sharp increase in saturation toward the late stages, whereas our schedule largely suppresses this growth. This behavior is consistent with the time-dependent stability trend identified in our analysis.

Scheduling schemes, additional quantitative results and the definition of the stability metrics are provided in Appendix B.4.

5.3 Real-World Evaluation on SDXL

We next evaluate whether the stability trends observed in the controlled experiments extend to practical text-to-image diffusion models. Our main experiments use SDXL, with additional results on SD1.5 provided in Appendix C.2.

Mitigation of Artifacts under High Guidance. As shown in Figure 6, at a high guidance scale of $\omega = 20$, constant guidance produces pronounced over-exposure and color saturation, consistent with the high-guidance instability observed in the controlled experiments. Guidance truncation mitigates these artifacts by removing guidance after a predefined timestep. Our schedule instead follows the temporal trend suggested by the derived boundary, and provides a smoother transition from strong guidance to weaker guidance, which mitigates late-stage artifacts better while maintaining comparable semantic alignment.

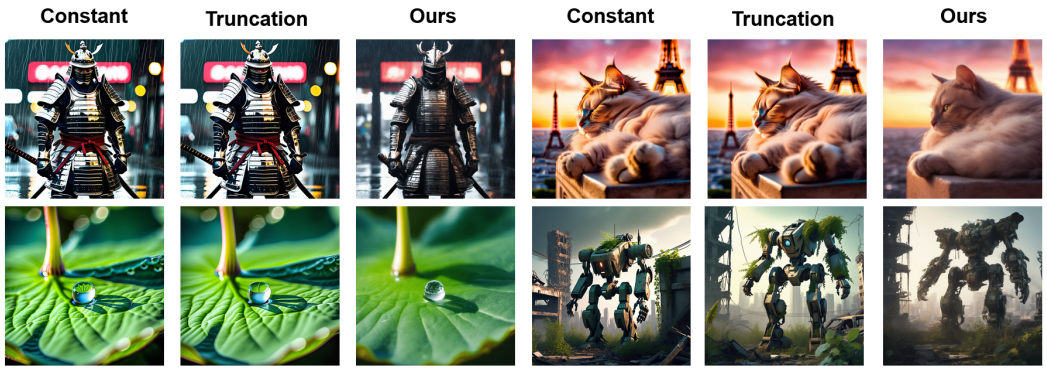

Figure 6: Visual comparisons of Stable Diffusion XL at high guidance scales ($\omega = 20$). These examples illustrate reduced saturation under our schedule.

Alignment-Stability Trade-offs. We quantify this trade-off using CLIP Score, saturation, and Out-of-Bounds Rate (OBR) in Fig. 7. Our schedule maintains a CLIP Score comparable to constant guidance while substantially reducing saturation and OBR. Detailed metric definitions are provided in Appendix B.4. To further demonstrate the robustness and consistency of our method, we provide more qualitative results across a wider range of prompts and random seeds in Appendix C.1.

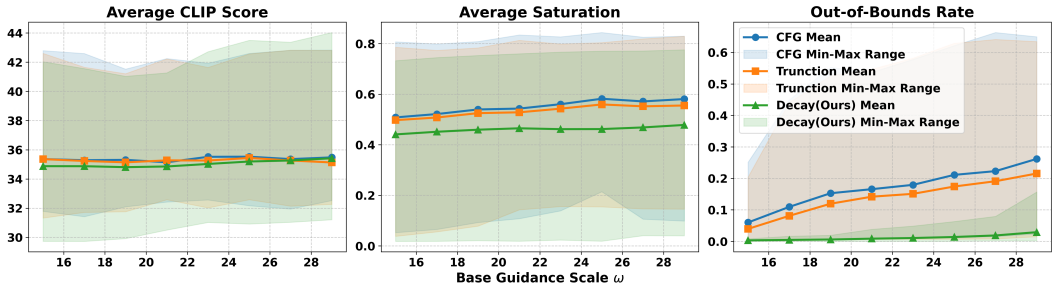

Figure 7: Comparison of guidance strategies on 50 random prompts using 50-step inference. Our schedule maintains comparable CLIP alignment while reducing saturation and OBR under high guidance. Detailed metric definitions are provided in Appendix B.4.

Effectiveness across Guidance strength. The stability boundary does not imply that boundary trend schedule is universally optimal. Its benefit should arise primarily when the guidance strength

becomes sufficiently large to approach or exceed the time-dependent stability boundary. To examine this regime dependence, we compare constant CFG and our strategy across a range of guidance scales. At relatively low guidance, where the constant scale remains within the stable regime, constant CFG can achieve comparable or better generation quality, while unnecessary decay may weaken conditional guidance. As the guidance scale increases, however, constant CFG becomes increasingly susceptible to boundary crossing, accompanied by rising saturation and out-of-bounds rates. In this regime, our strategy mitigates these instabilities while maintaining comparable semantic alignment. We report CLIP, FID, saturation, and OBR across guidance scales in Table 1.

Guidance	Method	CLIP ↑	FID ↓	Saturation ↓	OBR(%) ↓
$\omega_{max} = 5$	Constant	**35.11**	19.13	0.325	0.08
	Truncation	35.02	**18.99**	0.321	0.07
	Ours	34.11	19.24	**0.268**	**0.03**
$\omega_{max} = 15$	Constant	35.04	23.46	0.45	6.2
	Truncation	34.88	23.92	0.43	4.8
	Ours	34.53	**19.73**	**0.36**	**0.3**

Table 1: Comparison across different guidance regimes on SDXL. All methods use 50-step inference. CLIP Score, Saturation, and OBR are computed over 50 random prompts, while FID is evaluated on 10,000 generated samples from the COCO-30K benchmark.

6 Conclusion

In this paper, we developed a theoretical framework for understanding the stability dynamics of Classifier-Free Guidance (CFG) during diffusion sampling. By reducing the reverse dynamics to a tractable scalar stability process, we identified three qualitatively distinct regimes. This analysis shows that over-exposure is not merely an empirical artifact of large guidance scales, but can arise when the sampling trajectory enters an unstable regime during denoising. More importantly, we derived a time-dependent stability boundary that characterizes the range of guidance scales under which convergence can be maintained. Its temporal behavior depends on the diffusion parameterization, exhibiting distinct exponential or polynomial scaling behaviors under the considered VP, VE, and EDM settings.

The resulting boundary provides a principled reference for interpreting existing guidance heuristics and for designing practical guidance policies. As a concrete theory-inspired instantiation, we construct a simple schedule that follows the temporal trend of the derived boundary and empirically observe reduced late-stage saturation under high guidance. We emphasize that this construction is not intended as a universally optimal schedule; its purpose is to demonstrate how the theoretical boundary can be translated into a practical mechanism for mitigating the instability predicted by the analysis.

Limitations and Future Work. Our theoretical analysis relies on a structured shared-covariance GMM proxy, which provides tractability but does not capture the full geometry of real-world data distributions. Extending our stability analysis to richer mixture structures, such as multi-subspace, multimodal models [Yang et al., 2026a], remains an interesting direction. In addition, our current analysis focuses on deterministic PF-ODE dynamics. Extending the stability characterization to stochastic SDE sampling may reveal additional effects of noise on regime transitions and stability boundaries. Finally, the present schedule follows only the temporal trend of the derived boundary; developing adaptive guidance controllers that directly estimate or track local stability in real time remains an open problem.

Impact Statement

This work studies the stability of classifier-free guidance from a theoretical perspective and provides a framework for understanding how guidance strength interacts with the evolving denoising dynamics. We hope that a clearer understanding of these stability mechanisms can contribute to more predictable and reliable use of diffusion-based generative models. Potential negative impacts include the misuse of diffusion-based image generation to create misleading or deceptive synthetic imagery.

Acknowledgments and Disclosure of Funding

The corresponding author, Shuai Li, acknowledges support from the National Natural Science Foundation of China (Grant No. 92570111).

The authors declare no competing interests.

References

Yogesh Balaji, Seungjun Nah, Xun Huang, Arash Vahdat, Jiaming Song, Qinsheng Zhang, Karsten Kreis, Miika Aittala, Timo Aila, Samuli Laine, et al. ediff-i: Text-to-image diffusion models with an ensemble of expert denoisers. *arXiv preprint arXiv:2211.01324*, 2022.

Arwen Bradley and Preetum Nakkiran. Classifier-free guidance is a predictor-corrector. *arXiv preprint arXiv:2408.09000*, 2024.

Muthu Chidambaram, Khashayar Gatmiry, Sitan Chen, Holden Lee, and Jianfeng Lu. What does guidance do? a fine-grained analysis in a simple setting. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.

Kunyu Feng, Yue Ma, Bingyuan Wang, Chenyang Qi, Haozhe Chen, Qifeng Chen, and Zeyu Wang. Dit4edit: Diffusion transformer for image editing. *arXiv preprint arXiv:2411.03286*, 2024.

Jiayang Gao, Tianyi Zheng, Jiayang Zou, Fengxiang Yang, Shice Liu, Luyao Fan, Zheyu Zhang, Hao Zhang, Jinwei Chen, Peng-Tao Jiang, Bo Li, and Jia Wang. C^2fg: Control classifier-free guidance via score discrepancy analysis, 2026. URL `https://arxiv.org/abs/2603.08155`.

Xuefeng Gao, Hoang M. Nguyen, and Lingjiong Zhu. Wasserstein convergence guarantees for a general class of score-based generative models. *Journal of Machine Learning Research*, 26(43):1–54, 2025. URL `http://jmlr.org/papers/v26/24-0902.html`.

