BAYESIAN NONPARAMETRIC DIFFUSION MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Score-based diffusion models have achieved strong generative performance across a wide range of applications, but their perturbation law, which determines the noisy states and score targets, is typically fixed to a standard Gaussian form. Recent work shows that alternative perturbation laws can improve modeling in specific data regimes. However, existing approaches either prescribe a perturbation family or learn restricted forward-process parameterizations, leaving open how to adapt the perturbation law itself to a particular learner and training regime. We address this problem with a Bayesian nonparametric (BNP) diffusion model that learns the perturbation distribution directly. We show that standard denoising score matching (DSM) contains a perturbation-dependent irreducible term and derive a corrected criterion for generalized Bayesian learning of the perturbation distribution. Theoretically, we show that the corrected criterion consistently targets finite-resource score risk, that the resulting generalized posterior concentrates around the corresponding optimal perturbation laws, and that the same score risk controls reverse-process error. Empirically, controlled experiments show that BNP adapts across Gaussian and non-Gaussian regimes and that the correction enables more accurate comparisons of score error across perturbation laws. On CIFAR-10 and HRRR, comparisons based on estimated corrected score risk show that the preferred perturbation varies across tasks and training budgets. Together, these findings support treating the perturbation law as a learnable component of diffusion models, adapted to the score learner and available training resources.

1 INTRODUCTION

Score-based diffusion models have achieved strong generative performance across a wide range of applications (Ho et al., 2020; Song et al., 2021; Karras et al., 2022). Under denoising score matching, the forward perturbation law determines the noisy states and score targets used during training (Vincent, 2011; Song et al., 2021), which directly shapes the score-estimation problem faced by the learner. A growing body of work has therefore reconsidered the forward corruption process. Heavy-tailed and non-Gaussian alternatives replace standard Gaussian perturbations with Lévy, α-stable, Student-t, or multiplicative constructions, improving modeling in settings such as imbalanced or heavy-tailed data (YOON et al., 2023; Shariatian et al., 2025; Pandey et al., 2025; Gruhlke et al., 2026). Other approaches redesign the corruption mechanism itself, for example through blur–noise mixtures, or learn the forward process within a parameterized model class (Hsueh et al., 2025; Bartosh et al., 2024; 2026). These results establish that the forward corruption need not be treated as a fixed Gaussian construction.

However, existing approaches either prescribe a particular perturbation or corruption family, or learn the forward process within a predefined model class and training objective. They do not resolve which perturbation distribution is appropriate for a particular score learner under finite training resources. Different perturbation laws induce different noisy inputs and score targets, and therefore define different score-estimation problems for the learner. How well these problems can be learned depends on the hypothesis class, the amount of training data, and the accuracy reached during optimization (Zhang et al., 2025; Kumar et al., 2025; Wang et al., 2024; Xiao et al., 2025; Han et al., 2024; Gupta et al., 2024). Consequently, the perturbation law that minimizes score risk need not be the same across learners or training regimes, making a single fixed perturbation law generally unsuitable as a universal design choice.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

We therefore define the finite-resource optimal perturbation law for a score-training procedure $\mathcal{A}_s$ using m training samples as

$$G^{\star}_{m,\mathcal{A}_s} \in \arg\min_{G\in\mathcal{G}} \mathfrak{R}_{m,\mathcal{A}_s}(G),$$

where $\mathcal{G}$ is the admissible class of perturbation laws and $\mathfrak{R}_{m,\mathcal{A}_s}(G)$ is the expected score risk achieved by the estimator returned by $\mathcal{A}_s$ after training under perturbation law G. Accordingly, the preferred perturbation law depends on the score learner and the available training resources, and should be adapted to the specific score-training regime.

