EGCA: A Spectral Perspective on Forward Process Design in Diffusion Models

Chi Zhang [1], Guichao Chang[1], Sirui Liu [1], Xin Zhang [2], Guoqiang Zhong[3]*

[1] School of Mathematical Sciences, Ocean University of China [2] San Diego State University

[3] College of Computer Science and Technology, Ocean University of China

{zhangchi,changguichao,liusirui2883,gqzhong}@ouc.edu.cn; xzhang19@sdsu.edu

Abstract

The design of the forward diffusion process is a foundational yet under-theorized aspect of diffusion modeling. Although it determines how data are corrupted and how quickly the forward dynamics approach a reference distribution, forward-process design remains largely governed by canonical defaults rather than a unified analytical criterion. We propose Eigenvalue-Guided Control and Analysis (EGCA) for diffusion models, a spectral framework that uses the principal eigenvalue of the infinitesimal generator to analyze and guide forward-process design. For a structured class of admissible forward generators, this eigenvalue governs the exponential convergence rate to stationarity, providing a compact and interpretable summary of forward dynamics. We establish sufficient uniqueness and ergodicity conditions, derive eigenvalue-based convergence bounds, and adapt a practical numerical estimator for the principal eigenvalue. Experiments on multiple image datasets show that eigenvalue-guided generator choices can reduce training cost while maintaining or improving generation quality in the evaluated settings. Additional controlled experiments exhibit favorable within-pair trends with selected samplers, training formulations, and denoising architectures under the reported protocols. Beyond a single generator choice, our framework offers a spectral perspective for understanding and comparing forward-process choices in diffusion models.

1 Introduction

Recent advances in diffusion modeling have improved sampling efficiency, training stability, and sample quality through better reverse-time solvers, stronger architectures, and refined training objectives or parameterizations [8, 16, 26, 31, 36]. These developments largely assume that the forward process has already been specified. In contrast, we study the forward process itself.

The forward process plays a central role in diffusion models. It specifies how data are progressively corrupted [12, 37]. It also determines how quickly the system approaches its terminal distribution. In addition, it sets the temporal scale over which denoising must operate [16], and can therefore influence the effective computational budget under training or sampling protocols tied to the forward-process horizon.

In practice, forward-process design is still largely concentrated around canonical constructions, such as the discrete-time Gaussian noising process in DDPM [12] and VP/VE-type SDEs in score-based generative modeling [37]. Although later work has refined schedules and parameterizations [16, 18, 22, 29], forward-process design remains largely heuristic and fragmented. A unified criterion

*Corresponding author.

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

for comparing candidate forward generators and reasoning about their convergence behavior is still lacking.

This gap is consequential, as the lack of a principled understanding of forward-process behavior limits our ability to reason about convergence, compare different design choices, and systematically improve diffusion models. These observations raise a central question: *can forward-process design be organized around a computable quantity that captures convergence behavior and enables principled comparison across candidate generators?* In this work, we show that the spectral gap of the associated infinitesimal generator provides such a quantity. It directly characterizes the forward contraction rate in our theoretical setting and provides a convergence-oriented coordinate for comparing candidate generators.

This spectral perspective shifts the focus from raw coefficient parameterizations to intrinsic convergence behavior. Instead of reasoning only through drift and diffusion forms, we compare forward processes through a quantity that appears directly in convergence bounds. As a by-product, this spectral coordinate also enables retrospective analysis of existing designs. For example, under the normalization and default parameterization used in this paper, the standard VP-SDE associated with DDPM and SGM lies in a relatively low-spectral-gap regime.

Within our framework, the principal eigenvalue plays three roles. It provides a convergence-rate certificate through the exponential contraction bound, serves as a first-order descriptor for comparing admissible generators, and organizes the empirical efficiency–quality trade-off. We do not claim that the spectral gap independently determines sample quality; rather, it offers a compact and computable convergence-oriented coordinate for structuring forward-process design.

Our perspective is complementary to inference-stage acceleration methods, such as DDIM [36] and DPM-Solver-type samplers [26, 27], which improve reverse-time sampling under a fixed forward process. Training formulations and denoising architectures, such as EDM/EDM2-style training [16, 17] and diffusion transformers [31], act at different levels of the diffusion pipeline. In contrast, we focus on the forward process itself, aiming to characterize its convergence behavior and provide principled guidance for generator-level design. Subsection 4.6 reports controlled compatibility checks under specified paired protocols.

We also clarify the scope of our theory. Our convergence analysis applies to a high-dimensional forward process with a coordinate-wise separable generator, allowing the convergence behavior to be characterized through the principal eigenvalue of a shared one-dimensional coordinate generator. Thus, our results provide a tractable spectral theory for this structured high-dimensional setting, rather than a complete theory for arbitrary nonseparable or non-reversible diffusion dynamics. The experiments further show consistent within-pair trends across the evaluated image datasets and selected sampler, training-formulation, and architecture settings.

We instantiate this perspective in a framework termed Eigenvalue-Guided Control and Analysis (EGCA) for diffusion models. This framework combines spectral analysis, a practical numerical estimator for the principal eigenvalue, and empirical evaluation across multiple forward-process families. Experiments show that the spectral gap closely tracks forward convergence speed, and that eigenvalue-guided design can reduce training cost while maintaining competitive generation quality in the evaluated settings.

Our contributions are as follows:

- **A spectral framework for forward-process design.** We introduce EGCA, a framework that characterizes diffusion forward processes through the principal eigenvalue of the infinitesimal generator, providing a computable and interpretable coordinate for analyzing convergence behavior and organizing generator-level design.
- **Theoretical characterization of convergence.** Under suitable uniqueness and ergodicity conditions, we show that the principal eigenvalue governs the exponential convergence rate of the forward process, establishing a direct link between generator structure and convergence dynamics.
- **Practical estimation and control.** We adapt a practical numerical estimator for the principal eigenvalue and use it to guide forward-generator adjustment, making the spectral perspective computationally usable for generator selection and control.
- **Empirical evaluation across diffusion settings.** Across the evaluated image-generation settings, eigenvalue-guided generator choices exhibit favorable efficiency–quality trade-offs. Controlled

paired experiments further show consistent relative trends with selected samplers, EDM/EDM2-style training formulations, and a DiT denoiser under the reported protocols.

2 Preliminary

2.1 Score-Based Generative Model

In SGMs [37], the forward diffusion process progressively transforms an initial data sample $\boldsymbol{X}_0 \in \mathbb{R}^d$ into noise, such that the distribution of $\boldsymbol{X}_T$ approaches a reference distribution as $T \to \infty$. This evolution is described by the SDE

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{f}(\boldsymbol{X}_t, t)\,\mathrm{d}t + \boldsymbol{G}(\boldsymbol{X}_t, t)\,\mathrm{d}\boldsymbol{W}_t, \tag{1}$$

where $\boldsymbol{W}_t$ is a d-dimensional standard Wiener process, $\boldsymbol{f} : \mathbb{R}^d \times [0,T] \to \mathbb{R}^d$ is the drift field governing the deterministic component of the dynamics, and $\boldsymbol{G} : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{d\times d}$ is the diffusion coefficient controlling the injected noise.

Sample generation is then carried out by a reverse-time dynamics parameterized by the score of the perturbed data distribution. In this paper, we use this reverse-time view only as background and focus our theoretical analysis on the forward dynamics.

2.2 $\mathcal{L}_t$-generator

Beyond the SDE in Eq.(1), the forward diffusion process can be equivalently characterized by its infinitesimal generator $\mathcal{L}_t$. For any smooth, compactly supported test function $\phi : \mathbb{R}^d \to \mathbb{R}$, the generator is defined by $\mathcal{L}_t\phi(\boldsymbol{x}) = \lim_{h\to 0} \frac{\mathbb{E}[\phi(\boldsymbol{X}_{t+h})|\boldsymbol{X}_t=\boldsymbol{x}]-\phi(\boldsymbol{x})}{h}$, which captures the instantaneous evolution of the process [10, 30]. Applying Itô's formula to Eq.(1) yields the second-order differential form

$$\mathcal{L}_t\phi(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x},t)\cdot\nabla\phi(\boldsymbol{x}) + \frac{1}{2}\,\mathrm{Tr}\big(\boldsymbol{A}(\boldsymbol{x},t)\nabla^2\phi(\boldsymbol{x})\big),$$

where $\nabla\phi(\boldsymbol{x}) \in \mathbb{R}^d$ is the gradient, $\nabla^2\phi(\boldsymbol{x}) \in \mathbb{R}^{d\times d}$ is the Hessian, and $\boldsymbol{A}(\boldsymbol{x},t) = \boldsymbol{G}(\boldsymbol{x},t)\boldsymbol{G}(\boldsymbol{x},t)^\top$ [13, 30]. This representation makes explicit both the deterministic drift component (through $\boldsymbol{f}$) and the stochastic diffusion component (through $\boldsymbol{A}$). A special case arises when the coefficients are time-homogeneous, that is, $\boldsymbol{f}(\boldsymbol{x},t) \equiv \boldsymbol{b}(\boldsymbol{x})$, $\boldsymbol{A}(\boldsymbol{x},t) \equiv \boldsymbol{A}(\boldsymbol{x})$, in which case the generator becomes independent of t: $\mathcal{L}\phi(\boldsymbol{x}) = \boldsymbol{b}(\boldsymbol{x})\cdot\nabla\phi(\boldsymbol{x}) + \frac{1}{2}\,\mathrm{Tr}\big(\boldsymbol{A}(\boldsymbol{x})\nabla^2\phi(\boldsymbol{x})\big)$.

3 EGCA framework

This section develops EGCA as a spectral framework for forward-process design. We reduce the structured high-dimensional generator to a shared coordinate generator, connect its principal eigenvalue to convergence, and show how this quantity guides generator and schedule design.

3.1 Forward process structure and generator decomposition

For EGCA, the forward diffusion process $\boldsymbol{X}_t$ is designed to satisfy the SDE

$$\mathrm{d}\boldsymbol{X}_t = \beta(t)\boldsymbol{b}(\boldsymbol{X}_t)\,\mathrm{d}t + \sqrt{2\beta(t)}\,\boldsymbol{A}(\boldsymbol{X}_t)^{1/2}\,\mathrm{d}\boldsymbol{W}_t, \tag{2}$$

where $\boldsymbol{X}_t = (X_t^{(1)},\dots,X_t^{(d)})^\top \in \mathbb{R}^d$, $\boldsymbol{W}_t = (W_t^1, W_t^2, \dots, W_t^d)$ is a d-dimensional standard Brownian motion with independent components, $\beta(t)$ is a scalar time schedule, and

$$\boldsymbol{b}(\boldsymbol{x}) = \big(b(x_1),\dots,b(x_d)\big)^\top, \qquad \boldsymbol{A}(\boldsymbol{x}) = \mathrm{diag}\big(a(x_1),\dots,a(x_d)\big).$$

Consequently, each X_t^i $(i = 1,\dots,d)$ satisfies the identical one-dimensional SDE

$$\mathrm{d}X_t^{(i)} = \beta(t)b\big(X_t^{(i)}\big)\,\mathrm{d}t + \sqrt{2\beta(t)a\big(X_t^{(i)}\big)}\,\mathrm{d}W_t^{(i)}. \tag{3}$$

The associated generator of $\boldsymbol{X}_t$ is separable:

$$\mathcal{L}_t = \beta(t)\mathcal{L}_{\mathrm{sep}}, \quad \text{where } \mathcal{L}_{\mathrm{sep}} = \textstyle\sum_{i=1}^d \mathcal{L}^{(i)},\ \ \mathcal{L}^{(i)} = a(x_i)\frac{\partial^2}{\partial x_i^2} + b(x_i)\frac{\partial}{\partial x_i}.$$

We write $\mathcal{L}$ for the common one-dimensional coordinate generator. Once $\beta(t)$ is fixed, the high-dimensional dynamics are determined by the separable generator $\mathcal{L}_{\mathrm{sep}}$, and hence by the shared coordinate generator $\mathcal{L}$. Although (a, b) determines $\mathcal{L}$, inspecting the raw coefficient functions does not by itself yield a transparent convergence-rate criterion. By contrast, the principal eigenvalue λ_1 of $\mathcal{L}$ appears explicitly in the convergence bound derived below, making it a natural spectral coordinate for comparing candidate generators.

Therefore, EGCA treats forward-process design primarily from a generator-level perspective: (a, b) determines the underlying generator and its drift–diffusion structure, while $\beta(t)$ controls the operational time at which this generator is traversed. These two controls are related rather than disjoint, as uniform generator scaling is exactly equivalent to a corresponding scalar rescaling of $\beta(t)$. Indeed, for any $c > 0$, $\beta(t)(cL) = (c\beta(t))L$. Hence, uniform scaling of the generator lies in the overlap between generator-level and schedule-level control, whereas non-proportional changes of (a, b) generally cannot be reproduced by scalar schedule rescaling.

3.2 Ergodicity and spectral analysis of the forward process

A rigorous connection between the convergence rate of $\boldsymbol{X}_t$ and the spectral properties of $\mathcal{L}$ requires the forward SDE in Eq. (2) to be well defined and ergodic. The following assumption gives sufficient conditions on (a, b) and $\beta(t)$ for well-posedness, non-explosion, and existence of a finite stationary distribution.