Baoxiang He, Valentio Iverson, Shuai Li, Cheng Chen, and Bo Jiang. The accumulation of score estimation error in diffusion models. In Tong Zhang, Miroslav Dudik, Martin Jaggi, Alekh Agarwal, Sharon Li, Dale Schuurmans, Jerry Zhu, Felix Berkenkamp, Hanze Dong, and Alberto Bietti, editors, *Proceedings of the 43rd International Conference on Machine Learning*, volume 306 of *Proceedings of Machine Learning Research*, pages 41458–41478. PMLR, 06–11 Jul 2026. URL `https://proceedings.mlr.press/v306/he26q.html`.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022.

Susung Hong, Gyuseong Lee, Wooseok Jang, and Seungryong Kim. Improving sample quality of diffusion models using self-attention guidance. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7462–7471, 2023.

Xiao Huang, Xu Liu, Enze Zhang, Tong Yu, and Shuai Li. Offline-to-online reinforcement learning with classifier-free diffusion generation. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu, editors, *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pages 25581–25592. PMLR, 13–19 Jul 2025. URL `https://proceedings.mlr.press/v267/huang25v.html`.

Enze Jiang and Zheng Ma. Analytic distribution of classifier-free guidance for schedule design, 2026. URL `https://arxiv.org/abs/2607.19725`.

Cheng Jin, Qitan Shi, and Yuantao Gu. Stage-wise dynamics of classifier-free guidance in diffusion models. In *The Fourteenth International Conference on Learning Representations*, 2026. URL `https://openreview.net/forum?id=fP0s1TEow3`.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho, editors, *Advances in Neural Information Processing Systems*, 2022. URL `https://openreview.net/forum?id=k7FuTOWMOc7`.

Tuomas Kynkäänniemi, Miika Aittala, Tero Karras, Samuli Laine, Timo Aila, and Jaakko Lehtinen. Applying guidance in a limited interval improves sample and distribution quality in diffusion models. *Advances in Neural Information Processing Systems*, 37:122458–122483, 2024.

Gen Li and Yuchen Jiao. Provable efficiency of guidance in diffusion models for general data distribution. *arXiv preprint arXiv:2505.01382*, 2025.

Xiang Li, Rongrong Wang, and Qing Qu. Towards understanding the mechanisms of classifier-free guidance. *arXiv preprint arXiv:2505.19210*, 2025.

Haozhe Liu, Wentian Zhang, Jinheng Xie, Francesco Faccio, Mengmeng Xu, Tao Xiang, Mike Zheng Shou, Juan-Manuel Perez-Rua, and Jürgen Schmidhuber. Faster diffusion via temporal attention decomposition. *arXiv preprint arXiv:2404.02747*, 2024.

Haitong Ma, Haoran Yu, Haobo Fu, and Shuai Li. Learning preferences without interaction for cooperative AI: A hybrid offline-online approach. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=QVheAhJefR.

Dawid Malarz, Artur Kasymov, Maciej Zieba, Jacek Tabor, and Przemyslaw Spurek. Classifier-free guidance with adaptive scaling. *CoRR*, abs/2502.10574, February 2025. URL https://doi.org/10.48550/arXiv.2502.10574.

Pinelopi Papalampidi, Olivia Wiles, Ira Ktena, Aleksandar Shtedritski, Emanuele Bugliarello, Ivana Kajic, Isabela Albuquerque, and Aida Nematzadeh. Dynamic classifier-free diffusion guidance via online feedback. *arXiv preprint arXiv:2509.16131*, 2025.

Dustin Podell, Zion English, Kyle Lacey, Andreas Blattmann, Tim Dockhorn, Jonas Müller, Joe Penna, and Robin Rombach. SDXL: Improving latent diffusion models for high-resolution image synthesis. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=di52zR8xgf.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. *CoRR*, abs/2103.00020, 2021. URL https://arxiv.org/abs/2103.00020.

Kevin Rojas, Ye He, Chieh-Hsin Lai, Yuhta Takida, Yuki Mitsufuji, and Molei Tao. Improving classifier-free guidance in masked diffusion: Low-dim theoretical insights with high-dim impact. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=mMK9pvQJxf.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *CVPR*, pages 10674–10685, 2022. URL https://doi.org/10.1109/CVPR52688.2022.01042.

Seyedmorteza Sadat, Otmar Hilliges, and Romann M. Weber. Eliminating oversaturation and artifacts of high guidance scales in diffusion models. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=e2ONKX6qzJ.

Chitwan Saharia, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily L Denton, Kamyar Ghasemipour, Raphael Gontijo Lopes, Burcu Karagol Ayan, Tim Salimans, et al. Photorealistic text-to-image diffusion models with deep language understanding. *Advances in neural information processing systems*, 35:36479–36494, 2022.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *ICLR*, 2021. URL https://openreview.net/forum?id=PxTIG12RRHS.

Yuchen Wu, Minshuo Chen, Zihao Li, Mengdi Wang, and Yuting Wei. Theoretical insights for diffusion guidance: A case study for gaussian mixture models. *arXiv preprint arXiv:2403.01639*, 2024.

Ruofeng Yang, Bo Jiang, and Shuai Li. The polynomial iteration complexity for variance exploding diffusion models: Elucidating sde and ode samplers. In Yingzhen Li, Stephan Mandt, Shipra Agrawal, and Emtiyaz Khan, editors, *Proceedings of The 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pages 730–738. PMLR, 03–05 May 2025. URL https://proceedings.mlr.press/v258/yang25a.html.

Ruofeng Yang, Yongcan Li, Bo Jiang, Cheng Chen, and Shuai Li. Multi-subspace multi-modal modeling for diffusion models: Estimation, convergence and mixture of experts. In *The Fourteenth International Conference on Learning Representations*, 2026a. URL https://openreview.net/forum?id=MPWIM6rxxU.

Ruofeng Yang, Yiyu Qiu, Bo Jiang, Cheng Chen, and Shuai Li. Elucidating guidance in variance exploding diffusion models: Fast convergence and better diversity. In *ICLR 2026 2nd Workshop on Deep Generative Model in Machine Learning: Theory, Principle and Efficacy*, 2026b. URL https://openreview.net/forum?id=PmcV9vW2qZ.

Appendix

A Proofs and Detailed Derivations

A.1 GMM Generalization

Let us consider an arbitrary N Gaussian Mixture Model $p(x) = \sum_i^N \pi_i N(x; \mu_i, \Sigma)$, the forward SDE and reverse PFODE are given by

Forward SDE

$$\begin{aligned} dX_t = f(t)X_t dt + g(t)dW, \quad X_t|X_0 \sim N(\gamma(t)X_0, \sigma^2(t)I) \\ X_t|y \sim N(\gamma(t)\mu_y, \gamma^2(t)\Sigma + \sigma^2(t)I) \\ \gamma(t) = exp(\int_0^t f(s)ds), \quad \sigma^2(t) = \gamma^2(t)\int_0^t \frac{g^2(s)}{\gamma^2(s)}ds \end{aligned} \tag{9}$$

reverse PFODE

$$\begin{aligned} \frac{dx}{dt} &= f(t)x - \frac{1}{2}g^2(t)\nabla_x logp_t(x) - \frac{1}{2}g^2(t)\omega\nabla_x logp_t(y|x) \\ &= \underbrace{fx - \frac{1}{2}g^2\nabla_x logp_t(x|y)}_{Conditional\,Drift} - \underbrace{\frac{1}{2}g^2(\omega-1)(\nabla_x logp_t(x|y) - \nabla_x logp_t(x))}_{Guidance\,Trem} \end{aligned}$$

where $dt < 0$.

To further analyze the impact of the guidance term on the reverse process, we eliminated the Conditional drift through a linear transformation.

$$Z = H(t)X + b(t), \quad G(X) = \nabla_x logp_t(X|y) - \nabla_x logp_t(X)$$

$$\nabla_x logp_t(x|y) = -\Sigma_t^{-1}(x - \mu_{t,y}), \quad \mu_{t,y} = \gamma(t)\mu_y, \quad \Sigma_t = \sigma^2(t)I + \gamma^2(t)\Sigma$$

$$\frac{dz}{dt} = (\dot{H} + fH + \frac{1}{2}g^2 H\Sigma_t^{-1})x - \frac{1}{2}g^2 H\Sigma_t^{-1}\mu_{t,y} + \dot{b} - \frac{1}{2}g^2 H(\omega-1)G(x)$$

Therefore we need to ensure

$$\dot{H} = -H(f + \frac{1}{2}g^2\Sigma_t^{-1}), \quad \dot{b} = \frac{1}{2}g^2 H\Sigma_t^{-1}\mu_{t,y}$$

We orthogonally diagonalize the Σ, and since the eigenvectors do not change over time during the forward noise addition process of the diffusion model, we can find a particular solution under this basis.

$$\Sigma = Q\Lambda Q^T, \quad where\ \Lambda = diag(\lambda_1, ..., \lambda_d)$$

$$\Sigma_t = Q\Lambda_t Q^T, \quad where\ \Lambda_t = diag(\lambda_1(t), ..., \lambda_d(t)), \quad \lambda_j(t) = \gamma^2\lambda_j + \sigma^2$$

$$H = QhQ^T, \quad \dot{h}_j(t) = -h_j(t)(f + \frac{g^2}{2\lambda_j(t)})$$

According to the definition, we have

$$\frac{d\lambda_j(t)}{dt} = 2f\lambda_j(t) + g^2$$

$$let \quad h_j = \lambda_j(t)^{-\frac{1}{2}}, \quad \dot{h_j} = -\frac{1}{2\lambda_j(t)}(2f\lambda_j(t) + g^2) * h_j$$

$$let \quad \dot{b}_j = \frac{1}{2}g^2\lambda_j^{-\frac{3}{2}}(\mu_{t,y})_j \Rightarrow b_j = -\lambda_j^{-\frac{1}{2}}(\mu_{t,y})_j$$