Learning the perturbation law introduces a second difficulty: the standard denoising score matching (DSM) objective is not directly comparable across different perturbation laws. For a fixed perturbation law, DSM provides a tractable objective whose population minimizer is the corresponding marginal score (Vincent, 2011; Song et al., 2021). However, this agreement of minimizers does not imply that the objective values can be compared across different laws. In our setting, for any fixed lower time cutoff $\tau > 0$, the population DSM loss over $[\tau, T]$ decomposes as

$$\mathcal{L}_{\mathrm{DSM},\tau}(G, s) = \mathcal{R}_{\mathrm{score},\tau}(G, s) + \mathcal{C}_\tau(G),$$

where $\mathcal{C}_\tau(G)$ is a G-dependent irreducible variance term in the DSM target, capturing random variation caused by the sampled clean data and noise realizations. For a fixed G, it only adds a constant to the DSM objective, whereas across different perturbation laws it can change their relative DSM values and thus distort the ranking by score risk. Learning the perturbation law therefore requires a criterion that removes this G-dependent irreducible component before candidate laws are compared.

We address both difficulties with a Bayesian nonparametric (BNP) diffusion model. We introduce a trajectory-level latent scale

$$\Lambda \sim G, \qquad dX_t = f(X_t, t)\, dt + \sqrt{\Lambda}\, g(t)\, dW_t,$$

where Λ is drawn once per trajectory from the perturbation law G and held fixed throughout the forward process. For the VP process, $X^{\lambda}_t = a_t X_0 + \sqrt{\lambda v_t} Z$. Here, λ has a direct interpretation as a trajectory-level corruption scale. Specifically, at the same diffusion time, a larger λ produces a lower effective signal-to-noise ratio and stronger corruption along a trajectory, while a smaller λ preserves more signal, leading to different levels of difficulty in learning the score required for reverse denoising. The law G therefore determines how training trajectories are distributed across different levels of corruption. A continuous scale distribution spreads finite training resources across a continuum of noise levels, making it hard to concentrate learning on the corruption levels most useful for a particular learner and budget, while a discrete perturbation law allows many trajectories to share the same scales and enables repeated training on a smaller set of useful noise levels. We therefore use a Bayesian nonparametric prior to adapt these shared scales and their frequencies without fixing their number in advance.

Conditional on Λ, the process remains a Markov diffusion with the standard Brownian structure, while the VP specialization retains Gaussian conditional transitions, thereby preserving the conditional score and reverse-diffusion structure needed for standard diffusion analysis. To compare perturbation laws, we remove the G-dependent irreducible term from the DSM objective, yielding a corrected DSM criterion that targets score risk. We normalize G so that $\mathbb{E}_G[\Lambda] = 1$, keeping the average perturbation scale fixed while allowing its mass to be redistributed across different noise levels. We then combine the normalized Bayesian nonparametric prior (Ferguson, 1973; Sethuraman, 1994) with the corrected criterion through generalized Bayesian learning (Bissiri et al., 2016). The resulting generalized posterior takes the form

$$\Pi_n(dG) \propto \exp\left\{-\kappa_n \widehat{\mathfrak{R}}^{\mathrm{CF}}_{n,\mathcal{A}}(G)\right\} \Pi_0(dG),$$

where Π_0 is the prior, $\kappa_n > 0$ is the generalized-Bayes learning rate, and $\widehat{\mathfrak{R}}^{\mathrm{CF}}_{n,\mathcal{A}}(G)$ is the empirical cross-fitted corrected risk. This construction adapts the distribution of perturbation scales to the finite-resource score-learning problem.

Theoretically, we show that the corrected DSM criterion consistently targets finite-resource score risk, that the generalized posterior concentrates around the corresponding optimal perturbation laws, and that the same score risk controls reverse-process error. Empirically, experiments across synthetic settings, CIFAR-10, and HRRR show that BNP adapts across Gaussian and non-Gaussian regimes

and that the preferred perturbation varies across learning tasks and training budgets. These results support treating the perturbation law as a learnable component of the diffusion model and show that BNP can concentrate training trajectories on perturbation scales suited to the score learner and available budget.

Our contributions are summarized as follows. (1) We formulate finite-resource perturbation design through the score risk attained by a specified learner under finite training resources. (2) We introduce a latent-scale Bayesian nonparametric diffusion model with a corrected DSM criterion and establish consistency, posterior concentration, and reverse-process guarantees. (3) Synthetic, CIFAR-10, and HRRR experiments show that perturbation preferences vary across learning problems and training budgets, and that BNP adapts without fixing a parametric perturbation family in advance.