Assumption 1 (Admissible regime)**.** Throughout, we assume

(A1) $a(x) > 0$ and $b(x)$ are both continuous functions on $\mathbb{R}$;

(A2) $\kappa(+\infty) = +\infty = \kappa(-\infty)$, where $\kappa(x) = \int_0^x e^{-C(u)} \int_0^u \frac{1}{a(v)} e^{C(v)} \,\mathrm{d}v \,\mathrm{d}u$ and $C(x) = \int_0^x \frac{b(u)}{a(u)} \,\mathrm{d}u$ for $x \in \mathbb{R}$;

(A3) $Z := \int_{\mathbb{R}} \frac{1}{a(u)} e^{C(u)} \,\mathrm{d}u < +\infty$;

(A4) $\beta : [0, \infty) \to (0, \infty)$ is continuous, and $\int_0^\infty \beta(t) \,\mathrm{d}t = +\infty$.

Theorem 2 (Well-posedness and ergodicity)**.** *Under Assumption (1), for every initial condition $\boldsymbol{X}_0 \in \mathbb{R}^d$, the forward SDE in Eq. (2) admits a non-explosive weak solution $\boldsymbol{X}_t$ that is unique in law. The stationary distribution is*

$$\Pi(d\boldsymbol{x}) = \bigotimes_{i=1}^d \pi(dx_i) \;\; \textit{with} \;\; \pi(dx) = \tfrac{1}{Za(x)} e^{C(x)} \, dx. \tag{4}$$

Let $L^2(\pi)$ be the space of real-valued measurable functions $f : \mathbb{R} \to \mathbb{R}$ such that $\pi(f^2) < \infty$, equipped with the norm $\|f\|_{L^2(\pi)} = \left[\pi(f^2)\right]^{1/2}$ and inner product $\langle f, g\rangle_{L^2(\pi)} = \int_{\mathbb{R}} f(x)g(x) \, \pi(dx)$, where $\pi(g) = \int_{\mathbb{R}} g(x)\pi(dx)$ for general g. The principal eigenvalue λ_1 of $\mathcal{L}$ is defined as

$$\lambda_1 = \inf \left\{ \langle f, -\mathcal{L}f\rangle_{L^2(\pi)} : f \in \mathcal{D}(\mathcal{L}), \pi(f) = 0, \|f\|_{L^2(\pi)} = 1 \right\}, \tag{5}$$

with $\mathcal{D}(\mathcal{L})$ the domain of $\mathcal{L}$ in $L^2(\pi)$. Since $\mathcal{L}$ has one trivial eigenvalue $\lambda_0 = 0$, the spectral gap $\lambda_1 - \lambda_0$ is equal to λ_1 [5].

Denote by μ_t the law of $\boldsymbol{X}_t$. Since the raw data distribution μ_0 may be empirical or otherwise singular with respect to the continuous reference measure Π, we formulate the contraction result from a time $\tau \geq 0$ for which $\mu_\tau \ll \Pi$ and $\chi^2(\mu_\tau \| \Pi) < \infty$. For empirical data distributions, such a condition is naturally considered after a positive smoothing time $\tau > 0$. The theorem below gives the resulting convergence bound.

Theorem 3 (Convergence rate)**.** *Under Assumption 1, if $\mu_\tau \ll \Pi$ and $\chi^2(\mu_\tau \| \Pi) < \infty$ for some $\tau \geq 0$, then*

$$\chi^2(\mu_t \| \Pi) \leq \chi^2(\mu_\tau \| \Pi) \exp\!\big\{-2\lambda_1 \textstyle\int_\tau^t \beta(s) \, ds\big\}, \quad \forall \, t \geq \tau, \tag{6}$$

where $\chi^2(\mu_\tau \| \Pi) = \int_{\mathbb{R}^d} (\frac{d\mu_\tau}{d\Pi} - 1)^2 d\Pi$. In particular, if $\mu_0 \ll \Pi$ and $\chi^2(\mu_0 \| \Pi) < \infty$, then taking $\tau = 0$ recovers the corresponding bounds from the initial distribution μ_0.

Interpretation. This theorem decomposes the discrepancy to the reference law into a dataset-dependent prefactor $\chi^2(\mu_\tau\|\Pi)$ and a dynamics-dependent contraction term governed by λ_1. Thus, larger λ_1 implies faster forward contraction, while different datasets may still induce different distance levels through the prefactor. This makes λ_1 the natural spectral quantity for comparing candidate generators. This stronger χ^2-based contraction immediately yields a more interpretable consequence below.

Corollary 4 (Total-variation convergence)**.** *Under the assumptions of Theorem 3, it holds that*

$$\|\mu_t - \Pi\|_{\mathrm{TV}} \le \tfrac{1}{2}\big[\chi^2(\mu_\tau\|\Pi)\big]^{\frac{1}{2}} \exp\big\{-\lambda_1 \textstyle\int_\tau^t \beta(s)\,ds\big\}, \quad \forall\, t \ge \tau,$$

where the total variation distance $\|\mu_t - \Pi\|_{\mathrm{TV}} = \sup_{B\in\mathcal{B}(\mathbb{R}^d)} |\mu_t(B) - \Pi(B)|$.

However, obtaining λ_1 with exact precision is generally impractical, which motivates the numerical estimation approach developed in the following subsection.

3.3 Numerical estimation of the principal eigenvalue

This subsection outlines an iterative algorithm for numerically estimating λ_1, based on Chen's theoretical estimation for this principal eigenvalue [4]. Given an initialization f_1 for any $x < y$, the estimator constructs a sequence of functions f_n and corresponding lower bounds

$$\widehat{\lambda}_1^{[n]} = \sup_{x<y}\left[\inf_{z\le\theta_n} \frac{f_{n-1}(z)}{f_n^-(z)} \wedge \inf_{z>\theta_n} \frac{f_{n-1}(z)}{f_n^+(z)}\right], \tag{7}$$

which converges monotonically to λ_1. We defer the full iterative construction to Appendix B.1.

Since the estimator is applied to the one-dimensional generator, its cost is independent of the ambient data dimension. For a fixed generator family, it is also a one-time computation rather than a recurring training cost. In practice, the improper integrals are approximated on truncated domains using the composite trapezoidal rule on a uniform grid with 5000 subintervals. The truncation radius is enlarged when necessary to check numerical stability, and the iterative procedure is run for at most 50 iterations. Moreover, the computation takes about 1 minute on dedicated training hardware in our implementation. This makes the estimator a lightweight tool for comparing forward-process designs and motivates the eigenvalue-guided $\mathcal{L}$-generator tuning developed next.

3.4 Eigenvalue-guided forward-process design

Eq. (7) shows that the spectral gap λ_1 depends on the coefficient pair (a, b) through the generator and its invariant distribution. This provides a direct mechanism for generator design: candidate coefficient functions can be compared and adjusted using λ_1 as a computable spectral coordinate. We focus on parameterized polynomial and related coefficient families because they are expressive, easy to control, and amenable to stable numerical evaluation. The selection of a suitable generator can then be guided by the following theoretical and empirical principles:

Principle I (Ergodic admissibility). As established in Subsection 3.2, a candidate coefficient pair (a, b) should satisfy conditions (A1)–(A3) of Assumption 1, ensuring a well-defined, non-explosive diffusion with a positive stationary distribution. Combined with the schedule condition (A4), the resulting forward process converges toward stationarity. Representative candidate functions and their admissibility checks are summarized in Table 5 of Appendix B.2.

Principle II (Generator scaling). Proposition 5 provides a concrete principle for the $\mathcal{L}$-generator design: scaling the coefficient pair (a, b) by a common positive factor scales the principal eigenvalue by the same factor, making eigenvalue adjustment analytically predictable and changing the exponential convergence rate in a controlled way.

Proposition 5. *Let $a^*(x) = ka(x)$ and $b^*(x) = kb(x)$, where k is a positive constant. Then the principal eigenvalue λ_1^* of $\mathcal{L}^* = a^*(x)\frac{d^2}{dx^2} + b^*(x)\frac{d}{dx}$ is equal to $k\lambda_1$.*

Observation (Spectral variation under coefficient changes). Using the numerical scheme in Subsection 3.3, we estimate the spectral gap λ_1 for a range of admissible coefficient configurations (a, b); see Appendix Table 9. These values provide a practical reference grid for eigenvalue-guided tuning and allow nearby coefficient configurations to be explored in a structured manner.

Complementary mechanism (Schedule modulation). Since the convergence bound depends on the accumulated spectral time $\lambda_1 \int_\tau^t \beta(s)\, ds$, rescaling $\beta(t)$ changes the rate at which a fixed generator is traversed, whereas modifying the generator can change its spectral properties and, more generally, its coefficient structure. As clarified in Subsection 3.1, uniform generator scaling is a special case that is exactly equivalent, at the level of the scheduled forward dynamics, to a corresponding scalar rescaling of $\beta(t)$. More general non-proportional changes of (a, b), however, cannot in general be reproduced by scalar schedule rescaling. We empirically examine schedule modulation in Subsection 4.4.

Eigenvalue-guided workflow. Given a candidate coefficient pair (a, b), EGCA first verifies admissibility and estimates its spectral gap λ_1. The generator can then be adjusted either through uniform scaling or through more general changes of its coefficient structure. Uniform scaling provides a simple and analytically predictable way to change the spectral rate, while non-proportional coefficient changes explore generator variations that cannot in general be represented by scalar schedule rescaling.

The schedule $\beta(t)$ provides an additional control of the accumulated spectral time. Because uniform generator scaling is exactly equivalent to a corresponding scalar schedule rescaling, these two operations should not be interpreted as independent effects in that special case. The favorable spectral regime is empirical and dataset-dependent; Section 4 evaluates how λ_1 organizes the observed convergence–quality trade-off. Thus, EGCA turns forward-process design from an unstructured search over coefficient functions into a procedure guided by a computable spectral quantity.

3.5 Spectral view of existing forward processes

The preceding subsections develop EGCA as a spectral framework for admissible forward-process design. We now use this lens to position representative existing formulations; this is a scope analysis rather than a performance comparison.

We focus on generator-level forward-process choices. VP-SDE-type forward processes used in DDPM [12] and SGM [37] fit the present ergodic generator setting and can be assigned a principal eigenvalue. Under default parameterizations, they lie in a relatively low-eigenvalue regime, consistent with the conservative large-step configurations often used in practice. EGCA-style generators instead modify the generator itself, allowing the process to move to a different spectral regime. In contrast, VE-SDE [37] lacks a confining drift, and CLD [9] requires an augmented kinetic-state analysis; both therefore require different spectral tools. Appendix Table 6 provides a concise comparison.

4 Experiments

We evaluate whether the principal eigenvalue λ_1 provides a useful spectral coordinate for forward-process design, measuring convergence speed, training cost, and generation quality across datasets and diffusion settings.

4.1 Experiment setup and evaluation metrics

Datasets and settings. The main generator-level comparisons use tractable linear generators and vary the coefficient pair (a, b) under fixed dataset, training protocol, and schedule family. We evaluate representative image-generation settings, including CIFAR-10 [19], CelebA-HQ-64 [15], ImageNet-128/256 [7]. The main comparisons use a U-Net denoising backbone with a DDPM-style linear noise schedule [12]. We also test compatibility with representative reverse-time samplers, EDM/EDM2-style training formulations, DiT-based generation, and flow-based boundary comparisons. Additional generator-family diagnostics are provided in Appendix F. All main U-Net experiments are conducted on an NVIDIA GeForce RTX 4090 GPU.

Metrics. We employ four complementary metrics: *1) Fréchet Inception Distance (FID).* It measures the distributional discrepancy between real and generated images; lower values indicate better sample quality [11]. *2) Convergence Discrepancy (D_{disc}).* D_{disc} quantifies the proximity between the forward distribution μ_t and the stationary distribution Π. It serves as the guiding criterion for determining the required diffusion length T_{conv}. *3) Step Count T_{conv}.* T_{conv} denotes the number of noise-injection steps

Table 1: Main efficiency–quality results under low, moderate, and aggressive spectral regimes.

Dataset	Generator	Regime	λ_1	T_{conv}	T_{spend}	FID↓ (SDE solver)
CIFAR-10	$a = 1/2,\ b = -0.5x$	VP-type baseline	0.50	1000	45h	4.44
CIFAR-10	$a = 1/2,\ b = -x$	moderate EGCA	1.00	825	35h	3.15
CIFAR-10	$a = 1/2,\ b = -2x$	moderate EGCA	2.00	750	30h	3.19
CIFAR-10	$a = 1/2,\ b = -10x$	aggressive EGCA	10.00	350	20h	6.33
CelebA-HQ-64	$a = 1/2,\ b = -0.5x$	VP-type baseline	0.50	1000	104h	3.82
CelebA-HQ-64	$a = 1/2,\ b = -x$	moderate EGCA	1.00	800	67h	3.34
CelebA-HQ-64	$a = 1/2,\ b = -2x$	moderate EGCA	2.00	750	61h	3.67
CelebA-HQ-64	$a = 1/2,\ b = -10x$	aggressive EGCA	10.00	300	45h	7.41
ImageNet-128	$a = 1/2,\ b = -0.5x$	VP-type baseline	0.50	1000	182h	35.24
ImageNet-128	$a = 1/2,\ b = -x$	moderate EGCA	1.00	800	164h	31.63
ImageNet-128	$a = 1/2,\ b = -4x$	moderate EGCA	4.00	525	109h	37.15
ImageNet-128	$a = 1/2,\ b = -10x$	aggressive EGCA	10.00	325	92h	42.58

needed to reduce D_{disc} below a prescribed threshold, thus directly reflecting the convergence speed. *4) Training Time (T_{spend}).* The total wall-clock time required to train the model, used for comparing computational efficiency across different methods. The discretization protocol for D_{disc} and T_{conv} is reported in Appendix E.2; additional convergence diagnostics are provided in Appendix F.5.