Which perfectly meets the requirements.

$$H = \Sigma_t^{-\frac{1}{2}}, \quad b = -\Sigma_t^{-\frac{1}{2}}\mu_{t,y}, \quad Z = \Sigma_t^{-\frac{1}{2}}(X - \mu_{t,y})$$

$$\frac{dz}{dt} = -\frac{1}{2}g^2H(\omega - 1)G(x), \quad where\ X = H^{-1}(Z - b) \tag{10}$$

Under this transformation, our target Gaussian distribution can be viewed as $N(0, I)$ throughout the entire reverse process, while the overall data distribution changes as follows:

$$X_t \sim \sum_i \pi_i N(\mu_{t,i}, \Sigma_t) \Rightarrow Z_t \sim \sum_i \pi_i N(\mu'_{t,i}, \Sigma'_t)$$

$$\mu'_{t,i} = \Sigma_t^{-\frac{1}{2}}(\mu_{t,i} - \mu_{t,y}), \quad \mu'_{t,y} = 0, \quad \Sigma'_t = \Sigma_t^{-\frac{1}{2}}\Sigma_t\Sigma_t^{-\frac{1}{2}} = I$$

$$Z_t \sim \sum_i \pi_i N(\mu'_{t,i}, I)$$

The transformation also has a very strong interpretable meaning; it turns the Gaussian function at time t white during the inverse process, transforms its elliptical profile into a sphere, and reduces the Mahalanobis distance to the Euclidean norm.

Then we will show the $G(X)$ parameters in the form of Z.

$$G(x) = \nabla_x log p_t(x|y) - \nabla_x log p_t(x) = -\Sigma_t^{-\frac{1}{2}} \sum_i p_i \mu'_{k,i}$$

$$where \quad p_i = \frac{\pi_i exp(-(x - \mu_{t,i})^T\Sigma_t^{-1}(x - \mu_{t,i})/2)}{\sum_j \pi_j exp(-(x - \mu_{t,j})^T\Sigma_t^{-1}(x - \mu_{t,j})/2)} = \frac{\pi_i exp(-\|z - \mu'_{t,i}\|^2/2)}{\sum_j \pi_j exp(-\|z - \mu'_{t,j}\|^2/2)}$$

Bring back to 10, we have

$$\frac{dz}{dt} = \frac{1}{2}g^2\Sigma_t^{-1}(\omega - 1)\sum_i p_i \mu'_{t,i}$$

We show that the remaining part of $\hat{G}(z,t) = -\sum p_i\mu'_{t,i}$ depends entirely on the distribution parameters after transformation. At this point, we have successfully extracted the influencing factors of guidance:

- The influence of different SDEs affect the preceding coefficients g^2.
- Covariance affects the parameters of the preceding linear transformation $H\Sigma_t^{-\frac{1}{2}} = \Sigma_t^{-1}$.
- The positions of different Gaussians relative to target Gaussian (the influence of the absolute position of target Gaussian is eliminated)

We focus on the dependence of $\partial Z/\partial t$ on Z over the full domain R^d. Therefore, we focus on $\hat{G}(z,t)$, the last of the three influences mentioned above.

Specifically, examining the weight p_i within $\hat{G}(z,t) = -\sum_i p_i \mu'_{t,i}$, expanding the quadratic term $\|z - \mu'_{t,i}\|^2 = z^T z - 2\langle z, \mu'_{t,i}\rangle + \|\mu'_{t,i}\|^2$ reveals a key algebraic property: the $\|z\|^2$ term cancels out completely during the softmax normalization. Consequently, p_i—and thus the entire function $\hat{G}(z,t)$—does not depend on the absolute norm of Z in $\mathbb{R}^d$, but is entirely governed by the set of inner products $\{\langle z, \mu'_{t,i}\rangle\}$.

A.2 General Convergence Boundary via Collinear GMM

Scalar reduction under collinearity. Under Assumption 3.1, the transformed component means satisfy $\mu'_{t,i} = l'_{t,i}\hat{\mu}$, where $\|\hat{\mu}\| = 1$. The guidance-induced vector field can therefore be written as

$$Z_t \sim \sum_i \pi_i N(\mu'_{t,i}, I), \quad \mu'_{t,i} = \Sigma_t^{-1/2}(\mu_{t,i} - \mu_{t,y}) = \frac{\gamma(t)(l_i - l_y)}{\sqrt{\gamma(t)^2\lambda + \sigma(t)^2}}\hat{\mu} = l'_{t,i}\hat{\mu}$$

$$\frac{dz_t}{dt} = \frac{g^2(t)(\omega(t)-1)}{2(\gamma^2(t)\lambda + \sigma^2(t))}\sum_i p_{t,i}(z_t)\mu'_{t,i}, \tag{11}$$

where

$$p_{t,i}(z) = \frac{\pi_i \exp\big(\langle z, \mu'_{t,i}\rangle - \frac{1}{2}\|\mu'_{t,i}\|^2\big)}{\sum_j \pi_j \exp\big(\langle z, \mu'_{t,j}\rangle - \frac{1}{2}\|\mu'_{t,j}\|^2\big)}. \tag{12}$$

Since $\langle z_t, \mu'_{t,i}\rangle = l'_{t,i}\langle z_t, \hat{\mu}\rangle$, the posterior weights depend on z_t only through the scalar projection $v_t = \langle z_t, \hat{\mu}\rangle$. Projecting the vector field onto $\hat{\mu}$ yields

$$\frac{dv_t}{dt} = \frac{g^2(t)(\omega(t)-1)}{2(\gamma^2(t)\lambda + \sigma^2(t))}G_t(v_t), \tag{13}$$

$$G_t(v) = \sum_i p_{t,i}(v) l'_{t,i}, \qquad p_{t,i}(v) = \frac{\pi_i \exp\big(l'_{t,i}v - \frac{1}{2}(l'_{t,i})^2\big)}{\sum_j \pi_j \exp\big(l'_{t,j}v - \frac{1}{2}(l'_{t,j})^2\big)}. \tag{14}$$

Thus, the state dependence of the guidance-induced mode competition is fully characterized by the scalar field $G_t(v)$.

Monotonicity and instantaneous equilibrium. For each fixed t, differentiating $G_t(v)$ gives

$$\frac{\partial G_t(v)}{\partial v} = \sum_i p_{t,i}(v)(l'_{t,i})^2 - \left(\sum_i p_{t,i}(v) l'_{t,i}\right)^2 = \mathrm{Var}_{p_t(v)}(l'_{t,i}) \geq 0.$$

Hence, $G_t(v)$ is nondecreasing in v. Moreover,

$$\lim_{v\to-\infty} G_t(v) = \min_i l'_{t,i}, \qquad \lim_{v\to+\infty} G_t(v) = \max_i l'_{t,i}. \tag{15}$$

the monotonic property of $G(v_t)$, as established by its non-negative derivative (the variance of $l'_{t,i}$), allows us to characterize the global convergence of the parameter v_t. Given the update rule $\frac{dv_t}{dt} = \frac{g^2(\omega-1)}{2(\gamma^2\lambda+\sigma^2)}G(v_t)$, we distinguish two primary regimes based on the support of the distribution $\{l'_{t,i}\}$.

- When the target distribution resides at the "edges" of the distributions, specifically when $\forall i, l_{t,i} \geq 0$ (or $\forall i, l_i \leq 0$), the function $G(v_t)$ becomes sign-definite. In this regime, v_t does not possess a stationary point in $\mathbb{R}$. Divergence Tendency with Velocity Control: In the case where $l_i \geq 0$ (Due to symmetry, we consider only one case) and the target resides at the boundary ($\min(l'_{t,i}) = 0$), the parameter v_t exhibits a directed drift toward $-\infty$. However, the dynamics are governed by a self-regulating "velocity control" mechanism. When $v_t > 0$,

the substantial magnitude of $G(v_t)$ provides a strong corrective force. As v_t enters the negative regime, the probability mass p_i increasingly shifts toward the target samples where $l'_{t,y} = 0$. Consequently, while the gradient remains positive, it asymptotically vanishes ($G(v_t) \to 0$), effectively decelerating the divergence and preventing a violent escape of the parameter.

- The sign-indefinite case is a more common scenario($\exists i, j$ s.t. $l'_{t,i} > 0 > l'_{t,j}$), the parameter v_t exhibits guaranteed convergence to a unique equilibrium. By the Intermediate Value Theorem and the strict monotonicity of $G(v_t)$, there exists a unique root $v_c(t)$ that serves as a globally stable attractor for the dynamical system. The stability is intrinsic: since $sgn(G(v_t)) = sgn(v_t - v_c(t))$, the update rule $\frac{dv_t}{dt}$ always acts as a restorative force toward $v_c(t)$, which means $(v_t - v_c(t))(v_{t-1} - v_t) < 0$.(In a symmetric GMM where $\forall i, \exists j \quad l'_{t,i} + l'_{t,j} = 0$, $v_c(t)$ always remains 0)

A.3 Discrete-time overshoot and the Pushing set.