2 Bayesian Nonparametric Perturbation Learning

Let $X_0 \sim P_0$ denote the data distribution on $\mathbb{R}^d$. We consider the continuous-time forward diffusion

$$dX_t = f(X_t, t)\,dt + g(t)\,dW_t, \qquad t \in [0, T],$$

where W_t is a standard Wiener process (Song et al., 2021). We focus on the variance-preserving (VP) specialization (Ho et al., 2020), for which $a_t = \exp\{-\frac{1}{2}\int_0^t \beta(r)\,dr\}$ and $v_t = 1 - a_t^2$.

We extend this construction by treating the perturbation scale as a latent random variable. A single scale is sampled for each forward trajectory, while its distribution is shared across the population of trajectories. This turns the perturbation law into an unknown probability distribution that can be learned from data. We place a Bayesian nonparametric prior on this distribution, construct a score-risk criterion that is comparable across perturbation laws, and use the resulting loss to update the prior.

2.1 Latent-Scale Forward Diffusion

Let G be a probability distribution on positive scales. For each trajectory, we first draw $\Lambda \sim G$ and then evolve

$$dX_t = f(X_t, t)\,dt + \sqrt{\Lambda}\,g(t)\,dW_t.$$

The same value of Λ is retained throughout the trajectory. Conditional on $\Lambda = \lambda$, the process is an ordinary diffusion with diffusion coefficient $\sqrt{\lambda}\,g(t)$. Throughout, a superscript λ denotes a quantity conditional on the latent-scale realization $\Lambda = \lambda$. For the VP diffusion, the conditional transition is

$$X_t^\lambda = a_t X_0 + \sqrt{\lambda v_t}\,Z, \qquad Z \sim \mathcal{N}(0, I_d).$$

We write p_t^λ for the marginal density of X_t^λ and $s_t^\lambda(x) := \nabla_x \log p_t^\lambda(x)$ for its score. The corresponding transition density conditional on $X_0 = x_0$ will be denoted by $q_t^\lambda(\cdot \mid x_0)$. The score model is therefore conditioned on the latent scale, $s_\theta(x, t, \lambda)$. Conditioning on $\Lambda = \lambda$ yields an ordinary Markov diffusion, with Gaussian conditional transitions in the VP specialization. Because the same latent scale is shared across an entire trajectory, marginalizing it out need not preserve the Markov property of X_t alone. We therefore formulate score learning and reverse-time dynamics conditionally on Λ, or equivalently on the augmented Markov state (X_t, Λ). This conditional formulation also allows the reverse process to retain the same diffusion coefficient as the forward model.

2.2 Bayesian Nonparametric Prior over Perturbation Laws

The latent-scale construction reduces perturbation design to inference over the distribution G. Rather than fixing the number or locations of relevant scales in advance, we start from a Dirichlet process prior (Ferguson, 1973). Using its stick-breaking representation (Sethuraman, 1994), let the raw Dirichlet-process draw be

$$\widetilde{G} = \sum_{k=1}^{\infty} \pi_k \delta_{\widetilde{\lambda}_k}, \qquad \widetilde{G} \sim \mathrm{DP}(\alpha, H),$$

where α is the concentration parameter, H is the base distribution, and δ_a denotes a point mass at a.

A global rescaling of all atoms changes the overall noise magnitude without changing their relative structure. To keep this global scale fixed, we normalize the atoms by the mean of $\widetilde{G}$. Let $\bar{\lambda}(\widetilde{G}) := \sum_{k=1}^{\infty} \pi_k \widetilde{\lambda}_k$, and define $\lambda_k := \widetilde{\lambda}_k / \bar{\lambda}(\widetilde{G})$ and $G := \sum_{k=1}^{\infty} \pi_k \delta_{\lambda_k}$. The normalized distribution satisfies $\mathbb{E}_G[\Lambda] = 1$ by construction. Thus, the average perturbation variance remains matched to the unit-scale baseline, while G can still adapt its mass across different scales. The resulting G is the image of a Dirichlet-process draw under this normalization, rather than a Dirichlet-process draw itself. In computation, we use a truncated stick-breaking representation and apply the same normalization to the retained atoms. All subsequent learning is performed on the normalized distribution G.