4.2 Principal eigenvalue tracks forward convergence speed

Table 1 summarizes representative low, moderate, and high-eigenvalue regimes. Across the evaluated linear generator configurations, increasing λ_1 consistently reduces the convergence step T_{conv} and the training time T_{spend}. For example, on CIFAR-10, moving from the low-eigenvalue baseline $\lambda_1 = 0.5$ to $\lambda_1 = 1$ reduces training time from 45h to 35h while improving FID from 4.44 to 3.15. A similar trend appears on CelebA-HQ-64, where the same change reduces training time from 104h to 67h, and on ImageNet-128, where T_{conv} decreases from 1000 to 800 steps.

These trends agree with Theorem 3 and Corollary 4: the exponential decay rate is governed by λ_1, although the prefactor in the convergence bound depends on the initial and stationary distributions. Empirically, this makes the principal eigenvalue a useful descriptor of forward convergence speed. This efficiency gain does not imply that larger eigenvalues are always preferable: very large λ_1 values can degrade FID, motivating the efficiency–quality analysis in Subsection 4.4.

Full linear-generator results are provided in Appendix F.1, with additional fixed-horizon, stationary-law, and distance diagnostics in Appendix F.5.

Finding 1 (Eigenvalue-driven efficiency). Within each controlled generator family, the principal eigenvalue tracks forward convergence speed: larger λ_1 values generally reduce T_{conv} and T_{spend} under fixed dataset.

4.3 Principal eigenvalue as a first-order descriptor

We now test whether λ_1 remains informative when the coefficient form changes while the principal eigenvalue is kept approximately fixed. We first use tractable linear matched-λ_1 ablations, where different coefficient choices yield nearby estimated eigenvalues under the same dataset, in Table 2. Full matched-λ_1 ablations across linear and nonlinear coefficient parameterizations are provided in Appendix F.2.

The matched-λ_1 results show that nearby eigenvalues lead to comparable convergence costs, as measured by T_{conv} and T_{spend}. This supports using λ_1 as a first-order descriptor of forward convergence behavior beyond a single parameterization. At the same time, matched eigenvalues do not imply identical FID: generation quality can still vary with the coefficient structure, the induced forward trajectory, and the associated stationary distribution. Thus, λ_1 should be interpreted as an organizing spectral coordinate rather than a complete predictor of sample quality.

Table 2: Matched-λ_1 ablations across coefficient parameterizations.

Datasets	a	b	λ_1	T_{conv}	FID↓ (SDE solver)	T_{spend}
CIFAR-10	$1/2$	$-x$	1.00	825	3.15	35h
CIFAR-10	1	$-x$	1.00	825	3.23	34h
CelebA-HQ-64	$1/2$	$-x$	1.00	800	3.34	67h
CelebA-HQ-64	1	$-x$	1.00	800	3.39	67h

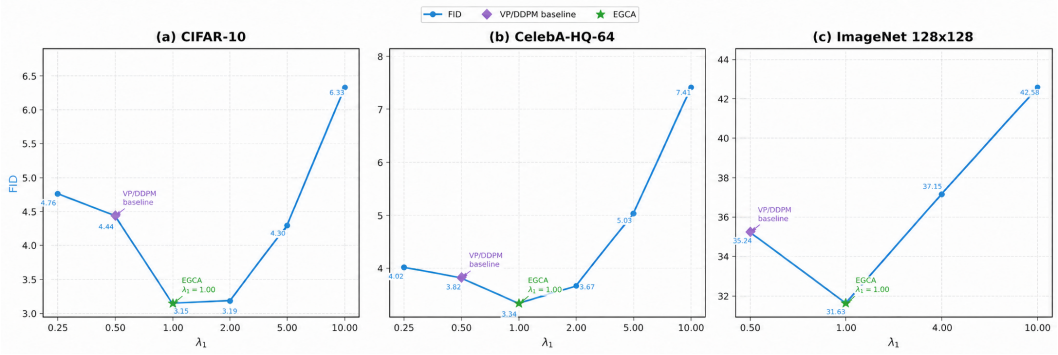


Figure 1: Effect of λ_1 on FID under the linear family $(a, b) = (1/2, -kx)$.

Finding 2 (First-order spectral descriptor). Matched or nearby principal eigenvalues yield comparable convergence costs, while FID may still vary across coefficient parameterizations. Thus, λ_1 captures the dominant convergence-rate behavior but does not fully determine generation quality.

4.4 Eigenvalue-organized efficiency–quality trade-off

The previous subsections show that λ_1 tracks forward convergence cost. We now examine how this spectral speed affects generation quality. Figure 1 and Table 1 show that FID is generally non-monotone in λ_1: moderate eigenvalue regimes often improve both efficiency and quality, whereas overly large eigenvalues continue to reduce T_{conv} and T_{spend} but degrade FID. Thus, λ_1 should not be maximized blindly; it should be used to select a favorable spectral regime.

Complementary sFID, Inception Score, Precision, and Recall evaluations, together with a seed-controlled qualitative comparison of representative low, moderate, and aggressive regimes, are reported in Appendix F.6.

When a generator is too aggressive, schedule modulation provides a complementary control route. Rescaling $\beta(t)$ changes the accumulated spectral time $\lambda_1 \int_\tau^t \beta(s)\, ds$ without changing the generator itself. Table 3 shows that this can recover sample quality for high-λ_1 generators while retaining part of the efficiency gain. Nonlinear-generator diagnostics in Appendix F.3 show a similar qualitative efficiency–quality trade-off.

Finding 3 (Eigenvalue-organized trade-off). The principal eigenvalue λ_1 organizes the efficiency–quality trade-off: moderate spectral regimes can maintain or improve FID while reducing training cost, whereas overly large λ_1 may further accelerate convergence at the expense of generation quality.

4.5 Factors beyond the spectral gap

Stationary law and forward trajectory. While λ_1 determines the contraction exponent in our theoretical setting, it does not by itself determine generation quality. Changing the spectral rate in the main linear family can also alter the invariant law, conditional noise variance, and log-SNR trajectory. To control for the invariant law, we compare $(a, b) = (1/2, -0.5x)$ and $(1, -x)$, which share the invariant law $\mathcal{N}(0, 1)$ but have $\lambda_1 = 0.5$ and 1, respectively. On both CIFAR-10 and CelebA-HQ-64, the larger-gap configuration reaches the convergence criterion earlier and yields lower FID, indicating

Table 3: Schedule modulation for fixed generators on CIFAR-10. Rescaling $\beta(t)$ changes the accumulated spectral time without changing λ_1.

Generator	λ_1	Schedule	T_{conv}	D_{disc}	FID↓ (SDE solver)	T_{spend}
$a = 1/2,\ b = -10x$	10.00	$\beta(t)$	350	0.03	6.33	20h
$a = 1/2,\ b = -10x$	10.00	$(1/10)\beta(t)$	825	0.03	3.94	35h
$a = 1/2,\ b = -10x$	10.00	$(0.5/10)\beta(t)$	950	0.03	3.74	41h

that invariant-law differences alone do not explain the observed quality trend. This comparison does not, however, isolate λ_1 from other coupled changes in the forward process. Appendix F.5.2 gives the corresponding stationary-law controls, analytical forward-trajectory statistics, and representative terminal values.

Convergence horizon and empirical diagnostics. The observed quality trend is not explained solely by differences in the forward horizon. In the fixed-T_{conv} control (Appendix F.5.1), setting $T_{\mathrm{conv}} = 1000$ across the linear generator sweep preserves the same qualitative non-monotone FID pattern, with intermediate spectral regimes remaining more favorable than both the low and highly aggressive regimes. We also emphasize that D_{disc} is used only as an empirical binned marginal-TV proxy for selecting a stable forward horizon, rather than as a measurement of the full high-dimensional discrepancy appearing in Theorem 3. Additional distance diagnostics along a fixed-generator forward trajectory are reported in Appendix F.5.3.

Dataset and configuration dependence. Taken together, these controls support using λ_1 as a convergence-oriented coordinate rather than as a complete predictor of generation quality. The favorable spectral regime is empirical and depends on the dataset and experimental configuration: increasing λ_1 can shorten the empirical forward convergence horizon, while the resulting FID can also vary with the invariant law, noise scale, log-SNR trajectory, coefficient structure, and denoising or sampling configuration. Accordingly, EGCA is intended to organize the search over forward-process choices rather than to prescribe a universal optimal value of λ_1.

4.6 Compatibility with existing diffusion settings

We finally examine whether eigenvalue-guided forward-process design can be used with existing components of diffusion pipelines. Table 4(a) combines the generator modification with representative reverse-time samplers, including DDIM [36], PNDM [24], EVODiff [21], DPM-Solver [26], and DPM-Solver++ [27]. Because these methods modify reverse-time integration rather than the forward generator, these comparisons provide evidence that the generator modification remains compatible with representative sampler-level acceleration methods.

Table 4(b) reports controlled within-pair checks under selected training formulations and denoising architectures. EDM/EDM2-style [16, 17] experiments use a T_{conv}-proportional reduced training budget, while DiT [31] is interpreted as a paired training-trajectory comparison. Flow-matching-style experiments [23] use an equal-epoch protocol and are treated as empirical boundary checks outside the stochastic-generator theory developed here. Detailed sampler/NFE diagnostics, training budgets, convergence status, and evaluation protocols are provided in Appendix F.4.

Finding 4 (Compatibility under evaluated protocols). Across the evaluated settings, the generator modification shows consistent relative trends with representative reverse-time samplers, training formulations, and denoising architectures. The broader architecture/training comparisons provide limited compatibility evidence under the reported protocols rather than fully converged or state-of-the-art comparisons.

Table 4: Compatibility of EGCA with representative diffusion components and settings.

(a) Representative reverse-time samplers.

Sampler	Dataset	Baseline (VP-type) FID↓	EGCA FID↓
DDIM	CIFAR-10	4.72	3.38
DPM-Solver	CIFAR-10	4.59	3.74
DPM-Solver++	CIFAR-10	4.44	3.15
PNDM	CIFAR-10	13.64	13.43
EVODiff	CIFAR-10	16.42	15.25

(b) Representative diffusion settings.

Paradigm	Dataset	Baseline (VP-type)		EGCA		Protocol
		FID↓	Epochs	FID↓	Epochs	
EDM	CIFAR-10	13.22	200	11.27	165	Reduced
EDM2	CIFAR-10	12.56	200	10.05	165	Reduced
FM	CIFAR-10	6.44	300	6.06	300	Equal
DiT	ImageNet-256	124	50	98	41	Paired

5 Related work

Diffusion forward processes and complementary improvements. Diffusion models prescribe a forward corruption process and learn a reverse-time generative process. DDPMs use a discrete Gaussian noising chain with a prescribed variance schedule [12], while score-based generative modeling formulates related dynamics through continuous-time SDEs, including VP- and VE-type processes [37]. Subsequent work improves schedules, parameterizations, and training objectives [16, 18, 22, 29]. Other advances accelerate reverse-time sampling [24, 26, 27, 36, 41] or improve denoising architectures and training recipes [8, 17, 31]. Representation-alignment methods such as REPA [40] act at the level of denoiser representation learning rather than forward-generator design; their joint use with EGCA is not evaluated here. These directions largely keep the forward generator fixed or choose it from canonical families. In contrast, EGCA studies the forward process at the generator level and uses the principal eigenvalue as a spectral coordinate for comparing and designing admissible generators.

Spectral analysis of Markov generators. Spectral gaps, Poincaré-type inequalities, and Markov semigroup techniques are classical tools for characterizing convergence to stationarity [1, 3, 5]. For reversible diffusions, the first nonzero eigenvalue of the generator controls exponential convergence in suitable norms, and one-dimensional diffusion theory provides practical stability-rate estimates [4]. Our contribution is to bring this spectral viewpoint into forward-process design for diffusion models: we connect the principal eigenvalue to forward convergence, adapt it into a practical generator-selection criterion, and empirically show that it organizes convergence cost and the efficiency–quality trade-off.

Additional discussion of adjacent formulations, including VE-SDEs, kinetic diffusions, and flow-matching-style probability paths, is provided in Appendix D.

6 Conclusion

We introduced EGCA, a spectral framework for forward-process design in diffusion models. By analyzing the infinitesimal generator and its principal eigenvalue, EGCA provides a computable coordinate that links generator structure to forward convergence under suitable ergodicity conditions. Empirically, this eigenvalue tracks convergence speed and organizes the efficiency–quality trade-off: moderate spectral regimes reduce training cost while preserving or improving generation quality, whereas overly aggressive regimes may degrade FID. The framework is complementary to existing advances in samplers, training formulations, and denoising architectures. Extending this spectral perspective beyond the separable ergodic setting to more general diffusion dynamics remains an important direction for future work.

Limitations and broader impacts are discussed in Appendix G.

Acknowledgments and Disclosure of Funding

We thank the reviewers and area chair for their constructive comments and suggestions, which helped improve the presentation and evaluation of this work. The authors from Ocean University of China were partially supported by the National Natural Science Foundation of China (Grants No. 42476194 and U24A20242), the Natural Science Foundation of Shandong Province (Grant No. ZR2024MF097), the Natural Science Foundation of Hainan Province (Grant No. 425MS106), and the Fundamental Research Funds for the Central Universities (Grant No. 202264005). The authors declare no competing interests.

References

[1] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and Geometry of Markov Diffusion Operators*. Springer, 2014.

[2] Adam Bobrowski. *Functional Analysis for Probability and Stochastic processes*. Cambridge University Press, 2005.