Theorem 4.6. *(Proof in App. A.3) Under Assumptions 3.1 and an explicit Euler discretization with reverse-time step size $\delta_t > 0$, we derive a time-dependent sufficient upper bound $\omega_{\text{upper}}(t)$. Whenever guidance scales satisfying $0 \le \omega(t) - 1 \le \omega_{\text{upper}}(t)$, the corresponding scalar update Φ_t is non-expansive. The upper bound scale obeys*

$$\omega_{\text{upper}}(t) = \Theta\left(\frac{[\gamma^2(t)\lambda + \sigma^2(t)]^2}{g^2(t)\gamma^2(t)\delta_t}\right). \tag{4}$$

up to constants independent of t. $g(t)$, $\gamma(t)$, and $\sigma(t)$ denote the corresponding diffusion coefficients under the transformed coordinate system(Section 3).

Proof. In practice, the reverse process is simulated using discrete updates with a finite step size. Let $\delta_t = -dt > 0$ denote the reverse-time step size. Applying an explicit Euler step to Eq. (13) yields

$$v_{t-\delta_t} = \Phi_t(v_t), \qquad \Phi_t(v) = v - \frac{g^2(t)(\omega(t)-1)\delta_t}{2(\gamma^2(t)\lambda + \sigma^2(t))} G_t(v). \tag{16}$$

The restorative structure of the continuous reverse-time dynamics does not by itself guarantee stability under discrete sampling, as a sufficiently large update may overshoot the instantaneous equilibrium. To prevent such overshooting, we seek a sufficient condition under which the distance to the equilibrium does not increase after one update:

$$|\Phi_t(v) - v_c(t)| \le |v - v_c(t)|, \qquad \forall v \in \mathbb{R}. \tag{17}$$

Given that $\Phi_t(v_c(t)) = v_c(t)$ and the function is monotonic in v, $|\Phi'_t(v)| \le 1$ for all $v \in \mathbb{R}$ is a sufficient condition for Eq. (17). Using $G'_t(v) = \text{Var}_{p_t(v)}(l'_{t,i})$, we obtain

$$\Phi'_t(v) = 1 - \frac{g^2(t)(\omega(t)-1)\delta_t}{2(\gamma^2(t)\lambda + \sigma^2(t))} \text{Var}_{p_t(v)}(l'_{t,i}).$$

For $\omega(t) \ge 1$, the variance non-negativity implies $\Phi'_t(v) \le 1$. Hence, non-expansiveness reduces to the condition $\Phi'_t(v) \ge -1$, or equivalently,

$$\frac{g^2(t)(\omega(t)-1)\delta_t}{2(\gamma^2(t)\lambda + \sigma^2(t))} \text{Var}_{p_t(v)}(l'_{t,i}) \le 2, \qquad \forall v \in \mathbb{R}.$$

Then we make a relatively coarse upper bound estimate for $Var_p(l'_{t,i})$.

$$Var_p(l'_{t,i}) \le \frac{1}{4}(max_i\ l'_{t,i} - min_i\ l'_{t,i})^2 = \frac{\gamma^2}{4(\gamma^2\lambda + \sigma^2)}(max_i\ l_i - min_i\ l_i)^2$$

Substituting this uniform variance bound yields a sufficient condition. We define the resulting upper bound for the excess guidance $\omega(t) - 1$ as $\omega_{\text{upper}}(t)$:

$$\omega(t) - 1 \leq \omega_{\text{upper}}(t) := \frac{16\big(\gamma^2(t)\lambda + \sigma^2(t)\big)^2}{\gamma^2(t)g^2(t)\delta_t\big(\max_i l_i - \min_i l_i\big)^2}.$$

When the mixture parameters are independent of t, the temporal scaling of this upper bound is

$$\omega_{\text{upper}}(t) = \Theta\left(\frac{\big(\gamma^2(t)\lambda + \sigma^2(t)\big)^2}{\gamma^2(t)g^2(t)\delta_t}\right).$$

■

A.4 General temporal form of the sufficient upper bound.

The expression above gives a general temporal representation of the sufficient upper bound. To characterize its temporal trend, let

$$D_\lambda(t) := \gamma^2(t)\lambda + \sigma^2(t).$$

Using

$$\frac{\mathrm{d}}{\mathrm{d}t}\gamma^2(t) = 2f(t)\gamma^2(t), \qquad \frac{\mathrm{d}}{\mathrm{d}t}D_\lambda(t) = 2f(t)D_\lambda(t) + g^2(t),$$

we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\log\omega_{\text{upper}}(t) = 2f(t) + \frac{2g^2(t)}{D_\lambda(t)} - \frac{\mathrm{d}}{\mathrm{d}t}\log g^2(t) - \frac{\mathrm{d}}{\mathrm{d}t}\log\delta_t. \tag{18}$$

Since reverse-time sampling proceeds from $t = T$ to $t = 0$, a positive value of the derivative in Eq. (18) corresponds to a decreasing upper bound along the reverse trajectory. The sign of this derivative depends on the diffusion schedule, the discretization step, and the covariance parameter λ.

A.5 Rigorous Analysis for VP and VE.

Corollary 4.7. *(Proof in App. A.5) For a constant discretization step $\delta_t = \delta$, the sufficient stability boundary exhibits distinct asymptotic scalings under different diffusion parameterizations in the large-forward-time regime:*

- ***VP (constant β, Gao et al. [2025]):*** *Exponential scaling $\Theta(e^{At})$.*
- ***VP (linear β, Song et al. [2021]):*** *Exponential-quadratic scaling $\Theta(e^{B(t)}/\beta(t))$, where $B(t)$ is quadratic and $\beta(t)$ is linear in t.*
- ***VE ($\sigma(t) = \sqrt{t}$, Song et al. [2021], Karras et al. [2022]):*** *Polynomial scaling $\Theta(t^2)$.*
- ***EDM ($\sigma(t) = t$,Karras et al. [2022]):*** *Polynomial scaling $\Theta(t^3)$.*

Proof. **VPSDE.** For a VPSDE, as

$$g^2(t) = \beta(t), \qquad f(t) = -\frac{1}{2}\beta(t), \qquad \gamma^2(t) := \exp\left(-\int_0^t \beta(s)\,\mathrm{d}s\right), \qquad \sigma^2(t) = 1 - \gamma^2(t).$$

The sufficient upper bound therefore becomes

$$\omega_{\text{upper}}^{\text{VP}}(t) = \frac{16}{\Delta_\ell^2}\frac{\big[1 + (\lambda - 1)\gamma^2(t)\big]^2}{\beta(t)\gamma^2(t)\delta_t}.$$

constant noise rate. For a constant noise rate $\beta(t) = A$ and a constant discretization step $\delta_t = \delta$, we have $\gamma^2(t) = e^{-At}$ and

$$\omega_{\text{upper}}^{\text{VP}}(t) = \frac{16}{A\delta\Delta_\ell^2}\frac{\big[1 + (\lambda - 1)e^{-At}\big]^2}{e^{-At}}.$$

Its logarithmic derivative is

$$\frac{\mathrm{d}}{\mathrm{d}t}\log\omega_{\mathrm{upper}}^{\mathrm{VP}}(t) = A\frac{1-(\lambda-1)e^{-At}}{1+(\lambda-1)e^{-At}}.$$

Hence, for $\lambda \le 1$, the upper bound increases with the forward time t and therefore decreases along the reverse trajectory. For $\lambda > 1$, the same reverse-time decay holds on any interval satisfying

$$(\lambda-1)e^{-At} \le 1.$$

When $\lambda \le 2$, this condition holds over the full interval $t \ge 0$. For $\lambda > 2$, the upper bound may exhibit a turning point at

$$t_\star = \frac{1}{A}\log(\lambda-1).$$

For large forward time, the dominant behavior is

$$\omega_{\mathrm{upper}}^{\mathrm{VP}}(t) = \Theta(\frac{e^{At}}{A\delta}),$$

up to constants independent of t.

(linear noise schedule). We next consider a linearly increasing noise rate

$$\beta(t) = \beta_{\min} + \frac{\beta_{\max}-\beta_{\min}}{T}t = \beta_{\min} + \kappa t, \qquad \kappa := \frac{\beta_{\max}-\beta_{\min}}{T} > 0.$$

The corresponding integrated noise rate is

$$B(t) := \int_0^t \beta(s)\,\mathrm{d}s = \beta_{\min}t + \frac{\kappa}{2}t^2,$$

and hence

$$\gamma^2(t) = e^{-B(t)}.$$

Substituting this expression into the general VP boundary with a constant discretization step $\delta_t = \delta$, we obtain

$$\omega_{\mathrm{upper}}^{\mathrm{VP-L}}(t) = \frac{16}{\Delta_\ell^2\delta}\frac{\left[1+(\lambda-1)e^{-B(t)}\right]^2}{(\beta_{\min}+\kappa t)e^{-B(t)}}.$$

Its logarithmic derivative is

$$\frac{\mathrm{d}}{\mathrm{d}t}\log\omega_{\mathrm{upper}}^{\mathrm{VP-L}}(t) = \beta(t)\frac{1-(\lambda-1)e^{-B(t)}}{1+(\lambda-1)e^{-B(t)}} - \frac{\kappa}{\beta(t)}.$$

Consequently, any turning point within the considered time horizon must satisfy

$$\beta^2(t)\left[1-(\lambda-1)e^{-B(t)}\right] = \kappa\left[1+(\lambda-1)e^{-B(t)}\right].$$

Unlike the constant-noise-rate case, this equation generally does not admit a simple closed-form solution. Therefore, the monotonicity over a finite time horizon depends on the parameters $\beta_{\min}$, $\beta_{\max}$, T, and λ.