2.3 Noise-Corrected Score Risk

For a fixed perturbation law G and scale λ, denoising score matching uses the conditional perturbation score (Vincent, 2011) $Y_t^\lambda := \nabla_{x_t} \log q_t^\lambda(X_t^\lambda \mid X_0)$, which satisfies $\mathbb{E}[Y_t^\lambda \mid X_t^\lambda] = s_t^\lambda(X_t^\lambda)$. For a fixed lower time cutoff $\tau > 0$, we use the weighting induced by the reverse diffusion and define

$$\mathcal{L}_{\mathrm{DSM},\tau}(G, s) := \frac{1}{2} \int G(d\lambda) \int_\tau^T \lambda g(t)^2 \, \mathbb{E}\left[\left\| s(X_t^\lambda, t, \lambda) - Y_t^\lambda \right\|^2 \right] dt.$$

The corresponding marginal score risk is

$$\mathcal{R}_{\mathrm{score},\tau}(G, s) := \frac{1}{2} \int G(d\lambda) \int_\tau^T \lambda g(t)^2 \, \mathbb{E}\left[\left\| s(X_t^\lambda, t, \lambda) - s_t^\lambda(X_t^\lambda) \right\|^2 \right] dt.$$

The conditional-expectation identity gives $\mathcal{L}_{\mathrm{DSM},\tau}(G, s) = \mathcal{R}_{\mathrm{score},\tau}(G, s) + \mathcal{C}_\tau(G)$, where

$$\mathcal{C}_\tau(G) := \frac{1}{2} \int G(d\lambda) \int_\tau^T \lambda g(t)^2 \, \mathbb{E}\left[\left\| Y_t^\lambda - s_t^\lambda(X_t^\lambda) \right\|^2 \right] dt.$$

For a fixed G, $\mathcal{C}_\tau(G)$ does not depend on the score model and therefore does not change its population minimizer. Across different perturbation laws, however, $\mathcal{C}_\tau(G)$ generally changes with G. Raw DSM values are therefore not directly comparable when the perturbation law itself is being learned. We remove this term and define the corrected criterion

$$\mathcal{L}_{\mathrm{corr},\tau}(G, s) := \mathcal{L}_{\mathrm{DSM},\tau}(G, s) - \mathcal{C}_\tau(G) = \mathcal{R}_{\mathrm{score},\tau}(G, s).$$

For the VP model, the correction can be estimated through ordinary denoising. Let $r_t^\lambda(x) := \mathbb{E}[X_0 \mid X_t^\lambda = x]$ denote the Bayes denoiser. Since $Y_t^\lambda = -(X_t^\lambda - a_t X_0)/(\lambda v_t)$, we have $Y_t^\lambda - s_t^\lambda(X_t^\lambda) = \frac{a_t}{\lambda v_t}\left[X_0 - r_t^\lambda(X_t^\lambda)\right]$. Hence

$$\mathcal{C}_\tau(G) = \frac{1}{2} \int G(d\lambda) \int_\tau^T \frac{a_t^2 g(t)^2}{\lambda v_t^2} \, \mathbb{E}\left[\left\| X_0 - r_t^\lambda(X_t^\lambda) \right\|^2 \right] dt.$$

In practice, we estimate r_t^λ with a denoising model $r_\psi(x, t, \lambda)$. We use cross-fitting (Chernozhukov et al., 2018): the observations are partitioned into folds, and the denoiser used to evaluate one fold is trained only on the remaining folds. This separates fitting from evaluation and yields the empirical criterion used below. Section 3 establishes its consistency.

2.4 Generalized Bayesian Learning of the Perturbation Law

The score estimator obtained in practice depends on both G and the available training data. We therefore evaluate each perturbation law using the score estimator produced by the same learning procedure used at training time. Let $\mathcal{A}_s$ denote the procedure used to train the score estimator and $\mathcal{A}_r$ the procedure used to train the denoiser for estimating the correction term, and write $\mathcal{A} := (\mathcal{A}_s, \mathcal{A}_r)$.