[3] Mu-Fa Chen. *Eigenvalues, Inequalities, and Ergodic theory*. Springer, 2005.

[4] Mu-Fa Chen. Basic estimates of stability rate for one-dimensional diffusions. In *Probability Approximations and Beyond*, pages 75–99, 2012.

[5] Mu-Fa Chen and Yong-Hua Mao. *Introduction to Stochastic Processes*. World Scientific, 2021.

[6] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural ordinary differential equations. *Advances in Neural Information Processing Systems*, 31:6572–6583, 2018.

[7] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[8] Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances in Neural Information Processing Systems*, 34:8780–8794, 2021.

[9] Tim Dockhorn, Arash Vahdat, and Karsten Kreis. Score-based generative modeling with critically-damped langevin diffusion. In *International Conference on Learning Representations*, 2022.

[10] Stewart N Ethier and Thomas G Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley & Sons, 1986.

[11] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in Neural Information Processing Systems*, 30:6626–6637, 2017.

[12] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

[13] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic Differential Equations and Diffusion Processes*. Elsevier, 2014.

[14] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1991.

[15] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. In *International Conference on Learning Representations*, 2018.

[16] Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. *Advances in Neural Information Processing Systems*, 35: 26565–26577, 2022.

[17] Tero Karras, Miika Aittala, Jaakko Lehtinen, Janne Hellsten, Timo Aila, and Samuli Laine. Analyzing and improving the training dynamics of diffusion models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 24174–24184, 2024.

[18] Diederik Kingma, Tim Salimans, Ben Poole, and Jonathan Ho. Variational diffusion models. *Advances in Neural Information Processing Systems*, 34:21696–21707, 2021.

[19] Alex Krizhevsky. Learning multiple layers of features from tiny images. In *Technical Report, University of Toronto*, 2009.

[20] Tuomas Kynkäänniemi, Tero Karras, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Improved precision and recall metric for assessing generative models. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

[21] Shigui Li, Wei Chen, and Delu Zeng. Evodiff: Entropy-aware variance optimized diffusion inference. *Advances in Neural Information Processing Systems*, 38:148134–148181, 2025.

[22] Shanchuan Lin, Bingchen Liu, Jiashi Li, and Xiao Yang. Common diffusion noise schedules and sample steps are flawed. In *IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 5404–5411, 2024.

[23] Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023.

[24] Luping Liu, Yi Ren, Zhijie Lin, and Zhou Zhao. Pseudo numerical methods for diffusion models on manifolds. In *International Conference on Learning Representations*, 2022.

[25] Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *International Conference on Learning Representations*, 2023.

[26] Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. Dpm-solver: A fast ode solver for diffusion probabilistic model sampling in around 10 steps. *Advances in Neural Information Processing Systems*, 35:5775–5787, 2022.

[27] Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. Dpm-solver++: Fast solver for guided sampling of diffusion probabilistic models. *Machine Intelligence Research*, 22(4):730–751, 2025.

[28] Charlie Nash, Jacob Menick, Sander Dieleman, and Peter Battaglia. Generating images with sparse representations. In *International Conference on Machine Learning*, pages 7958–7968, 2021.

[29] Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *International Conference on Machine Learning*, pages 8162–8171, 2021.

[30] Bernt Oksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer Science & Business Media, 2013.

[31] William Peebles and Saining Xie. Scalable diffusion models with transformers. In *IEEE/CVF International Conference on Computer Vision*, pages 4195–4205, 2023.

[32] Danilo Rezende and Shakir Mohamed. Variational inference with normalizing flows. In *International Conference on Machine Learning*, pages 1530–1538, 2015.

[33] L. C. G. Rogers and David Williams. *Diffusions, Markov Processes and Martingales, Volume 2: Itô Calculus*. Cambridge University Press, 2000.

[34] Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models. *International Conference on Learning Representations*, 2022.

[35] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training GANs. In *Advances in Neural Information Processing Systems*, volume 29, 2016.

[36] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021.

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

[38] Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. In *International Conference on Machine Learning*, pages 32211–32252, 2023.

[39] Kai Wang, Mingjia Shi, Yukun Zhou, Zekai Li, Zhihang Yuan, Yuzhang Shang, Xiaojiang Peng, Hanwang Zhang, and Yang You. A closer look at time steps is worthy of triple speed-up for diffusion model training. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Conference*, pages 12934–12944, 2025.

[40] Sihyun Yu, Sangkyung Kwak, Huiwon Jang, Jongheon Jeong, Jonathan Huang, Jinwoo Shin, and Saining Xie. Representation alignment for generation: Training diffusion transformers is easier than you think. In *International Conference on Learning Representations*, 2025.

[41] Qinsheng Zhang and Yongxin Chen. Fast sampling of diffusion models with exponential integrator. In *International Conference on Learning Representations*, 2023.

A Proofs of ergodicity and spectral convergence

Throughout the appendix, vectors in $\mathbb{R}^d$ are written in bold, while scalar coordinates remain unbolded. We use lower-case functions such as f, g for functions on $\mathbb{R}$ and upper-case functions such as F, G for functions on $\mathbb{R}^d$. Generator identities are first checked on $C_c^\infty(\mathbb{R}^d)$ or $C_c^\infty(\mathbb{R})$, the spaces of compactly supported smooth test functions, and then extended to the relevant generator domains by density and closure. We prove Theorem 2, Theorem 3, and Corollary 4.

For completeness, we first recall the forward dynamics and the corresponding generator notation used throughout the proofs. The scheduled forward process in Eq. (2) satisfies

$$\mathrm{d}\boldsymbol{X}_t = \beta(t)\boldsymbol{b}(\boldsymbol{X}_t)\,\mathrm{d}t + \sqrt{2\beta(t)}\,\boldsymbol{A}(\boldsymbol{X}_t)^{1/2}\,\mathrm{d}\boldsymbol{W}_t,$$

with

$$\boldsymbol{b}(\boldsymbol{x}) = \big(b(x_1), \dots, b(x_d)\big)^\top, \quad \boldsymbol{A}(\boldsymbol{x}) = \mathrm{diag}\big(a(x_1), \dots, a(x_d)\big).$$

Hence each coordinate satisfies

$$\mathrm{d}X_t^{(i)} = \beta(t)b\big(X_t^{(i)}\big)\,\mathrm{d}t + \sqrt{2\beta(t)a\big(X_t^{(i)}\big)}\,\mathrm{d}W_t^{(i)}, \quad i = 1, \dots, d.$$

For $F \in C_c^\infty(\mathbb{R}^d)$ and $\boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$, the separable generator is

$$\mathcal{L}_t = \beta(t)\mathcal{L}_{\mathrm{sep}}, \quad \mathcal{L}_{\mathrm{sep}}F(\boldsymbol{x}) = \sum_{i=1}^{d} \mathcal{L}^{(i)}F(\boldsymbol{x}),$$

where

$$\mathcal{L}^{(i)}F(\boldsymbol{x}) = a(x_i)\partial_{ii}F(\boldsymbol{x}) + b(x_i)\partial_i F(\boldsymbol{x}).$$

The common one-dimensional coordinate generator is

$$\mathcal{L}f(x) = a(x)f''(x) + b(x)f'(x), \quad f \in C_c^\infty(\mathbb{R}).$$

Define the accumulated operational time

$$B(t) := \textstyle\int_0^t \beta(s)\,\mathrm{d}s, \quad B(s,t) := B(t) - B(s) = \int_s^t \beta(u)\,\mathrm{d}u.$$

By (A4) in Assumption 1, B(t) is continuous and strictly increasing to ∞ as $t \to \infty$.

A.1 Auxiliary lemmas

We first collect the main ingredients used in the proofs: deterministic time change, the invariant and reversible structure of the homogeneous coordinate and product diffusions, and tensorization of the one-dimensional spectral gap [1–3]. The time-change argument first removes the explicit schedule dependence in $\mathcal{L}_t = \beta(t)\mathcal{L}_{\mathrm{sep}}$, after which the invariant measure, Dirichlet form, and spectral gap can be analyzed through the time-homogeneous operator $\mathcal{L}_{\mathrm{sep}}$.

Lemma 6 (Time change)**.** *Let $(Q_r)_{r\geq 0}$ be the Markov semigroup generated by $\mathcal{L}_{\mathrm{sep}}$, and let $(P_{s,t})_{0\leq s\leq t}$ be the evolution family generated by $\mathcal{L}_t = \beta(t)\mathcal{L}_{\mathrm{sep}}$. Then, for $0 \leq s \leq t$,*

$$P_{s,t} = Q_{B(t)-B(s)}.$$

Conversely, let $T = B^{-1}$, then

$$Q_{r_2-r_1} = P_{T(r_1),T(r_2)}, \quad 0 \leq r_1 \leq r_2.$$

Thus the two evolutions determine each other through the deterministic time change B.

Proof. Let $\boldsymbol{Y}_r$ solve the homogeneous martingale problem for $\mathcal{L}_{\mathrm{sep}}$ and define $\boldsymbol{X}_t := \boldsymbol{Y}_{B(t)}$. For every $F \in C_c^\infty(\mathbb{R}^d)$,

$$F(\boldsymbol{Y}_r) - F(\boldsymbol{Y}_0) - \textstyle\int_0^r \mathcal{L}_{\mathrm{sep}}F(\boldsymbol{Y}_u)\,\mathrm{d}u$$

is a martingale. Setting $r = B(t)$ and using $u = B(v)$ gives

$$\textstyle\int_0^{B(t)} \mathcal{L}_{\mathrm{sep}}F(\boldsymbol{Y}_u)\,\mathrm{d}u = \int_0^t \beta(v)\mathcal{L}_{\mathrm{sep}}F(\boldsymbol{X}_v)\,\mathrm{d}v = \int_0^t \mathcal{L}_v F(\boldsymbol{X}_v)\,\mathrm{d}v.$$

Hence $\boldsymbol{X}$ solves the martingale problem for $\mathcal{L}_t = \beta(t)\mathcal{L}_{\mathrm{sep}}$. Since $\beta > 0$, B is strictly increasing; applying the inverse clock $T = B^{-1}$ gives the converse correspondence. See [10] for time changes of martingale problems. Therefore the two equations hold. □

We next show that tensorization from the one-dimensional coordinate diffusion to the product diffusion preserves the spectral gap. Consequently, the same λ_1 governs the d-dimensional convergence bound.

Lemma 7 (One-dimensional diffusion)**.** *Under (A1)–(A3), the coordinate SDE*

$$\mathrm{d}Y_r = b(Y_r)\,\mathrm{d}r + \sqrt{2a(Y_r)}\,\mathrm{d}W_r \tag{8}$$

admits a conservative weak solution that is unique in law. Its invariant probability measure is π in Eq. (4), and the diffusion is ergodic with respect to π. Moreover, $\mathcal{L}$ is symmetric in $L^2(\pi)$ and

$$-\langle g, Lg\rangle_{\mathcal{L}^2(\pi)} = \textstyle\int_{\mathbb{R}} a(x)|g'(x)|^2\,\pi(\mathrm{d}x), \quad g \in C_c^\infty(\mathbb{R}). \tag{9}$$

Proof. Under (A1), $1/a$ and b/a are locally integrable. Standard one-dimensional diffusion theory therefore yields a weak solution of Eq. (8) that is unique in law up to its explosion time [14]. The scale and speed densities are proportional to $e^{-C(x)}$ and $e^{C(x)}/a(x)$, respectively. By the Feller boundary test, (A2) rules out explosion at both boundaries, while (A3) makes the speed measure finite. Hence the diffusion is conservative and ergodic with invariant probability measure π [14, 33]. Writing $\rho(x) = Z^{-1}e^{C(x)}/a(x)$, we have $\mathcal{L}g = \frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}x}\big(a\rho g'\big)$. Thus, for $g, h \in C_c^\infty(\mathbb{R})$, integration by parts gives

$$\langle h, \mathcal{L}g\rangle_{L^2(\pi)} = -\textstyle\int_{\mathbb{R}} a(x)h'(x)g'(x)\,\pi(dx).$$

Therefore $\mathcal{L}$ is symmetric in $L^2(\pi)$, and taking $h = g$ yields Eq. (9). □

Lemma 8 (Product process and invariant law)**.** *Under (A1)–(A3), the process for L_{sep} is conservative and well posed. Its transition kernel is*

$$Q_r(\boldsymbol{x}, \mathrm{d}\boldsymbol{y}) = \textstyle\bigotimes_{i=1}^d q_r(x_i, \mathrm{d}y_i). \tag{10}$$

Consequently, Π in (4) is invariant and reversible for Q_r.

Proof. By Lemma (7), the one-dimensional process for $\mathcal{L}$ is conservative and well posed, with transition kernel $q_r(x, \mathrm{d}y)$ and reversible invariant measure π. The standard finite-product construction for well-posed martingale problems therefore gives a conservative and well-posed process with generator $\mathcal{L}_{\mathrm{sep}}$, whose transition kernel is Q_r in (10) [10].