For large forward time, we have $e^{-B(t)} \to 0$ and $\beta(t) \sim \kappa t$. Consequently,

$$\frac{\mathrm{d}}{\mathrm{d}t}\log\omega_{\mathrm{upper}}^{\mathrm{VP-L}}(t) \sim \beta(t) - \frac{\kappa}{\beta(t)} > 0.$$

Thus, the boundary is eventually increasing in forward time and decreasing along the reverse trajectory. The corresponding time asymptotic behavior is

$$\omega_{\mathrm{upper}}^{\mathrm{VP-L}}(t) \sim \frac{16}{\Delta_\ell^2\delta}\frac{e^{B(t)}}{\beta(t)} = \Theta\left(\frac{e^{\beta_{\min}t+\frac{\kappa}{2}t^2}}{t}\right),$$

up to constants independent of t.

VE-type diffusion schedules. For VE-type diffusion processes, $\gamma(t) = 1$, and therefore

$$\sigma^2(t) = \int_0^t g^2(s)\,\mathrm{d}s.$$

The sufficient upper bound reduces to

$$\omega_{\mathrm{upper}}^{\mathrm{VE}}(t) = \frac{16}{\Delta_\ell^2}\frac{\left(\lambda + \sigma^2(t)\right)^2}{g^2(t)\delta_t}. \tag{19}$$

Its temporal derivative satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\log\omega_{\mathrm{upper}}^{\mathrm{VE}}(t) = \frac{2g^2(t)}{\lambda + \sigma^2(t)} - \frac{\mathrm{d}}{\mathrm{d}t}\log g^2(t) - \frac{\mathrm{d}}{\mathrm{d}t}\log\delta_t. \tag{20}$$

The temporal behavior therefore depends on the particular VE parameterization.

EDM-type parameterization. For the EDM-type parameterization

$$\gamma(t) = 1, \qquad \sigma^2(t) = t^2, \qquad g^2(t) = 2t,$$

and a constant discretization step $\delta_t = \delta$, we obtain

$$\omega_{\mathrm{upper}}^{\mathrm{EDM}}(t) = \frac{8}{\Delta_\ell^2\delta}\frac{(\lambda + t^2)^2}{t}. \tag{21}$$

Its logarithmic derivative is

$$\frac{\mathrm{d}}{\mathrm{d}t}\log\omega_{\mathrm{upper}}^{\mathrm{EDM}}(t) = \frac{3t^2 - \lambda}{t(\lambda + t^2)}. \tag{22}$$

Consequently, the bound exhibits a turning point at

$$t_\star = \sqrt{\lambda/3}.$$

For large forward time, the dominant behavior is

$$\omega_{\mathrm{upper}}^{\mathrm{EDM}}(t) = \Theta(t^3),$$

up to constants independent of t.

SMLD-type parameterization. For the SMLD-type parameterization

$$\gamma(t) = 1, \qquad \sigma^2(t) = t, \qquad g^2(t) = 1,$$

we obtain

$$\omega_{\mathrm{upper}}^{\mathrm{SMLD}}(t) = \frac{16}{\Delta_\ell^2\delta}(\lambda + t)^2. \tag{23}$$

Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}t}\log\omega_{\mathrm{upper}}^{\mathrm{SMLD}}(t) = \frac{2}{\lambda + t} > 0. \tag{24}$$

Therefore, under a constant discretization step, the sufficient upper bound decreases monotonically along the reverse trajectory.

For large forward time, the dominant behavior is

$$\omega_{\mathrm{upper}}^{\mathrm{SMLD}}(t) = \Theta(t^2),$$

up to constants independent of t.

■

A.6 Detailed Analysis on Symmetry 3GMM

In this section, we provide a more concrete description of our proposed 3-regime based on a specific 3GMM.

A.6.1 Reference Formula

We firstly present some parameters and derivation formulas that will be used in the reverse process. $y \in \mathcal{Y} := \{-, 0, +\}$ denotes the class label. In VPSDE we take $\beta(t) = 2$

$$p(X) = \frac{1}{3}N(X; \mu_-, I) + \frac{1}{3}N(X; \mu_0, I) + \frac{1}{3}N(X; \mu_+, I)$$
$$\mu_+ = l\hat{\mu}, \quad \mu_{0(target)} = 0, \quad \mu_- = -l\hat{\mu} \quad \langle \hat{\mu}, \hat{\mu} \rangle = 1$$

Then the parameter changes and reverse sampling in the PFODE process are

$$\mu_{t,y} = \mu_y e^{-t}, \quad x_{t-1} = x_t + \delta_t \Big(x_t + \nabla_x \log p(x_t, y = 0) + (\omega - 1)\nabla_x \log p(y = 0 \mid x_t) \Big)$$

We can further expand the closed expressions for $\nabla_x \log p(x_t, y = 0)$ and $\nabla_x \log p(y = 0|x_t)$.

$$\nabla_x \log p(x_t, y = 0) = -(x_t - \mu_{t,0}), \quad \nabla_X \log p(y = 0|x_t) = (\mu_{t,0} - \sum p(y|x_t)\mu_{t,y})$$

$$f(y) = \frac{1}{3}\, exp(\frac{-\|X - \mu_{t,y}\|^2}{2}), \quad p(y = 0|x) = f(0)/\sum f(y)$$

We set $e^{-t} = \theta_t$

$$x_{t-1} = x_t - \delta_t(\omega - 1)\theta_t l\hat{\mu} \frac{exp(\theta_t l\langle x_t, \hat{\mu}\rangle) - exp(-\theta_t l\langle x_t, \hat{\mu}\rangle)}{exp(\theta_t^2 l^2 \|\hat{\mu}\|^2/2) + exp(\theta_t l\langle x_t, \hat{\mu}\rangle) + exp(-\theta_t l\langle x_t, \hat{\mu}\rangle)} \tag{25}$$

$$l_t = \theta_t l, \quad v_t = \langle x_t, \hat{\mu}\rangle,$$

We multiply both sides of equation (25) by $\hat{\mu}$, simplify the vector x into the form of a scalar v_t, and simultaneously move a v_t to the left side of the equation.

$$v_{t-1} - v_t = -\delta_t(\omega - 1)l_t \frac{exp(l_t v_t) - exp(-l_t v_t)}{exp(l_t^2/2) + exp(l_t v_t) + exp(-l_t v_t)} \tag{26}$$

A.6.2 Analysis of v_t

We will find that the right side of the equation is an odd function of v_t, which is always smaller than 0 in the part $v_t > 0$, which means

$$v_t(v_{t-1} - v_t) < 0 \;\; v_t \neq 0$$

So if we want to compare $|v_t|$ and $|v_{t-1}|$ when $v_t \neq 0$, we only need to consider $h = (-v_{t-1}) - v_t$. Meanwhile, since v_{t-1} is also an odd function of v_t, the relationship between $|v_{t-1}|$ and $|v_t|$ is the same in cases $v_t > 0$ and $v_t < 0$. Therefore, we only need to consider case $v_t > 0$.

$$h(v_t, t) = \delta_t(\omega - 1)l_t \frac{exp(l_t v_t) - exp(-l_t v_t)}{exp(l_t^2/2) + exp(l_t v_t) + exp(-l_t v_t)} - 2v_t \tag{27}$$

when $h > 0$, $-v_{t-1} > v_t$, so $|v_{t-1}| > |v_t|$. when $h < 0$, $-v_t < -v_{t-1} < v_t$, so $|v_{t-1}| < |v_t|$. Therefore, we only need to focus on the properties of h.

$$\frac{\partial h}{\partial v_t} = \delta_t(\omega - 1)l_t^2 \frac{4 + exp(l_t^2/2)(exp(l_t v_t) + exp(-l_t v_t))}{(exp(l_t^2/2) + exp(l_t v_t) + exp(-l_t v_t))^2} - 2$$

we set $P = \delta_t(\omega - 1)l_t$, and $M_t = exp(l_t^2/2)$, and while separating variables, we only consider v_t (hereinafter written as x).

$$h(x) = P * \frac{2sinh(l_t x)}{M_t + 2cosh(l_t x)} - 2x, \quad \frac{dh(x)}{dx} = P * \frac{2l_t(M_t cosh(l_t x) + 2)}{(2cosh(l_t x) + M_t)^2} - 2$$

we can change $cosh(l_t x) + \frac{2}{M_t}$ to X,and $C = M_t - \frac{4}{M_t}$. x and X increase or decrease in the same manner when $x \geq 0$, only the rate of increase differ.

$$\frac{dh(x)}{dx} = H(X) = 2Pl_t M_t \frac{1}{4X + 4C + \frac{C^2}{X}} - 2$$

Here, the denominator exhibits the characteristics of a hook function. This implies that $H(x)$ has only three possible situations.

1. $H(X)$ Always less than 0 (no intersection with the x-axis)

2. $H(X)$ Initially greater than 0, subsequently less than 0 (intersecting the x-axis only once)

3. $H(X)$ Initially less than 0, then greater than 0, and finally less than 0 (intersecting the x-axis twice).

A.6.3 Situation Definitions

In situation 1, we have $h(x) \leq 0 \; when \; x > 0$ as $\frac{dh(x)}{dx} \leq 0$ and $h(0) = 0$, illustrating that $|v_{t-1}| \leq |v_t|$.