Let n denote the total number of observations used to construct the empirical criterion, and partition them into K folds. For fold k, let $\widehat{s}_G^{(-k)}$ and $\widehat{r}_G^{(-k)}$ denote the score estimator and denoiser trained without fold k. We evaluate both quantities on the held-out fold and combine the results as

$$\widehat{\mathfrak{R}}_{n,\mathcal{A}}^{\mathrm{CF}}(G) := \frac{1}{K} \sum_{k=1}^{K} \left[\widehat{\mathcal{L}}_{\mathrm{DSM},\tau}^{(k)}\big(G, \widehat{s}_G^{(-k)}\big) - \widehat{\mathcal{C}}_\tau^{(k)}\big(G, \widehat{r}_G^{(-k)}\big) \right].$$

Thus, $\widehat{\mathfrak{R}}^{\mathrm{CF}}_{n,\mathcal{A}}(G)$ is the empirical cross-fitted corrected risk for perturbation law G under the procedures collected in $\mathcal{A}$.

We use this criterion to update the prior over G through generalized Bayesian inference. Generalized Bayes combines a prior with a loss when the inferential target is defined by risk rather than by a likelihood (Bissiri et al., 2016). The generalized posterior is

$$\Pi_n(dG) \propto \exp\left\{-\kappa_n \widehat{\mathfrak{R}}^{\mathrm{CF}}_{n,\mathcal{A}}(G)\right\} \Pi_0(dG),$$

where Π_0 and Π_n denote the prior and generalized posterior over perturbation laws G, respectively, and $\kappa_n > 0$ is the generalized-Bayes learning rate. During BNP adaptation, the perturbation law and the score estimator are learned through alternating updates. Given the current variational distribution over G, the score estimator is updated using perturbations drawn from that distribution, after which the variational parameters of G are updated using held-out corrected risk. Thus, perturbation-law learning and score learning remain coupled throughout the adaptation stage. The resulting procedure learns the full perturbation distribution rather than selecting a single scale or a fixed parametric family. Its target is determined by the score learner and the available training resources. Section 3 shows that the cross-fitted criterion converges to the corresponding algorithm-dependent score risk and that the generalized posterior concentrates around its finite-sample minimizers.

3 Theoretical Guarantees

We now establish the statistical properties of the learning procedure introduced in Section 2. The analysis follows the same order as the method itself. We first show that the noise correction recovers the population score risk. We then connect the cross-fitted criterion to the score risk achieved by the actual training algorithm and derive concentration of the generalized posterior. Next, we isolate the finite-sample terms through which the perturbation law affects score estimation. Finally, we translate score error into reverse-process and generation error.

3.1 Population Corrected-Risk Consistency

We first study the corrected criterion in Section 2.3 at the population level and show that it recovers the corresponding score risk uniformly over perturbation laws. For a perturbation law G, define the profiled score risk, obtained by optimizing over the score class $\mathcal{F}$, as $\mathfrak{R}^{\mathrm{prof}}(G) := \inf_{s\in\mathcal{F}} \mathcal{R}_{\mathrm{score},\tau}(G, s)$. Its empirical counterpart is obtained by minimizing the empirical DSM loss and subtracting the cross-fitted estimate of $\mathcal{C}_\tau(G)$. For the VP diffusion, the correction can be written in terms of the Bayes denoiser $r_t^\lambda(x) = \mathbb{E}[X_0 \mid X_t^\lambda = x]$. This representation allows the irreducible DSM variance to be estimated by ordinary denoising regression. The following result shows that the resulting profiled criterion is uniformly consistent over the perturbation family.

Theorem 3.1 (Consistency of the profiled corrected risk)**.** *Suppose the DSM loss class satisfies a uniform law of large numbers, the cross-fitted correction class is Glivenko–Cantelli (Shorack & Wellner, 2009), and the denoiser is uniformly consistent in the weighted L^2 norm. Then*

$$\sup_{G\in\mathcal{G}} \left|\widehat{\mathfrak{R}}^{\mathrm{prof}}_n(G) - \mathfrak{R}^{\mathrm{prof}}(G)\right| \xrightarrow{\mathbb{P}} 0.$$

The key point is that the correction removes the part of DSM that varies with G but does not measure marginal score error. Thus the corrected criterion supports valid population comparison across perturbation laws. The full assumptions and proof are given in Appendix A.