Since π is invariant for q_r, Fubini's theorem yields

$$\int_{\mathbb{R}^d} Q_r F(\boldsymbol{x})\, \Pi(\mathrm{d}\boldsymbol{x}) = \int_{\mathbb{R}^d} F(\boldsymbol{x})\, \Pi(\mathrm{d}\boldsymbol{x}).$$

Hence Π is invariant. Moreover, reversibility of each coordinate semigroup tensorizes, giving

$$\langle F, Q_r G\rangle_{L^2(\Pi)} = \langle Q_r F, G\rangle_{L^2(\Pi)}.$$

Thus Q_r is reversible with respect to Π. □

Lemma 9 (Product Dirichlet form)**.** *Under (A1)–(A3), for $F \in C_c^\infty(\mathbb{R}^d)$,*

$$-\langle F, \mathcal{L}_{\mathrm{sep}} F\rangle_{L^2(\Pi)} = \textstyle\sum_{i=1}^d \int_{\mathbb{R}^d} a(x_i) |\partial_i F(\boldsymbol{x})|^2\, \Pi(d\boldsymbol{x}). \tag{11}$$

Proof. Applying the same identity coordinatewise Eq.(9) and using Fubini's theorem gives (11). □

The final auxiliary step shows that passing from one coordinate to the product diffusion does not change the spectral exponent. This justifies using the same λ_1 in the d-dimensional convergence statement.

Lemma 10 (Tensorization of the spectral gap)**.** *The product generator $\mathcal{L}_{\mathrm{sep}}$ has the same spectral gap in $L^2(\Pi)$:*

$$\inf\Big\{ \langle F, -\mathcal{L}_{\mathrm{sep}} F\rangle_{L^2(\Pi)} : F \in \mathcal{D}(\mathcal{L}_{\mathrm{sep}}),\ \Pi(F) = 0,\ \|F\|_{L^2(\Pi)} = 1 \Big\} = \lambda_1, \tag{12}$$

where λ_1 is the one-dimensional spectral gap of $\mathcal{L}$ in Eq.(5).

Proof. We use the following variance notation:

$$\mathrm{Var}_\pi(g) := \textstyle\int_{\mathbb{R}} \big(g(x) - \pi(g)\big)^2\, \pi(\mathrm{d}x), \quad \mathrm{Var}_\Pi(F) := \int_{\mathbb{R}^d} \big(F(\boldsymbol{x}) - \Pi(F)\big)^2\, \Pi(\mathrm{d}\boldsymbol{x}).$$

The variational definition of λ_1 is equivalent to the one-dimensional Poincare inequality

$$\lambda_1 \mathrm{Var}_\pi(g) \le \textstyle\int_{\mathbb{R}} a(x) |g'(x)|^2\, \pi(\mathrm{d}x), \tag{13}$$

with the usual interpretation when the infimum is approached by an approximating sequence [1, 3]. For $F \in C_c^\infty(\mathbb{R}^d)$, variance tensorization under Π gives

$$\mathrm{Var}_\Pi(F) \le \textstyle\sum_{i=1}^d \int_{\mathbb{R}^{d-1}} \mathrm{Var}_\pi\big(F(x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_d)\big)\, \Pi_{-i}(\mathrm{d}\boldsymbol{x}_{-i}), \tag{14}$$

where $\Pi_{-i}(\mathrm{d}\boldsymbol{x}_{-i}) = \bigotimes_{j \ne i} \pi(\mathrm{d}x_j)$. Explicitly, for fixed $\boldsymbol{x}_{-i}$,

$$\mathrm{Var}_\pi\big(F(x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_d)\big) := \textstyle\int_{\mathbb{R}} \big(F(x_1, \dots, z, \dots, x_d) - m_i(\boldsymbol{x}_{-i})\big)^2\, \pi(\mathrm{d}z), \tag{15}$$

where $m_i(\boldsymbol{x}_{-i}) := \int_{\mathbb{R}} F(x_1, \dots, u, \dots, x_d)\, \pi(\mathrm{d}u)$. Applying Eq.(13) to each conditional variance and then using (11), we obtain

$$\lambda_1 \mathrm{Var}_\Pi(F) \le \textstyle\sum_{i=1}^d \int_{\mathbb{R}^d} a(x_i) |\partial_i F(\boldsymbol{x})|^2\, \Pi(\mathrm{d}\boldsymbol{x}) = \langle F, -\mathcal{L}_{\mathrm{sep}} F\rangle_{L^2(\Pi)}.$$

Thus the product spectral gap is at least λ_1.

For the reverse inequality, choose a sequence $g_n \in \mathcal{D}(\mathcal{L})$ with $\pi(g_n) = 0$, $\|g_n\|_{L^2(\pi)} = 1$, and $\langle g_n, -\mathcal{L} g_n\rangle_{L^2(\pi)} \to \lambda_1$. Define $F_n(\boldsymbol{x}) = g_n(x_1)$. Then $\Pi(F_n) = 0$, $\|F_n\|_{L^2(\Pi)} = 1$, and

$$\langle F_n, -\mathcal{L}_{\mathrm{sep}} F_n\rangle_{L^2(\Pi)} = \langle g_n, -\mathcal{L} g_n\rangle_{L^2(\pi)} \to \lambda_1.$$

Hence the product spectral gap is at most λ_1, proving Eq.(12). □

Lemma 11 (L^2 spectral contraction)**.** *Under (A1)–(A3), for $G \in L^2(\Pi)$ with $\Pi(G) = 0$,*

$$\|Q_r G\|_{L^2(\Pi)} \le e^{-\lambda_1 r} \|G\|_{L^2(\Pi)}, \qquad r \ge 0. \tag{16}$$

Proof. By Lemma 10, the product generator $\mathcal{L}_{\mathrm{sep}}$ has spectral gap λ_1 in $L^2(\Pi)$. Let $G \in L^2(\Pi)$ satisfy $\Pi(G) = 0$, and set $u_r = Q_r G$. Since Π is invariant, $\Pi(u_r) = 0$. Using symmetry and the variational definition of the spectral gap,

$$\tfrac{\mathrm{d}}{\mathrm{d}r} \|u_r\|_{L^2(\Pi)}^2 = 2\langle u_r, \mathcal{L}_{\mathrm{sep}} u_r\rangle_{L^2(\Pi)} = -2\langle u_r, -\mathcal{L}_{\mathrm{sep}} u_r\rangle_{L^2(\Pi)} \le -2\lambda_1 \|u_r\|_{L^2(\Pi)}^2.$$

Gronwall's inequality gives the $L^2(\Pi)$ spectral contraction

$$\|Q_r G\|_{L^2(\Pi)} \le e^{-\lambda_1 r} \|G\|_{L^2(\Pi)}.$$

This is the standard reversible spectral-gap contraction [1, 3]. □

A.2 Proof of Theorem 2

Proof. By Lemma 8, under (A1)–(A3), the homogeneous d-dimensional martingale problem for $\mathcal{L}_{\mathrm{sep}}$ is conservative and well posed. Lemma 6 therefore implies that the scheduled SDE (2) admits a conservative weak solution that is unique in law, with evolution family $P_{s,t} = Q_{B(t)-B(s)}$. Since Π is invariant for Q_r, it is a common invariant probability measure for the scheduled evolution:

$$\Pi P_{s,t} = \Pi Q_{B(s,t)} = \Pi, \qquad 0 \le s \le t.$$

It remains to prove convergence. By Lemma 7, for each coordinate, $q_r(x_i, \cdot) \Longrightarrow \pi$ as $r \to \infty$. Hence, by the product representation (10),

$$Q_r(\boldsymbol{x}, \cdot) \Longrightarrow \Pi, \qquad r \to \infty,$$

for $\boldsymbol{x} \in \mathbb{R}^d$. By (A4), $B(t) - B(s) \to \infty$ as $t \to \infty$ for every fixed $s \ge 0$. Therefore,

$$P_{s,t}(\boldsymbol{x}, \cdot) = Q_{B(s,t)}(\boldsymbol{x}, \cdot) \Longrightarrow \Pi, \qquad t \to \infty.$$

This proves the well-posedness and ergodicity statement of Theorem 2. □

A.3 Proof of Theorem 3

Proof. By Lemma 6 and Lemma 11,

$$\|P_{\tau,t} G\|_{L^2(\Pi)} \le \exp\left(-\lambda_1 \textstyle\int_\tau^t \beta(s)\,\mathrm{d}s\right) \|G\|_{L^2(\Pi)}. \tag{17}$$

Since $\mu_\tau \ll \Pi$ with $\chi^2(\mu_\tau \| \Pi) < \infty$, we could define $h_\tau := \frac{\mathrm{d}\mu_\tau}{\mathrm{d}\Pi} \in L^2(\Pi)$, $h_t := \frac{\mathrm{d}\mu_t}{\mathrm{d}\Pi}$. Because the semigroup is symmetric in $L^2(\Pi)$, the density evolves by the same operator: $h_t = P_{\tau,t} h_\tau$. Moreover, $P_{\tau,t} 1 = 1$ and $\Pi(h_\tau - 1) = 0$, hence $h_t - 1 = P_{\tau,t}(h_\tau - 1)$. Applying (17) with $G = h_\tau - 1$ gives

$$\|h_t - 1\|_{L^2(\Pi)} \le \exp\left(-\lambda_1 \textstyle\int_\tau^t \beta(s)\,\mathrm{d}s\right) \|h_\tau - 1\|_{L^2(\Pi)}.$$

Since

$$\chi^2(\mu_t \| \Pi) = \textstyle\int_{\mathbb{R}^d} \left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\Pi} - 1\right)^2 \Pi(\mathrm{d}\boldsymbol{x}) = \|h_t - 1\|^2_{L^2(\Pi)},$$

squaring the preceding bound yields

$$\chi^2(\mu_t \| \Pi) \le \chi^2(\mu_\tau \| \Pi) \exp\left(-2\lambda_1 \textstyle\int_\tau^t \beta(s)\,\mathrm{d}s\right), \qquad \forall\, t \ge \tau.$$

Taking $\tau = 0$ gives the corresponding bound from the initial distribution. □

A.4 Proof of Corollary 4

Proof. Since $\mu_t \ll \Pi$, total variation has the density representation

$$\|\mu_t - \Pi\|_{\mathrm{TV}} = \tfrac{1}{2} \textstyle\int_{\mathbb{R}^d} |h_t - 1|\, \Pi(\mathrm{d}\boldsymbol{x}),$$

where $h_t = \mathrm{d}\mu_t / \mathrm{d}\Pi$. By the Cauchy–Schwarz inequality,

$$\textstyle\int_{\mathbb{R}^d} |h_t - 1|\, \Pi(\mathrm{d}\boldsymbol{x}) \le \|h_t - 1\|_{L^2(\Pi)} \|1\|_{L^2(\Pi)} = \|h_t - 1\|_{L^2(\Pi)}.$$

Then the $L^2(\Pi)$ estimate of Theorem 3 gives

$$\|\mu_t - \Pi\|_{\mathrm{TV}} \le \tfrac{1}{2} \|h_t - 1\|_{L^2(\Pi)} \le \tfrac{1}{2} \exp\left(-\lambda_1 \textstyle\int_\tau^t \beta(s)\,\mathrm{d}s\right) \|h_\tau - 1\|_{L^2(\Pi)}.$$

Finally, $\|h_\tau - 1\|_{L^2(\Pi)} = \chi^2(\mu_\tau \| \Pi)^{1/2}$ gives the conclusion

$$\|\mu_t - \Pi\|_{\mathrm{TV}} \le \tfrac{1}{2} \chi^2(\mu_\tau \| \Pi)^{1/2} \exp\left(-\lambda_1 \textstyle\int_\tau^t \beta(s)\,\mathrm{d}s\right), \quad \forall\, t \ge \tau.$$

□

A.5 Proof of Proposition 5

Proof. Let $a^*(x) = ka(x)$ and $b^*(x) = kb(x)$ with $k > 0$. Then

$$C^*(x) = \int_0^x \frac{b^*(u)}{a^*(u)}\,\mathrm{d}u = \int_0^x \frac{b(u)}{a(u)}\,\mathrm{d}u = C(x),$$

and

$$Z^* = \int_{\mathbb{R}} \frac{e^{C^*(x)}}{a^*(x)}\,\mathrm{d}x = \frac{1}{k}Z.$$

So the invariant probability measure is unchanged:

$$\pi^*(\mathrm{d}x) = \frac{1}{Z^*}\frac{e^{C^*(x)}}{a^*(x)}\,\mathrm{d}x = \pi(\mathrm{d}x).$$

The scaled generator satisfies $\mathcal{L}^* = k\mathcal{L}$ on the same $L^2(\pi)$ domain. Therefore its spectral gap is

$$\begin{aligned}\lambda_1^* &= \inf\big\{\langle f, -\mathcal{L}^* f\rangle_{L^2(\pi)} : f \in D(\mathcal{L}),\ \pi(f) = 0,\ \|f\|_{L^2(\pi)} = 1\big\} \\ &= k\inf\big\{\langle f, -\mathcal{L} f\rangle_{L^2(\pi)} : f \in D(\mathcal{L}),\ \pi(f) = 0,\ \|f\|_{L^2(\pi)} = 1\big\} = k\lambda_1.\end{aligned}$$

This completes the proof. □

B Numerical estimation and generator admissibility

B.1 Additional details on the eigenvalue estimator

Iterative Construction. Section 3.3 gives the main estimator used to compute the principal eigenvalue. For completeness, we provide the full iterative construction here, which is given in [4].