In situation 2, let the unique zero be x_0, then h increases in interval $(0, x_0)$ and decreases in interval (x_0, ∞). At the same time, as x approaches infinity, H approaches -2, therefore h must also intersect the x-axis, and this is the only zero point in the positive semi-axis. We set this zero point as b_t, so $|v_{t-1}| < |v_t|$ when $v_t > b_t$, and $|v_{t-1}| > |v_t|$ when $v_t \in (0, b_t)$

In situation 3, We set the two zero points as x_0, x_1, then h increases in interval (x_0, x_1) and decreases in interval $(0, x_0) \cup (x_1, \infty)$. This case requires further distinction, depending on whether $h(x_1) > 0$. if $h(x_1) \leq 0$, then $h \leq 0$, same as Situation 1. If $h(x_1) > 0$, then h have two zero point in the positive semi-axis as a_t, b_t. So $|v_{t-1}| < |v_t|$ when $v_t \in (0, a_t) \cup (b_t, \infty)$, and $|v_{t-1}| > |v_t|$ when $v_t \in (a_t, b_t)$

Here we must clarify: Situation 1 corresponds to the convergent regime, Situation 2 corresponds to the divergent regime, while Situation 3 actually encompasses the possibilities of both the mixed regime and convergence regime.

A.6.4 Threshold Functions Analysis

We now discuss the constraints corresponding to the three situations. As $X_{min} = 1 + \frac{2}{M_t}$, and $X_{opt} = argmax_X H(X) = max(X_{min}, |C/2|)$, then we can list the conditions for the three situations.

- $H(X_{opt}) \leq 0$;
- $H(X_{min}) \geq 0 \quad and \quad H(X_{opt}) > 0$
- $H(X_{opt}) > 0 \quad and \quad H(X_{min}) < 0$

In situation 1, $H < 0$, $h(0, k) = 0$, so $h < 0$ when $v_k > 0$

$C = M_t - \frac{4}{M_t}$. We first consider $C > 0, \; M_t > 2, when \; X_{min} = \frac{C}{2}, M_t = 4$, so

$$M_t \geq 4, \; X_{opt} = C/2$$

$$4 > M_t > 2, \; X_{opt} = X_{min}$$

and when $C < 0$, $0 < M_t \leq 2$, $X_{min} > -\frac{C}{2}$, so $X_{opt} = X_{min}$. Finally we have

$$M_t \geq 4, \; X_{opt} = C/2$$

$$4 > M_t > 0, \; X_{opt} = X_{min}$$

Then we consider each case separately. When $4 > M_t > 0$, only case 1 and 2 are possible.

$$H(X_{opt}) = 2Pl_tM_t\frac{1}{2M_t + M_t^2} - 2 = 2B\frac{1}{M_t + 2} - 2$$

where $B = l_t^2(\omega - 1)\delta_t$ is constant for sampling phase, and is mainly controlled by Sampling Settings Parameters.

Situation 1(Convergent Regime): $H_{max} \leq 0$

$$B \leq M_t + 2$$

Situation 2(Divergent Regime): $H(X_{min}) \geq 0 \; and \; H(X_{opt}) > 0$

$$B > M_t + 2$$

When $M_t \geq 4$

$$H(X_{opt}) = \frac{BM_t^2}{4(M_t^2 - 4)} - 2$$

$$H(X_{min}) = 2B\frac{1}{M_t + 2} - 2$$

Situation 1(Convergent Regime): $H_{max} \leq 0$

$$B \leq 8 - \frac{32}{M_t^2}$$

Situation 2(Divergent Regime): $.H(X_{min}) \geq 0 \; and \; H(X_{opt}) > 0$

$$B \geq M_t + 2$$

$$B > 8 - \frac{32}{M_t^2}(The \; scope \; is \; almost \; encompassed \; by \; the \; former.)$$

Situation 3(Convergent Regime + Mixed Regime): $H(X_{opt}) > 0 \; and \; H(X_{min}) < 0$

$$B < M_t + 2$$

$$B > 8 - \frac{32}{M_t^2}$$

Then we re-expand B, and propose $\omega - 1$, we can summarize the constraints of the three situations.

Situation 1(Convergent Regime):

when $M_t < 4$:

$$\omega - 1 \leq \frac{M_t + 2}{\delta_t l_t^2}$$

when $M_t \geq 4$:

$$\omega - 1 \leq \frac{8(1 - 4/M_t^2)}{\delta_t l_t^2}$$

Situation 2(Divergent Regime):

when $M_t < 4$:

$$\omega - 1 > \frac{M_t + 2}{\delta_t l_t^2}$$

when $M_t \geq 4$:

$$\omega - 1 \geq \frac{M_t + 2}{\delta_t l_t^2}$$

Situation 3(Convergent Regime + Mixed Regime):

only $M_t \geq 4$

$$\frac{M_t + 2}{\delta_t l_t^2} > \omega - 1 > \frac{8(1 - 4/M_t^2)}{\delta_t l_t^2}$$

A.6.5 Boundary Time-varying Trend Analysis

We first consider the expression for g. Expanding the expression of M_t, we have that

$$g(l, \delta_t, t) = \frac{e^{l_t^2/2} + 2}{\delta_t l_t^2}, \quad l_t = \theta_t l = e^{-t} l$$

We focus on analyzing the boundary change trend under the convergence regime, that is, under the $M_t = e^{l_t^2/2} < 4$ case.

$$g \approx \frac{(1 + \frac{l_t^2}{2}) + 2}{\delta_t l_t^2} = \frac{3}{\delta_t l_t^2} + \frac{1}{2\delta_t}$$

Here we assume the sampling step size δ_t is constant with respect to k, therefore the dynamic change of g is approximately equal to:

$$g \approx \frac{3}{\delta_t l^2 e^{-2t}} \propto e^{2t} \delta_t^{-1}$$

As for f, we have that

$$f(l, \delta_t, t) = 8 \frac{1 - 4e^{-l_t^2}}{l_t^2 \delta_t}$$

For the function f, we only consider the case where $M_t > 4$.

$$f \approx \frac{8}{\delta_t (l^2 e^{-2t})} = \frac{8}{\delta_t l^2} \cdot e^{2t} \propto e^{2t} \delta_t^{-1}$$

This confirms our subsequent experiments (C) and past sampling experience: sampling with larger step sizes tends to be more unstable, which conforms to our conclusions $O(e^{At} \delta_t^{-1})$.

B Experimental Settings

B.1 Unified setup for low-dimensional 3GMM experiments

To validate the theoretical regime transition map in a controlled setting, we conduct experiments on an one-dimensional Gaussian Mixture Model (3GMM) consisting of three components. The data distribution is defined as $p_{data}(x) = \frac{1}{3} \sum_{i=1}^{3} \mathcal{N}(x; \mu_i, \sigma^2)$, where the means are symmetrically placed at $\mu \in \{-5, 0, 5\}$ with a shared standard deviation $\sigma = 1$. We parameterize the diffusion process with a total time horizon $T = 10$ and $N = 51$ discretization steps. To ensure statistical robustness and accurately capture the distribution shift, we generate $n = 10,000$ independent samples (trajectories) for each configuration. This large-scale batch sampling allows for a precise estimation of the Confidence Evolution and the empirical boundary of the mixed regime.

B.2 Regime Visualizing Details

For the regime map visualizations in Figure 3, we use the discrete values of $f(\cdot)$ and $g(\cdot)$ mentioned in Equation (6) to compute the boundaries of situations convergence, divergence and mixed. As we divide the one step sampling status into three regimes in the main theory, h cannot be expressed in a closed form. For rigor, we use the situation divided by f and g instead of the regime for labeling. Situation 1(convergence) corresponds to the convergent regime, Situation 2(divergence) corresponds to the divergent regime, while Situation 3(mixed) actually encompasses the possibilities of both the mixed regime and convergence regime. (See details in Appendix A.6.3). We will also use this method in subsequent plotting.

B.3 Guidance Schedule and Comparison Strategies

Theory-inspired schedule. As analyzed in App. A.5, under the constant-β VPSDE, the sufficient stability boundary exhibits an exponential dependence on the forward diffusion time. In particular, for the unit-variance setting ($\lambda = 1$), the boundary reduces to

$$\omega_{\text{upper}}^{\text{VP}}(t) = \frac{16}{A \delta \Delta_\ell^2} e^{At}. \tag{28}$$

We consider the standard OU normalization with $A = 2$. To translate the temporal trend of the boundary into a practical guidance schedule, we normalize its shifted exponential profile to the interval $[0, 1]$:

$$k(t) = \frac{\omega_{\text{upper}}^{\text{VP}}(t) - \omega_{\text{upper}}^{\text{VP}}(0)}{\omega_{\text{upper}}^{\text{VP}}(T) - \omega_{\text{upper}}^{\text{VP}}(0)} = \frac{e^{At} - 1}{e^{AT} - 1}. \tag{29}$$

Setting $\kappa = AT$, the resulting guidance schedule is

$$\omega_{\text{ours}}(t) = 1 + \omega_{\text{base}} \frac{\exp(\kappa(t/T)) - 1}{\exp(\kappa) - 1}, \qquad t : T \to 0^+, \tag{30}$$

where $\omega_{\text{base}} = \omega_{\text{start}} - 1$ denotes the maximum excess guidance scale. This schedule satisfies $\omega_{\text{ours}}(T) = 1 + \omega_{\text{base}}$ and $\omega_{\text{ours}}(0) = 1$.