3.2 Algorithm-Dependent Risk and BNP Posterior Concentration

The population profiled risk describes approximation within a function class, but the score used in practice is the output of a finite-sample training procedure. We therefore make the learning algorithm explicit. Let $\widehat{s}_{G,D_m} = \mathcal{A}_s(D_m; G)$ be the score estimator returned by $\mathcal{A}_s$ after training on a sample D_m of size m under perturbation law G. Its expected score risk is

$$\mathfrak{R}_{m,\mathcal{A}_s}(G) := \mathbb{E}_{D_m,\mathcal{A}_s}\left[\mathcal{R}_{\mathrm{score},\tau}\left(G, \widehat{s}_{G,D_m}\right)\right],$$

where the expectation is over the training sample and any randomness in the training procedure. We denote the full minimizer set by $\mathcal{G}^\star_{m,\mathcal{A}_s} := \arg\min_{G\in\mathcal{G}} \mathfrak{R}_{m,\mathcal{A}_s}(G)$, and use $G^\star_{m,\mathcal{A}_s} \in \mathcal{G}^\star_{m,\mathcal{A}_s}$ for

an individual finite-resource optimal perturbation law. The cross-fitted criterion from Section 2.4 is designed to estimate this algorithm-dependent risk. Each score estimator and denoiser is trained on the complementary folds and evaluated only on held-out observations. The resulting consistency statement is the direct link between the practical learning rule and the finite-resource target.

Theorem 3.2 (Consistency of the cross-fitted algorithm risk)**.** *Assume that the held-out DSM and correction losses satisfy uniform laws of large numbers, the cross-fitted denoiser is uniformly consistent, and the trained score risks concentrate uniformly around their expectation. If m_n denotes the number of observations used to train each learner when cross-fitting a total sample of size n, then*

$$\sup_{G\in\mathcal{G}} \left|\widehat{\mathfrak{R}}^{\mathrm{CF}}_{n,\mathcal{A}}(G) - \mathfrak{R}_{m_n,\mathcal{A}_s}(G)\right| \xrightarrow{\mathbb{P}} 0.$$

We now combine this result with the generalized posterior in Section 2.4. For $\varepsilon > 0$, let $\Delta_{n,\varepsilon}$ be the risk gap between the minimizer set $\mathcal{G}^\star_{m_n,\mathcal{A}_s}$ and perturbation laws at least ε away from it, where dist denotes the metric on the admissible law space $\mathcal{G}$.

Theorem 3.3 (Generalized BNP concentration)**.** *Suppose $\Delta_{n,\varepsilon} > 0$ for all sufficiently large n. If the uniform estimation error is $o_{\mathbb{P}}(\Delta_{n,\varepsilon})$, the prior assigns sufficient mass to near-optimal perturbation laws, and $\kappa_n\Delta_{n,\varepsilon} \to \infty$, then*

$$\Pi_n\left(G : \mathrm{dist}\left(G, \mathcal{G}^\star_{m_n,\mathcal{A}_s}\right) \geq \varepsilon\right) \xrightarrow{\mathbb{P}} 0.$$

Hence the posterior tracks the perturbation laws that minimize the expected score risk of the actual learner at the corresponding training size. Unlike a resource-independent population target, this target is allowed to vary with the available training resources. The complete error decomposition and posterior proof are given in Appendices B and C.