For any $x < y$, initialize a function $f_1(z)$ as follows:

$$f_1(z) = \begin{cases} \left[\dfrac{\int_y^\infty \frac{e^{C(u)}}{a(u)}\mathrm{d}u \cdot \int_{-\infty}^z \frac{e^{C(u)}}{a(u)}\mathrm{d}u}{\int_{-\infty}^x \frac{e^{C(u)}}{a(u)}\mathrm{d}u}\right]^{1/2} & \text{if } z \le x, \\ \left[\displaystyle\int_y^\infty \frac{e^{C(u)}}{a(u)}\mathrm{d}u\right]^{1/2} & \text{if } x \le z \le y, \\ \left[\displaystyle\int_z^\infty \frac{e^{C(u)}}{a(u)}\mathrm{d}u\right]^{1/2} & \text{if } z \ge y. \end{cases}$$

For $n \ge 2$, define

$$f_n(z) = \begin{cases} f_n^-(z) & \text{if } z \le \theta_n; \\ f_n^+(z) & \text{if } z > \theta_n, \end{cases}$$

where

$$\begin{aligned} f_n^-(z) &= \int_{-\infty}^z \frac{e^{C(u)}}{a(u)}\left[\int_u^{\theta_n} e^{-C(t)} f_{n-1}(t)\mathrm{d}t\right]\mathrm{d}u; \\ f_n^+(z) &= \int_z^{+\infty} \frac{e^{C(u)}}{a(u)}\left[\int_{\theta_n}^u e^{-C(t)} f_{n-1}(t)\mathrm{d}t\right]\mathrm{d}u, \end{aligned}$$

and θ_n is obtained by solving the below equation for every (x, y):

$$f_n^-(\theta_n) = f_n^+(\theta_n).$$

Then the n-th estimation for the λ_1 is

$$\widehat{\lambda}_1^{[n]} = \sup_{x<y}\left[\inf_{z\le\theta_n}\frac{f_{n-1}(z)}{f_n^-(z)} \wedge \inf_{z>\theta_n}\frac{f_{n-1}(z)}{f_n^+(z)}\right].$$

This sequence asymptotically and increasingly approaches to the true value of λ_1 as n tends to $+\infty$.

Closed-form reference case. To validate the numerical estimator, we use the following analytically tractable calibration cases to verify that the numerical estimator recovers the correct spectral scale before applying it to more general generators.

Proposition 12 (Spectral gap of a closed-form benchmark)**.** *Let $a(x) = a_0 > 0$ and $b(x) = -kx$, where k is a positive constant. Then the principal eigenvalue $\lambda_1 = k$.*

Proof. For $f \in C_c^\infty(\mathbb{R})$, integration by parts gives the reversible Dirichlet form

$$\langle f, -\mathcal{L}f \rangle_{L^2(\pi)} = a_0 \int_{\mathbb{R}} |f'(x)|^2 \, \pi(dx),$$

and thus the principal eigenvalue admits the variational characterization

$$\lambda_1 = \inf_{\substack{f \in D(\mathcal{L}) \\ \pi(f)=0}} \frac{a_0 \int |f'(x)|^2 \, d\pi(x)}{\int |f(x)|^2 \, d\pi(x)}.$$

From Eq.(4), the stationary distribution π is $N(0, a_0/k)$, the Gaussian Poincaré inequality gives

$$\mathrm{Var}_\pi(f) := \int_{\mathbb{R}} \big(f(x) - \pi(f)\big)^2 \, \pi(\mathrm{d}x) \le \frac{a_0}{k} \int |f'(x)|^2 \, d\pi(x),$$

which implies for $f \in D(\mathcal{L})$ with $\pi(f) = 0$ that

$$a_0 \int |f'(x)|^2 \, d\pi(x) \ge k \int |f(x)|^2 \, d\pi(x),$$

and therefore $\lambda_1 \ge k$.

On the other hand, taking $f(x) = x \in \mathcal{D}(\mathcal{L})$ gives

$$-\mathcal{L}x = -a_0 \cdot 0 + kx = kx, \quad \pi(f) = 0.$$

Hence $\lambda_1 \le k$. Combining the two inequalities yields $\lambda_1 = k$. □

B.2 Admissibility verification for candidate generators

This subsection verifies whether representative coefficient pairs considered in our generator search satisfy conditions (A1)–(A3) of Assumption 1. For each candidate generator $\mathcal{L} = a(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}}{\mathrm{d}x}$, we check continuity and positivity in (A1), the boundary condition in (A2), and the integrability condition in (A3). Condition (A4) concerns only the common schedule $\beta(t)$ and is therefore not repeated in these generator-wise checks.

Table 5 summarizes the resulting classification for the nonlinear coefficient families considered here, with $c > 0$, ✓ indicating that all of conditions (A1)–(A3) are satisfied and × indicating otherwise.

Table 5: Admissibility verification for polynomial drift families $b(x) = -x^m/2$.

$a(x)$ \ $b(x)$	$-\frac{x}{2}$	$-\frac{x^2}{2}$	$-\frac{x^3}{2}$	$-\frac{x^4}{2}$	$-\frac{x^5}{2}$	$-\frac{x^6}{2}$	$-\frac{x^7}{2}$
$\frac{\lvert x\rvert^{\frac{1}{20}}+c}{2}$	✓	×	✓	×	✓	×	✓
$\frac{\lvert x\rvert^{\frac{1}{3}}+c}{2}$	✓	×	✓	×	✓	×	✓
$\frac{\lvert x\rvert^{\frac{1}{2}}+c}{2}$	✓	×	✓	×	✓	×	✓
$\frac{\lvert x\rvert+c}{2}$	✓	×	✓	×	✓	×	✓
$\frac{\lvert x\rvert^{2}+c}{2}$	✓	×	✓	×	✓	×	✓
$\frac{\lvert x\rvert^{3}+c}{2}$	✓	✓	✓	×	✓	×	✓
$\frac{\lvert x\rvert^{4}+c}{2}$	✓	✓	✓	×	✓	×	✓

C Extended spectral positioning of existing forward processes

Table 6 summarizes representative forward-process formulations most relevant to EGCA. It is a scope clarification rather than a ranking: EGCA focuses on generator-level stochastic forward processes with an ergodic stationary regime, while training recipes, architectures, samplers, and deterministic transport paths are treated as orthogonal unless they specify such a forward generator.

Table 6: Spectral positioning of representative forward-process formulations. The comparison is a scope analysis for the ergodic generator framework used in EGCA, not a leaderboard.

Formulation	**Forward-process role**	**Within current framework?**	**Spectral interpretation**
VP-SDE/ DDPM	Confining VP-type stochastic forward process	Yes	The associated generator admits a stationary ergodic interpretation. Under the default parameterization, its principal eigenvalue is relatively small, consistent with conservative large-step settings.
EGCA-style generator	Generator-level modification of the forward process	Yes	The coefficient pair (a, b) changes the generator and hence the principal eigenvalue. This is the direct design axis studied by EGCA.
VE-SDE	Variance-exploding forward process without confining drift	No	The process does not fit the present stationary ergodic generator setting. It requires a different non-stationary or boundary-based analysis.
CLD	Kinetic or augmented-state diffusion	Not directly	The dynamics are defined on an enlarged state space and require a kinetic or hypoelliptic spectral analysis beyond the coordinate-generator setting used here.

Relation to nearby formulations. Training recipes, architectures, reverse-time samplers, and alternative transport formulations are not included in the generator-level scope table. EDM/EDM2-style recipes [17, 16], and reverse-time samplers [24, 26, 27, 36] modify other parts of the generative pipeline under a specified forward process. Flow-matching and rectified-flow formulations [23, 25] instead define alternative probability paths, and are therefore treated as boundary comparisons rather than as generator-level stochastic forward processes in the sense analyzed here.

D Extended related work and scope discussion

Forward schedules and parameterizations. A number of diffusion-model improvements can be interpreted as refinements of the forward noising schedule or its parameterization. Improved DDPM studies learned variances and cosine schedules [29], variational diffusion models revisit the variational parameterization of the diffusion process [18], and EDM reorganizes diffusion design through noise-level parameterization, preconditioning, and training choices [16]. Recent schedule and timestep analyses further show that commonly used schedules and sample-step choices may be suboptimal [22, 39]. These works are closely related in motivation, since they highlight that the forward corruption path matters. However, they do not provide a generator-level spectral quantity that directly appears in a convergence bound. EGCA addresses this complementary issue by analyzing the infinitesimal generator and using its principal eigenvalue as a computable coordinate for forward-process comparison.

Reverse-time acceleration. Sampler acceleration methods improve the numerical solution of the reverse-time dynamics under a given or approximately fixed forward process. DDIM introduces a non-Markovian sampling procedure that enables faster generation [36], PNDM develops pseudo

numerical methods for diffusion sampling [24], and DPM-Solver and DPM-Solver++ design high-order solvers for diffusion ODEs [26, 27]. Exponential-integrator-based samplers provide another line of fast reverse-time integration [41]. These methods are not alternatives to generator-level forward-process design: they accelerate the reverse trajectory after the forward process has been specified. This distinction explains why EGCA can be combined with such samplers in our compatibility experiments.

Training formulations and architectures. Other improvements modify the training recipe or the denoising model rather than the forward generator. EDM and EDM2 improve parameterization, preconditioning, and training dynamics [16, 17]; progressive distillation and consistency models reduce sampling cost through training-level reformulations [34, 38]. Architectural advances such as improved U-Nets [8], and diffusion transformers [31] improve expressiveness or computational efficiency of the denoising network. These works act at different levels of the diffusion pipeline. By contrast, EGCA changes the forward generator itself, while leaving room for these training and architectural improvements to be used on top. Representation-alignment methods provide another training-level intervention. REPA [40], for example, aligns intermediate representations of noisy inputs with clean-image features from a pretrained visual encoder through an auxiliary objective. EGCA instead modifies the forward generator and induced corruption trajectory. The two mechanisms act at different levels and are conceptually composable, but their joint effect is not evaluated here.

Alternative probability paths. Some generative modeling frameworks depart more substantially from the stochastic ergodic forward processes analyzed in this paper. Normalizing flows and neural ODEs construct deterministic transformations between data and latent variables [6, 32]. Flow matching and rectified flow learn deterministic probability paths rather than stochastic forward diffusions with stationary ergodic behavior [23, 25]. Within score-based modeling, VE-SDEs are important but do not have the same confining-drift structure as the ergodic generator setting considered here [37]. Critically damped or kinetic diffusion models introduce augmented state variables and require hypoelliptic or kinetic spectral tools [9]. These directions clarify the boundary of our current theory and suggest possible extensions of spectral forward-process design beyond the reversible or coordinate-separable setting.

Classical spectral theory. The theoretical tools used in EGCA are rooted in the classical analysis of Markov diffusion operators. Spectral gaps and Poincaré-type inequalities characterize the convergence rate of Markov semigroups to stationarity [1, 3, 5]. For reversible diffusions, the first nonzero eigenvalue of the generator governs exponential contraction in suitable norms. One-dimensional diffusion theory also provides practical estimates of stability rates [4]. Our work differs from these classical results in its goal: rather than studying convergence only as a property of a given Markov process, we use the principal eigenvalue as a design coordinate for choosing and comparing forward processes in diffusion models.

E Experimental setup and metrics

E.1 Datasets, models, and experimental protocols

The main generator-level experiments use CIFAR-10, CelebA-HQ-64, and ImageNet-128/256, as described in Section 4. For the primary comparisons, the denoising network is kept fixed and the forward generator is varied through the one-dimensional coordinate generator

$$\mathcal{L} = a(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}}{\mathrm{d}x}.$$

The VP-type baseline corresponds to the linear family with $a(x) = 1/2$ and $b(x) = -0.5x$. EGCA-style rows use the same training recipe but replace the coefficient pair with the reported generator. For each compatibility comparison, the baseline and EGCA configurations use the same dataset, model framework, denoising architecture, sampler, and number of function evaluations where applicable. The corresponding training-budget protocol is reported with each experiment. These results therefore provide paired relative evidence under the stated controls, rather than a complete causal isolation of the forward-generator choice.

Table 7: Extended linear-generator results on CIFAR-10 and CelebA-HQ-64.

(a) CIFAR-10

a	b	λ_1	T_{conv}	D_{disc}	FID↓	T_{spend}
$\frac{1}{2}$	$-0.25x$	0.25	1000	0.03	4.76	52h
$\frac{1}{2}$	$-0.5x$	0.50	1000	0.03	4.44	45h
$\frac{1}{2}$	$-x$	1.00	825	0.03	**3.15**	35h
$\frac{1}{2}$	$-2x$	2.00	750	0.03	3.19	30h
$\frac{1}{2}$	$-5x$	5.00	525	0.03	4.30	27h
$\frac{1}{2}$	$-5.25x$	5.25	512	0.03	4.48	27h
$\frac{1}{2}$	$-5.5x$	5.50	500	0.03	4.63	26h
$\frac{1}{2}$	$-7x$	7.00	425	0.03	5.47	26h
$\frac{1}{2}$	$-8x$	8.00	375	0.03	5.92	24h
$\frac{1}{2}$	$-10x$	10.00	350	0.03	6.33	**20h**

(b) CelebA-HQ-64

a	b	λ_1	T_{conv}	D_{disc}	FID↓	T_{spend}
$\frac{1}{2}$	$-0.25x$	0.25	1000	0.04	4.02	136h
$\frac{1}{2}$	$-0.5x$	0.50	1000	0.03	3.82	104h
$\frac{1}{2}$	$-x$	1.00	800	0.03	**3.34**	67h
$\frac{1}{2}$	$-2x$	2.00	750	0.03	3.67	61h
$\frac{1}{2}$	$-3x$	3.00	650	0.03	3.78	58h
$\frac{1}{2}$	$-4x$	4.00	550	0.03	3.80	55h
$\frac{1}{2}$	$-4.5x$	4.50	525	0.03	3.97	55h
$\frac{1}{2}$	$-4.75x$	4.75	515	0.03	4.45	55h
$\frac{1}{2}$	$-5x$	5.00	500	0.03	5.03	51h
$\frac{1}{2}$	$-10x$	10.00	300	0.03	7.41	**45h**

E.2 Distance-to-stationarity metric

The theoretical results bound the distance between the noised law μ_t and the stationary distribution Π. In experiments, we use a discretized empirical proxy because the exact high-dimensional densities are unavailable. Given samples $\{x_i(t)\}_{i=1}^n$ from the forward process at time step t and samples $\{y_i\}_{i=1}^n$ from the corresponding stationary distribution, we flatten the samples and, when needed, apply the same PCA map to both sets. Let $\tilde{x}_i, \tilde{y}_i \in \mathbb{R}^K$ denote the resulting representations. For each PCA coordinate k, we discretize the two empirical distributions into $M = 30$ bins and denote the resulting probabilities by $p_{k,j}(t)$ and $q_{k,j}$, $j = 1, \dots, M$.