Comparison strategies. For a fair comparison of different temporal strategies, we match their maximum guidance scales. Specifically, all strategies use the same ω_{base}, while differing in how the guidance is distributed along the reverse trajectory.

Constant strategy:

$$\omega_{\text{const}}(t) - 1 = \omega_{\text{base}}.$$

Truncation strategy:

$$\omega_{\text{trunc}}(t) - 1 = \begin{cases} \omega_{\text{base}}, & t \geq 0.4T, \\ 0, & t < 0.4T. \end{cases}$$

Since sampling proceeds along the reverse trajectory from $t = T$ to $t = 0$, the truncation strategy disables guidance during the final 40% of sampling steps.

All strategies are matched by their maximum guidance scale, with $\max_t \omega(t) = 1 + \omega_{\text{base}}$. The comparisons therefore control for peak guidance strength, while allowing the cumulative guidance exposure to differ across strategies.

B.4 Parameter Setup and Evaluation Metrics

B.4.1 Common Stability Metrics

These metrics are used to quantify the numerical stability and artifact levels during the generation process.

Out-of-Bounds Rate (OBR). OBR measures the frequency of extreme pixel values that indicate numerical divergence or severe clipping. We define OBR as the proportion of pixels in an image that fall into the extreme intensity ranges (near-zero or near-saturation):

$$\text{OBR} = \frac{1}{H \times W \times C} \sum_{i,j,c} 1(x_{i,j,c} < 5 \text{ or } x_{i,j,c} > 250)$$

where $1(\cdot)$ is the indicator function. A high OBR typically correlates with "color-crushing" and loss of structural detail.

Saturation. To evaluate the perceptual fidelity and prevent "color-crushing" artifacts, we measure the color saturation of generated images. For each pixel, the saturation S is calculated based on the distance between the maximum and minimum color channel intensities relative to the brightness:

$$S = \begin{cases} \frac{\max(C) - \min(C)}{\max(C)}, & \text{if } \max(C) > 0 \\ 0, & \text{otherwise} \end{cases}$$

where $C = \{R, G, B\}$ denotes the color channels. In our analysis, we report the average S across all pixels. An abnormal surge in this metric indicates that the guidance scale is forcing the model to produce overly-simplistic, "neon-like" colors that deviate from the natural data distribution.

B.4.2 Model-Specific Metrics

Classifier-based Confidence (for CIFAR-10) To quantify the semantic alignment of generated samples on CIFAR-10, we employ a pre-trained **ResNet-32** classifier.

- Architecture: The model follows the standard ResNet design for CIFAR, consisting of three residual stages with filter sizes of $\{16, 32, 64\}$ and $n = 5$ blocks per stage, totaling 32 weighted layers.
- Training and Performance: The classifier was trained for 200 epochs using SGD with momentum (0.9), weight decay ($1e-4$), and a multi-step learning rate schedule. It achieves a top-1 accuracy of 93.4%

CLIP Score (for Stable Diffusion) To evaluate the semantic alignment between the generated images and the input text prompts, we calculate the CLIP Score. Specifically, we employ the pre-trained CLIP model with a ViT-B/32 backbone (openai/clip-vit-base-patch32). The score is defined as the cosine similarity between the visual embedding of the generated image and the textual embedding of the prompt, scaled by a factor of 100 to enhance interpretability. This metric serves as a proxy for the model's instruction-following capability.

B.5 Model Licenses and Citations

We use the following pretrained models:

- **Stable Diffusion v1.5** (runwayml/stable-diffusion-v1-5): Licensed under the CreativeML OpenRAIL-M License. Model available at `https://huggingface.co/sd-legacy/stable-diffusion-v1-5`. The underlying diffusion model architecture is described in [Rombach et al., 2022].
- **Stable Diffusion XL Base 1.0** (stabilityai/stable-diffusion-xl-base-1.0): Licensed under the CreativeML Open RAIL++-M License. Model available at `https://huggingface.co/stabilityai/stable-diffusion-xl-base-1.0`. See [Podell et al., 2024] for model details.
- **CLIP ViT-B/32** (openai/clip-vit-base-patch32): Licensed under the MIT License. Model available at `https://huggingface.co/openai/clip-vit-base-patch32`. Original paper: [Radford et al., 2021].

C More Visual Results

C.1 SDXL Experiments

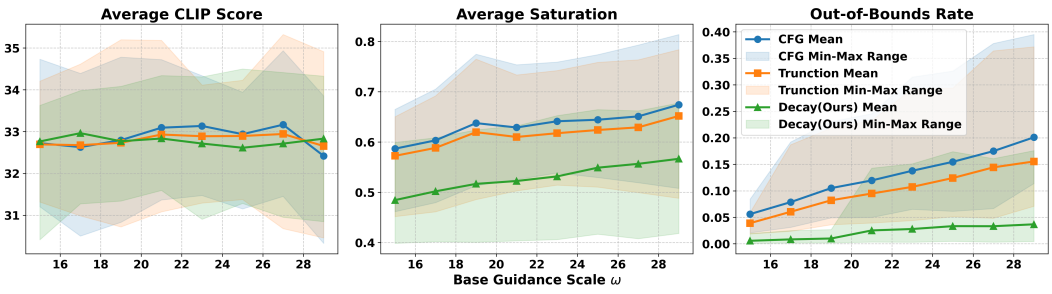

Figure 8: Comparison of guidance strategies. Metrics are evaluated over 50 random seed over same prompt, using 50-step inference.

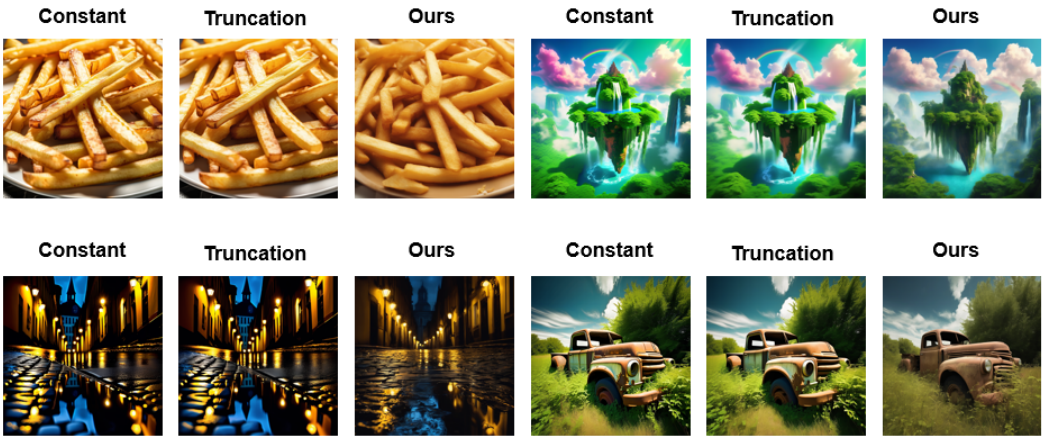

Figure 9: Additional SDXL sampling results across different prompts.

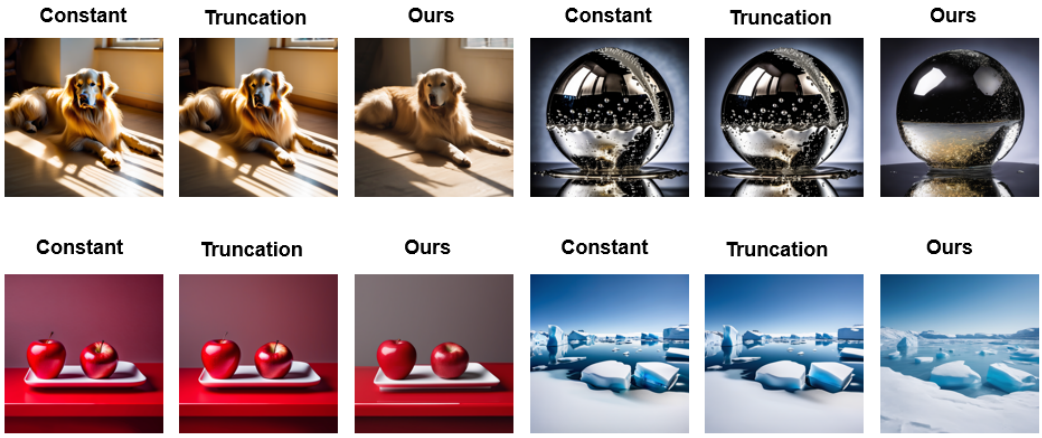

Figure 10: Additional SDXL sampling results across different prompts.

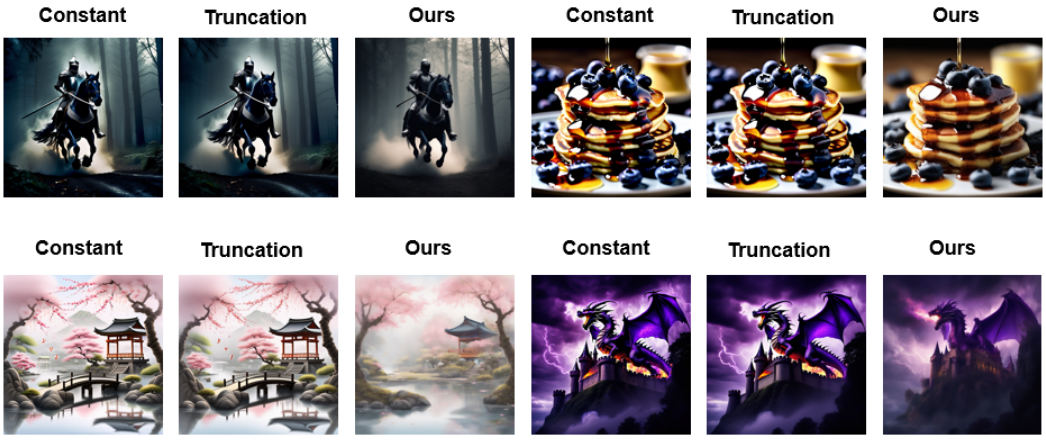

Figure 11: Additional SDXL sampling results across different prompts.