The reported convergence discrepancy $D_{\text{disc}}(t)$ is the averaged discretized total-variation distance

$$D_{\text{disc}}(t) = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{2} \sum_{j=1}^{M} \left| p_{k,j}(t) - q_{k,j} \right|. \tag{18}$$

This metric is used only as an empirical diagnostic for selecting a sufficiently converged forward length; it is not used as a training loss. For the chi-square diagnostics related to Theorem 3, we also compute

$$\widehat{\chi^2}(t) = \frac{1}{K} \sum_{k=1}^{K} \sum_{j=1}^{M} \frac{\left(p_{k,j}(t) - q_{k,j}\right)^2}{q_{k,j} + \varepsilon}, \tag{19}$$

where a small $\varepsilon > 0$ is used only to avoid division by zero in empty empirical bins.

E.3 Choice of T_{conv} and reporting conventions

The convergence step T_{conv} is chosen as the first step after which $D_{\text{disc}}(t)$ reaches the prescribed low-discrepancy regime and remains stable over a short subsequent window. This avoids selecting a step based on a transient fluctuation. Once T_{conv} is fixed, the denoising model is trained and evaluated under the corresponding forward process. FID is reported with the same evaluation convention within each comparison block. The training time T_{spend} is wall-clock time and is used as an empirical measure of computational cost.

F Additional experimental results

F.1 Extended results for linear generators

The linear generator family $a(x) = 1/2$, $b(x) = -kx$ provides the cleanest controlled setting because its principal eigenvalue is analytically equal to k. Consistent with Finding 1, Table 7 extends the main results to additional eigenvalue regimes and FID evaluations using the same SDE solver.

F.2 Matched-λ_1 ablations across generator families

Table 8 complements the first-order descriptor analysis in the main text. We compare different coefficient parameterizations whose estimated principal eigenvalues are matched around $\lambda_1 \approx 1$.

Table 8: Matched-λ_1 ablations across coefficient parameterizations.

Datasets	a	b	λ_1	T_{conv}	D_{disc}	FID↓	T_{spend}
CIFAR-10	0.5	-x	1.00	825	0.03	3.15	35h
CIFAR-10	1	$-x$	1.00	825	0.03	3.23	34h
CIFAR-10	$0.5x^2 + 0.1$	$-1.05x$	1.03	820	0.03	3.12	34h
CIFAR-10	$0.3\|x\|^3 + 0.2$	$-0.95x$	1.02	830	0.03	3.18	36h
CelebA-HQ-64	0.5	$-x$	1.00	800	0.03	3.34	67h
CelebA-HQ-64	1	$-x$	1.00	800	0.03	3.39	67h
CelebA-HQ-64	$0.5x^2 + 0.1$	$-1.05x$	1.03	795	0.03	5.05	89h
CelebA-HQ-64	$0.3\|x\|^3 + 0.2$	$-0.95x$	1.02	805	0.03	5.11	88h

Table 9: Exploratory nonlinear-generator diagnostics on CIFAR-10 and CelebA-HQ-64.

(a) CIFAR-10

a	b	λ_1	T_{conv}	D_{disc}	FID↓	T_{spend}
$\frac{\|x\|+c}{2}$	-0.4x	0.40	800	0.07	3.45	45h
$\frac{\|x\|+c}{2}$	-0.5x	0.50	730	0.05	**3.22**	41h
$\frac{\|x\|+c}{2}$	-0.6x	0.60	680	0.04	3.85	38h
$\frac{\|x\|+c}{2}$	-x	1.00	650	0.03	3.96	35h
$\frac{\|x\|+c}{2}$	-2x	2.00	625	0.03	4.60	**33h**
$\frac{x^2+c}{2}$	-0.5x	0.52	700	0.16	**3.45**	43h
$\frac{x^2+c}{2}$	-0.65x	0.65	650	0.11	4.12	40h
$\frac{x^2+c}{2}$	-x	1.00	625	0.03	4.17	38h
$\frac{x^2+c}{2}$	-2x	2.00	600	0.03	4.85	**35h**

(b) CelebA-HQ-64

a	b	λ_1	T_{conv}	D_{disc}	FID↓	T_{spend}
$\frac{\|x\|+c}{2}$	-0.3x	0.30	900	0.22	4.12	110h
$\frac{\|x\|+c}{2}$	-0.5x	0.50	825	0.10	**3.88**	101h
$\frac{\|x\|+c}{2}$	-0.7x	0.70	750	0.06	5.07	92h
$\frac{\|x\|+c}{2}$	-x	1.00	725	0.03	5.60	**88h**
$\frac{x^2+c}{2}$	-0.4x	0.43	850	0.25	4.05	105h
$\frac{x^2+c}{2}$	-0.5x	0.52	775	0.22	**3.86**	96h
$\frac{x^2+c}{2}$	-0.6x	0.61	725	0.19	5.33	92h
$\frac{x^2+c}{2}$	-x	1.00	700	0.04	5.77	**84h**

Across both CIFAR-10 and CelebA-HQ-64, matched eigenvalues lead to similar T_{conv}, D_{disc}, and training time, supporting the use of λ_1 as a convergence-oriented descriptor. At the same time, FID varies across coefficient forms, especially for nonlinear generators on CelebA-HQ-64. This confirms that λ_1 captures the dominant convergence-rate behavior but does not fully determine sample quality; the coefficient structure and induced forward trajectory still matter.

F.3 Exploratory diagnostics for nonlinear generator families

The nonlinear results in this subsection are included as exploratory spectral diagnostics rather than as primary evidence for efficient score training with arbitrary non-Gaussian forward kernels. They extend the controlled linear studies in the main text by testing whether the qualitative role of the principal eigenvalue persists when the coefficient functions are nonlinear. As shown in Table 9 with $c > 0$, increasing λ_1 within each nonlinear family generally reduces T_{conv} and T_{spend}, consistent with Finding 1. At the same time, FID is non-monotone across spectral regimes, which is consistent with the efficiency–quality trade-off discussed in Finding 3. Thus, these experiments support the use of λ_1 as a convergence-oriented spectral coordinate beyond the linear family, while also showing that nonlinear coefficient structure and the induced forward trajectory still affect generation quality.

F.4 Compatibility details with existing diffusion components

This subsection provides additional results and protocol details for the compatibility experiments in Table 4. We first examine compatibility with representative reverse-time samplers and the effect of sampling budget, and then report setting-specific details for the EDM/EDM2-style, DiT, and flow-matching-style comparisons. These experiments are intended as controlled compatibility checks rather than comprehensive or fully optimized reproductions of the corresponding pipelines.

Table 10: DDIM compatibility on CIFAR-10 under matched NFE.

a	b	λ_1	T_{conv}	NFE	D_{disc}	FID↓	T_{spend}
$\frac{1}{2}$	-0.25x	0.25	1000	50	0.03	5.08	52h
$\frac{1}{2}$	-0.5x	0.50	1000	50	0.03	4.72	45h
$\frac{1}{2}$	-x	1.00	825	50	0.03	**3.38**	35h
$\frac{1}{2}$	-10x	10.00	350	50	0.03	6.57	**20h**

Table 11: DPM-Solver and DPM-Solver++ compatibility under matched NFE.

(a) CIFAR-10, NFE=15

a	b	λ_1	T_{conv}	FID↓ (DPM)	FID↓ (DPM++)
$\frac{1}{2}$	$-0.25x$	0.25	1000	4.89	4.76
$\frac{1}{2}$	$-0.5x$	0.50	1000	4.59	4.44
$\frac{1}{2}$	$-x$	1.00	825	3.74	**3.15**
$\frac{1}{2}$	$-2x$	2.00	750	**3.23**	3.19
$\frac{1}{2}$	$-5x$	5.00	525	4.43	4.30
$\frac{1}{2}$	$-10x$	10.00	350	6.55	6.33

(b) CelebA-HQ-64, NFE=20

a	b	λ_1	T_{conv}	FID↓ (DPM)	FID↓ (DPM++)
$\frac{1}{2}$	$-0.25x$	0.25	1000	4.29	4.06
$\frac{1}{2}$	$-0.5x$	0.50	1000	3.95	3.90
$\frac{1}{2}$	$-x$	1.00	800	**3.88**	**3.54**
$\frac{1}{2}$	$-2x$	2.00	750	3.96	3.73
$\frac{1}{2}$	$-5x$	5.00	500	5.71	5.34
$\frac{1}{2}$	$-10x$	10.00	300	8.09	7.99

Reverse-time sampler compatibility. We provide additional sampler-level results complementing Table 4(a). Within each comparison, the forward generator is varied while the reverse-time sampler and evaluation protocol are held fixed. The results show consistent relative trends with representative samplers, including DDIM and DPM-Solver-type methods, providing evidence that the generator modification can be used with these reverse-time solvers. See Tables 10–11.

NFE and step-count diagnostics. We further examine FID under different reverse-time sampling budgets. Figure 2 studies the low-NFE regime, where limited network-evaluation budgets can make the results sensitive to solver discretization. Figure 3 varies the number of reverse-time integration steps and shows that FID becomes increasingly stable as the step count grows. In these diagnostics, the linear generators are sampled with DPM-Solver, whereas the nonlinear generators use the SDE solver. Across the evaluated budgets, the relative behavior of the moderate spectral regime remains qualitatively consistent with the main comparison.

Training formulations and denoising architectures. Table 4(b) reports controlled within-pair checks under selected training formulations and denoising architectures. Unless otherwise stated, the comparisons use the baseline generator $(a, b) = (1/2, -0.5x)$ with $T_{\text{conv}} = 1000$ and the EGCA generator $(a, b) = (1/2, -x)$ with $T_{\text{conv}} = 825$. The resource-matching convention differs across settings.

For EDM- and EDM2-style training on CIFAR-10, the baseline is trained for 200 epochs and EGCA for $165 = [200 \times 825/1000]$ epochs, following the T_{conv}-proportional reduced-budget convention. Both use batch size 32 and Heun sampling with 50 steps, with FID evaluated from 50,000 generated samples. These comparisons therefore assess relative performance under a prescribed reduced training budget rather than full-convergence performance.

For DiT on ImageNet-256, we compare four paired training checkpoints using 50 baseline epochs and $41 = [50 \times 825/1000]$ EGCA epochs. Both configurations use batch size 2 and DDIM sampling with 250 steps, and FID is evaluated from 50,000 generated samples. The checkpoint-wise results are reported in Table 12. Since the FID trajectories are still improving over the evaluated interval, these results are interpreted as paired architecture-level trajectory evidence rather than as a fully converged DiT benchmark.

The flow-matching-style experiment uses an equal-epoch comparison (300 epochs for both configurations), with batch size 32 and Heun sampling with 50 steps. Since flow matching defines an

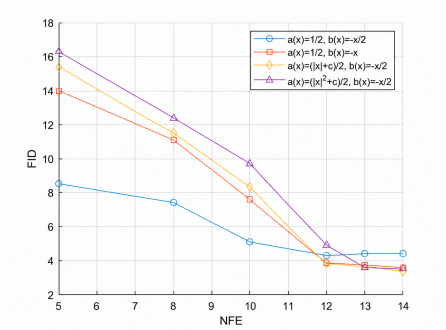


Figure 2: Low-NFE sampling diagnostics on CIFAR-10.

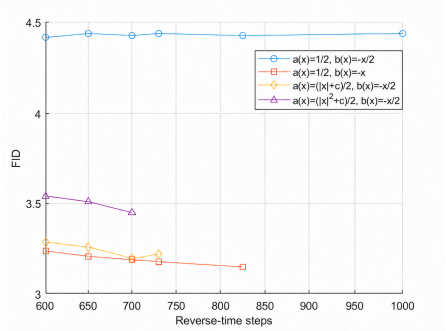


Figure 3: Reverse-time step-count diagnostics on CIFAR-10.

Table 12: Compatibility with DiT on ImageNet-256 under paired training checkpoints. FID1–FID4 denote the FID values measured at the four checkpoint pairs.

a	b	λ_1	T_{conv}	FID1	FID2	FID3	FID4
$\frac{1}{2}$	$-0.5x$	0.50	1000	270	220	198	124
$\frac{1}{2}$	$-x$	1.00	825	243	200	182	98

alternative probability path rather than the stochastic ergodic forward process analyzed here, this experiment is treated only as an empirical boundary comparison.

Interpretation. Taken together, these experiments provide setting-specific compatibility evidence at two levels: the forward-generator modification can be used with representative reverse-time samplers, and similar relative trends are observed under selected training formulations and denoising architectures. The latter experiments use setting-specific budget conventions and should be interpreted as controlled within-pair checks, not as fully converged or state-of-the-art reproductions.