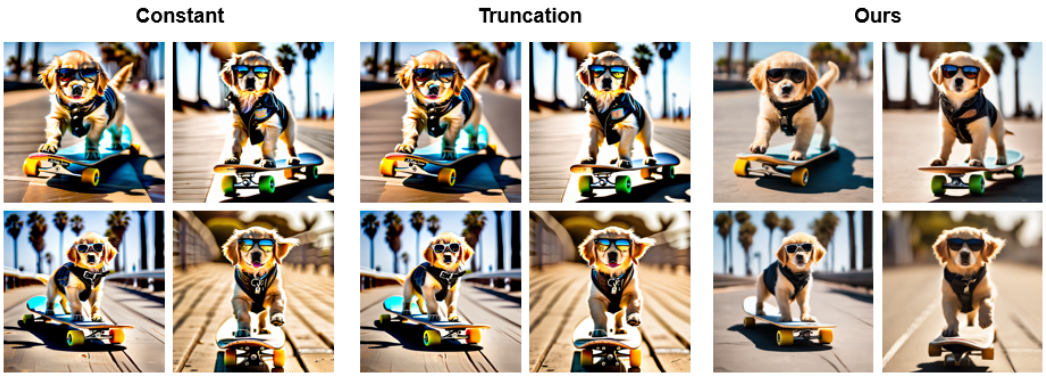

Figure 12: Additional SDXL sampling results across different random seeds.

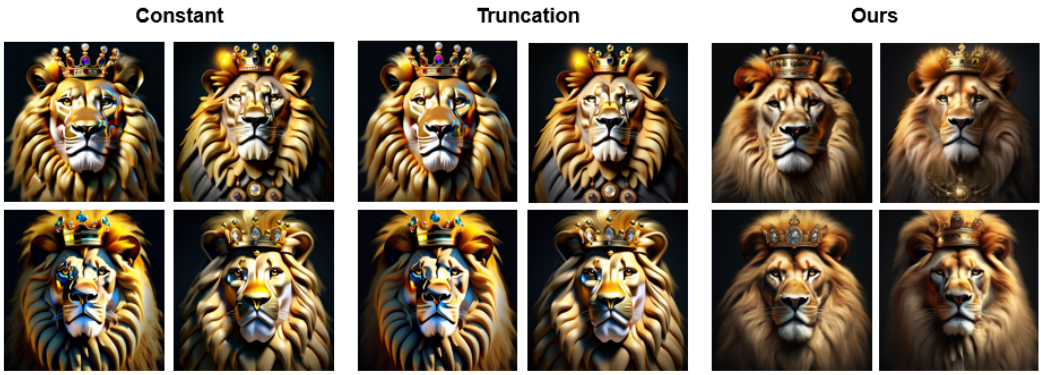

Figure 13: Additional SDXL sampling results across different random seeds.



Figure 14: Additional SDXL sampling results across different random seeds.

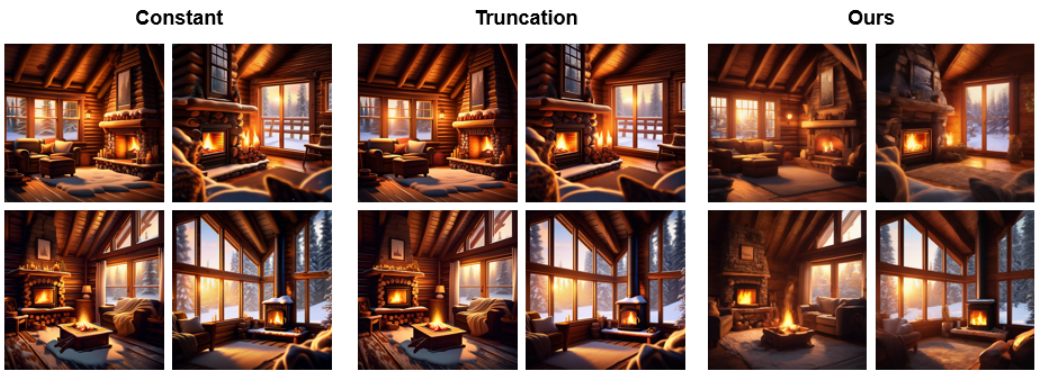

Figure 15: Additional SDXL sampling results across different random seeds.

C.2 SD1.5 Experiments

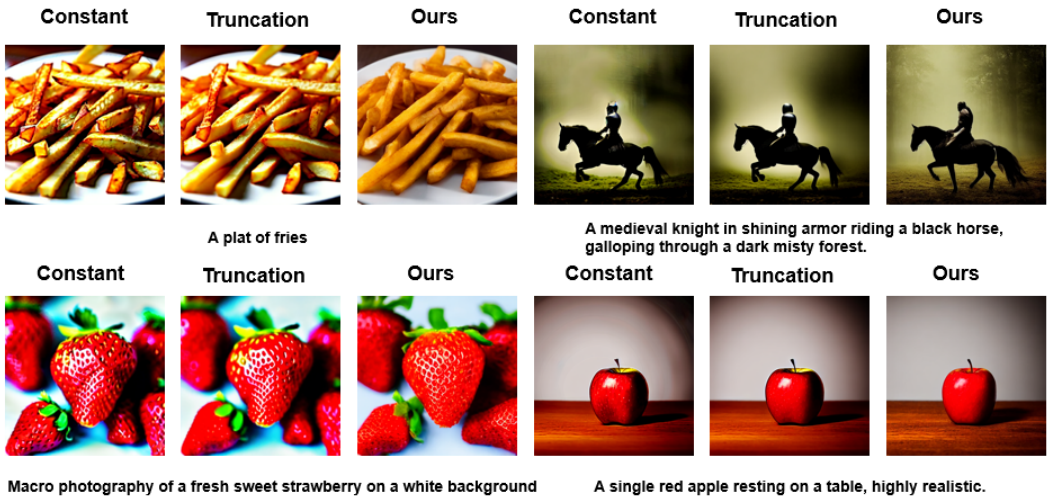

Figure 16: Showcasing the compatibility of our schedule using SD1.5. Compared to CFG, it is also better than Guidance Truncation at eliminating artifacts.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: This work studies the stability of guidance in diffusion sampling. We derive a time-dependent stability boundary in a simplified GMM setting and investigate its behavior under different diffusion parameterizations. Motivated by these findings, we explore a practical guidance schedule and evaluate its effectiveness through synthetic experiments and real-world image generation tasks.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We discuss future work and limitation at Section 6

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: We have shown all assumptions, theorem and proof sketch in the main paper. The detailed proof appears in the appendix.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: The experimental settings are documented throughout the appendix. Appendix Section B.1 specifies the GMM distributions, diffusion time horizon, discretization steps, and sampling configurations. Appendix Section B.3 provides the hyperparameter settings for the dynamic sampling experiments. Appendix Section B.4 describes the evaluation metrics and model-specific evaluation procedures, while Appendix Section B.5 identifies the pretrained models and their corresponding sources and licenses. Additional image-generation results are provided in Appendix C.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [No]

   Justification: This paper is primarily a theoretical contribution. The simulation experiments strictly follow the mathematical settings defined in the text. Our real-world experiments are intended as heuristic explorations to verify the theoretical validity of the derived boundary $\omega_{upper}(t)$, rather than as a refined, production-ready sampling algorithm. All necessary mathematical formulations for this schedule are fully disclosed, allowing for independent verification of our theoretical claims.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: We provide the experimental settings and evaluation details in Appendix.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

Justification: Figure 7 reports the mean and min-max range over 50 random prompts. Figure 8 reports the mean and min-max range over 50 random seeds for a fixed prompt.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: All experiments were conducted on consumer-grade hardware. Training on CIFAR-10 required approximately 40 GPU hours ($4 \times$ NVIDIA RTX 3090 for 10 hours), while all SDXL inference and SD-v1.5 evaluations were performed on a single NVIDIA RTX 3090.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: We have checked the code of ethics and ensured that our work satisfies the code of ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.

- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: We have discussed the broader impacts of our work at the end of main paper.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
    - If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: Since this work does not introduce or redistribute new models or datasets, additional release safeguards are not applicable.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

Justification: We use Stable Diffusion v1.5 (CreativeML OpenRAIL-M License), Stable Diffusion XL Base 1.0 (CreativeML Open RAIL++-M License), and CLIP ViT-B/32 (MIT License). Detailed license information, model URLs, and citations are provided in Appendix B.5. Original papers are cited accordingly in the references section.

Guidelines:

- The answer [N/A] means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [N/A]

Justification: This paper does not release new assets.

Guidelines:

- The answer [N/A] means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [N/A]

Justification: This paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [N/A]

Justification: This paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: This research does not involve LLMs as any important, original, or non-standard components.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.