F.5 Complementary diagnostics for T_{conv}, D_{disc}, and stationary effects

This subsection provides complementary diagnostics for the empirical convergence criterion used in the main experiments. We examine the role of the forward horizon and stationary law, and assess D_{disc} as an empirical convergence proxy. These analyses support interpreting λ_1 as a convergence-oriented spectral coordinate rather than a complete predictor of generation quality.

F.5.1 Fixed-T_{conv} controls

To examine whether the observed FID differences can be explained solely by the use of different forward horizons, we fix $T_{\text{conv}} = 1000$ across the evaluated generator configurations. As shown in Table 13, the qualitative non-monotone FID pattern of the main linear sweep is preserved, with intermediate spectral regimes remaining more favorable than both the low and highly aggressive regimes. The nonlinear configurations also retain noticeable FID differences under the common forward horizon. These results indicate that differences in T_{conv} alone do not account for the observed variation in generation quality.

F.5.2 Stationary-law and forward-trajectory diagnostics

We further examine the invariant law, conditional noise variance, and log-SNR trajectory for the linear generators used in the main experiments. Consider the coordinate generator

$$\mathcal{L} = a_0 \frac{\mathrm{d}^2}{\mathrm{d}x^2} - kx \frac{\mathrm{d}}{\mathrm{d}x}, \qquad a_0 > 0,\ k > 0.$$

Then the corresponding forward transition can be written as

$$X_t = \alpha_k(t) X_0 + \sigma(t)\epsilon, \qquad \epsilon \sim \mathcal{N}(0, 1),$$

Table 13: Fixed-T_{conv} control across generator choices on CIFAR-10.

a	b	λ_1	T_{conv}	FID↓	T_{spend}
$\frac{1}{2}$	-0.25x	0.25	1000	4.76	52h
$\frac{1}{2}$	-0.5x	0.50	1000	4.44	45h
$\frac{1}{2}$	-x	1.00	1000	3.15	40h
$\frac{1}{2}$	-2x	2.00	1000	3.19	42h
$\frac{1}{2}$	-5x	5.00	1000	4.27	44h
$\frac{1}{2}$	-10x	10.00	1000	6.32	49h
$\frac{\lvert x\rvert+c}{2}$	-0.5x	0.50	1000	3.22	50h
$\frac{\lvert x\rvert^2+c}{2}$	-0.5x	0.52	1000	3.33	57h

Table 14: Stationary-law control with matched invariant law $\mathcal{N}(0,1)$ and different spectral gaps.

Dataset	a	b	λ_1	T_{conv}	FID↓	T_{spend} ↓
CIFAR-10	$1/2$	$-0.5x$	0.50	1000	4.44	45 h
CIFAR-10	1	$-x$	1.00	825	3.23	34 h
CelebA-HQ-64	$1/2$	$-0.5x$	0.50	1000	3.82	104 h
CelebA-HQ-64	1	$-x$	1.00	800	3.39	67 h

where

$$\alpha_k(t) = e^{-kB(t)}, \quad \sigma^2(t) = \frac{a_0}{k}\left(1 - e^{-2kB(t)}\right), \quad B(t) = \int_0^t \beta(s)\,ds,$$

and the invariant law is

$$\pi = \mathcal{N}\Big(0, \frac{a_0}{k}\Big).$$

If the initial coordinate has variance v_0, the marginal variance is

$$\mathrm{Var}(X_t) = \alpha_k^2(t)v_0 + \sigma^2(t).$$

For the log-SNR values reported in Table 15, we use the common unit-variance reference $v_0 = 1$ across all generator configurations. We use

$$\log \mathrm{SNR}(t) = \log \frac{\alpha_k^2(t)v_0}{\sigma^2(t)}$$

for the corresponding log-SNR trajectory. The same expressions apply coordinate-wise to the separable high-dimensional process considered in our experiments.

Stationary-law control. In the main linear sweep with $a_0 = 1/2$, varying k changes both the spectral gap and the invariant variance. To control for the invariant law, we additionally compare $(a, b) = (1/2, -0.5x)$ and $(1, -x)$, which share the invariant law $\mathcal{N}(0, 1)$ but have $\lambda_1 = 0.5$ and 1. As shown in Table 14, the larger-gap configuration reaches the convergence criterion earlier and yields lower FID on both CIFAR-10 and CelebA-HQ-64. Thus, invariant-law differences alone do not explain the observed quality trend. This comparison does not, however, isolate λ_1 from other coupled changes in the forward process.

Terminal forward statistics. The closed-form expressions above characterize the complete forward trajectory of the linear family. To complement them with values at the actual empirical stopping horizons, Table 15 reports the terminal conditional noise variance and log-SNR for representative low, moderate, and aggressive regimes in the CIFAR-10 main linear sweep.

The terminal statistics differ across the three regimes, showing that the main spectral sweep changes not only the contraction rate but also the forward noise scale and log-SNR at the empirical stopping horizon. Together with the stationary-law control above, these diagnostics support interpreting λ_1 as a convergence-oriented organizing coordinate rather than as an isolated predictor of generation quality.

Table 15: Terminal statistics across representative CIFAR-10 spectral regimes.

a	b	λ_1	T_{conv}	$\sigma^2(T_{\mathrm{conv}})$	$\log \mathrm{SNR}(T_{\mathrm{conv}})$
$1/2$	$-0.5x$	0.5	1000	1.00	-10.12
$1/2$	$-x$	1	825	0.5	-16.00
$1/2$	$-10x$	10	350	0.05	-67.83

Table 16: Distance diagnostics for $(a, b) = (1/2, -0.5x)$ on CIFAR-10.

Distance	$t = 300$	$t = 350$	$t = 500$	$t = 750$	$t = 825$	$t = 1000$
D_{disc}	0.2618	0.2029	0.0862	0.0340	0.0299	0.0298
$\widehat{\chi^2}$	1.3376	0.8362	0.1239	0.0078	0.0054	0.0057
$\widehat{\mathrm{SW}}$	1.0261	0.7140	0.2316	0.0506	0.0402	0.0322

F.5.3 Distance diagnostics

As described in Appendix E.2, D_{disc} is an empirical binned marginal-TV proxy used to select a stable forward horizon T_{conv}. It is not intended as an estimator of the full high-dimensional χ^2 divergence appearing in Theorem 3. To examine whether the stabilization detected by D_{disc} is also reflected by additional diagnostic measures, we evaluate D_{disc}, a binned marginal χ^2 diagnostic, and sliced Wasserstein distance (SWD) along the same forward trajectory for the fixed generator $(a, b) = (1/2, -0.5x)$ on CIFAR-10. SWD is used only as a complementary multivariate diagnostic and does not enter the theoretical convergence bound.

Sliced Wasserstein diagnostic. Let $\widetilde{X}_i(t), \widetilde{Y}_i \in \mathbb{R}^K$ denote the PCA representations of samples from the forward distribution at step t and from the corresponding stationary distribution, respectively. For L random unit directions $\{\theta_\ell\}_{\ell=1}^{L} \subset \mathbb{S}^{K-1}$, define

$$u_{i,\ell}(t) = \theta_\ell^\top \widetilde{X}_i(t), \qquad v_{i,\ell} = \theta_\ell^\top \widetilde{Y}_i.$$

Let $u_{(i),\ell}(t)$ and $v_{(i),\ell}$ denote the corresponding sorted one-dimensional projections. For equal sample size n, we estimate the order-2 sliced Wasserstein distance by

$$\widehat{\mathrm{SW}}(t) = \left[\frac{1}{Ln} \sum_{\ell=1}^{L} \sum_{i=1}^{n} \left(u_{(i),\ell}(t) - v_{(i),\ell} \right)^2 \right]^{1/2}.$$

Unlike the coordinate-wise D_{disc}, SWD probes multiple linear projections of the joint PCA representation and therefore serves as a complementary multivariate diagnostic.

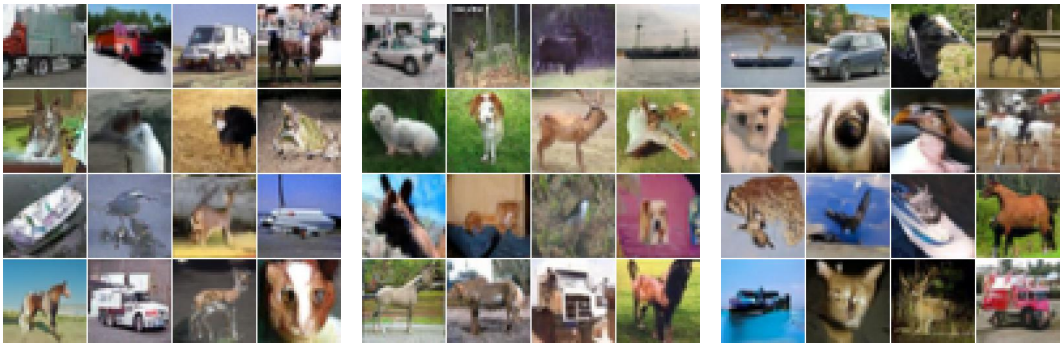
For consistency, the PCA representation is fitted once and reused across all evaluated forward steps, and the same set of random projection directions is used throughout the comparison. We use the same sample set and PCA dimension as for D_{disc}, together with $L = 256$ fixed random projection directions. Table 16 reports the three diagnostics at representative forward steps.

All three diagnostics decrease substantially along the forward trajectory and become nearly stable in the late-stage region. Over the six reported steps, D_{disc} has Pearson correlation $r = 0.987$ and Spearman correlation $\rho = 0.943$ with the binned marginal χ^2 diagnostic, and $r = 0.998$ and $\rho = 1.000$ with SWD. These correlations are used only as descriptive diagnostics for this single controlled trajectory: the reported steps arise from the same forward process and should not be regarded as independent experimental replicates or as evidence that the three discrepancies are generally equivalent. In particular, D_{disc} and $\widehat{\chi^2}$ are marginal empirical diagnostics rather than measurements of the full high-dimensional $\chi^2(\mu_t \| \Pi)$ appearing in Theorem 3, while SWD is an additional empirical multivariate check rather than a quantity entering our theoretical convergence analysis.

Finally, D_{disc} is used only as a forward-convergence diagnostic, not as a predictor of FID; similar stabilized values of D_{disc} can correspond to noticeably different generation quality across generator configurations.

Table 17: Complementary generation-quality metrics across representative CIFAR-10 spectral regimes.

Regime	(a, b)	λ_1	FID↓	sFID↓	IS↑	Precision↑	Recall↑
Low	$(1/2, -0.5x)$	0.5	4.44	5.12	9.1	0.64	0.59
Moderate	$(1, -x)$	1	3.23	4.15	9.3	0.65	0.60
Aggressive	$(1/2, -10x)$	10	6.33	6.14	8.8	0.64	0.56



(a) Low, $\lambda_1 = 0.5$ (b) Moderate, $\lambda_1 = 1$ (c) Aggressive, $\lambda_1 = 10$

Figure 4: Qualitative comparison across the CIFAR-10 spectral regimes in Table 17.

F.6 Complementary quality metrics and qualitative samples

To complement the FID results in the main text, we evaluate representative low, moderate, and aggressive spectral regimes on CIFAR-10 using sFID [28], Inception Score (IS) [35], Precision, and Recall [20] in Table 17. The same three generator configurations are also used for the seed-controlled qualitative comparison in Figure 4.

Each configuration is evaluated using 50,000 generated images under the same evaluation protocol. The complementary metrics broadly agree with the FID comparison: the moderate regime performs best overall, whereas the aggressive regime degrades several quality and coverage metrics. No single metric is intended to fully characterize generation quality.

For the qualitative comparison, we use the same low, moderate, and aggressive configurations as in Table 17. The same SDE solver is used across the three configurations, and the same random seeds are used at corresponding image positions. The visual differences are relatively subtle in this finite sample grid. Accordingly, these examples are intended as qualitative illustrations, while distribution-level differences are assessed using the metrics in Table 17.

G Limitations and broader impacts

Limitations. Our theoretical analysis focuses on structured high-dimensional forward processes with coordinate-wise separable generators. This setting enables a clean reduction to a shared one-dimensional coordinate generator, but it does not provide a complete spectral theory for arbitrary nonseparable, non-reversible, or kinetic diffusion dynamics. The principal eigenvalue should also be interpreted as a convergence-oriented descriptor rather than a complete predictor of generation quality; overly aggressive spectral regimes can reduce convergence cost while degrading FID.

Broader impacts. EGCA may reduce the computational cost and energy use of diffusion-model training by guiding forward-process design through a computable spectral quantity. This can make diffusion modeling more accessible, but efficiency improvements may also lower the cost of generating synthetic media. The method does not introduce a new data source or content modality, yet it should be used with standard safeguards for generative models, including provenance tracking, misuse evaluation, and deployment-specific content policies.

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: We demonstrate the trend of model performance with different L-generator, verifying the effectiveness of EGCA.

   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: See Appendix G.

   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: See Appendix A .

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 methodology and experimental settings needed to reproduce the main results are provided in Section 4, Appendix E and F.

   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: The experiments use publicly available datasets, but the implementation code is not publicly available.

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: The datasets, model architectures, training-budget conventions, sampling settings, and evaluation protocols are described in Section 4, Appendix E and F.

   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: [No]

   Justification: Due to the computational cost of the main training experiments, each configuration was evaluated using a single training run; therefore, we do not report error bars or statistical significance tests.

   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: See Section 4.

   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 reviewed the NeurIPS Code of Ethics and confirm that the research reported in this paper conforms to it.

   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: See Appendix G.

    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: This paper poses no such risks.

    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: All datasets, models, and external implementations used in this work are properly cited and used in accordance with their respective licenses and terms of use.

    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: Our core method development in 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.