3.3 Finite-Sample Score Estimation

The previous result identifies the target of perturbation learning. We next examine why that target can differ across perturbation laws. To make the finite-sample effect explicit, fix G. Let T denote a sampled diffusion time and define the score-regression input $\mathsf{Z} := (X^\Lambda_{\mathsf{T}}, \mathsf{T}, \Lambda)$. For a feature map $\phi(\mathsf{Z}) \in \mathbb{R}^{d_\phi}$, consider the d_ϕ-dimensional score class $\mathcal{F}_{d_\phi} = \left\{s_\Theta : s_\Theta(\mathsf{Z}) = \Theta\phi(\mathsf{Z}), \quad \Theta \in \mathbb{R}^{d\times d_\phi}\right\}$. Let $\mathcal{E}_{\mathrm{app},d_\phi}(G)$ denote the weighted approximation error of the exact conditional score within this class, and let ℓ_G denote the corresponding leverage constant. Writing $S^\star(\mathsf{Z})$ for the exact conditional score, the DSM target admits the regression decomposition $Y = S^\star(\mathsf{Z}) + \xi$ with $\mathbb{E}[\xi \mid \mathsf{Z}] = 0$, and its weighted conditional variance is exactly $\mathcal{C}_\tau(G)$. The finite-sample risk therefore contains both approximation and estimation effects.

Theorem 3.4 (Finite-sample score-risk bound)**.** *Suppose the weighted population Gram matrix is positive definite and the weighted leverage is bounded by $\ell_G d_\phi$. Let $\widehat{\Theta}_G$ denote the weighted least-squares coefficient estimated from m training samples. Then, for every $\gamma \in (0,1)$, with probability at least $1 - \gamma - d_\phi \exp\left\{-\frac{c_0 m}{\ell_G d_\phi}\right\}$, the corresponding score estimator satisfies*

$$\mathcal{R}_{\mathrm{score},\tau}\left(G, s_{\widehat{\Theta}_G}\right) \leq \mathcal{E}_{\mathrm{app},d_\phi}(G) + \frac{4\ell_G d_\phi}{m\gamma}\left[\mathcal{E}_{\mathrm{app},d_\phi}(G) + \mathcal{C}_\tau(G)\right],$$

where $c_0 > 0$ is a universal constant.

The bound separates three effects of the perturbation law. First, $\mathcal{E}_{\mathrm{app},d_\phi}(G)$ measures how well the induced score field can be represented by the chosen function class. Second, $\mathcal{C}_\tau(G)$ measures the variance of the DSM target. Third, ℓ_G captures the conditioning of the weighted regression problem. Thus a change in G can alter both the approximation and estimation difficulty of finite-sample score learning. In the realizable case, $\mathcal{E}_{\mathrm{app},d_\phi}(G) = 0$, the bound reduces to a pure estimation term proportional to $\mathcal{C}_\tau(G)/m$. The proof is given in Appendix D.

3.4 Reverse-Process Guarantees

We finally connect the score risk to the generative process. This follows a broader line of diffusion theory linking score-estimation error to sampling accuracy (Chen et al., 2023). For a fixed latent

scale λ, write $u = T - t$ and define the reverse-time state $\bar{X}_u^\lambda := X_{T-u}^\lambda$. The exact reverse diffusion (Anderson, 1982) is

$$d\bar{X}_u^\lambda = \left[-f(\bar{X}_u^\lambda, T-u) + \lambda g(T-u)^2 s_{T-u}^\lambda(\bar{X}_u^\lambda)\right] du + \sqrt{\lambda}\, g(T-u)\, d\overline{W}_u.$$

The learned reverse process replaces s_t^λ with $\widehat{s}_t^\lambda$ while keeping the same latent scale and diffusion coefficient. Let $\mathbf{P}_G^{\mathrm{rev}}$ and $\mathbf{Q}_{G,\widehat{s}}^{\mathrm{rev}}$ denote the corresponding joint laws of the latent scale and the entire reverse trajectory. The learned reverse process may be initialized from a reference terminal density $\widetilde{p}_T^\lambda$ rather than the exact terminal marginal p_T^λ. Define the resulting terminal mismatch by $\mathcal{D}_T(G) := \int \mathrm{KL}\left(p_T^\lambda \,\|\, \widetilde{p}_T^\lambda\right) G(d\lambda)$.

Theorem 3.5 (Reverse-path relative entropy)**.** *Under the regularity conditions stated in Appendix E,*

$$\mathrm{KL}\left(\mathbf{P}_G^{\mathrm{rev}} \,\|\, \mathbf{Q}_{G,\widehat{s}}^{\mathrm{rev}}\right) = \mathcal{D}_T(G) + \mathcal{R}_{\mathrm{score},\tau}(G, \widehat{s}).$$

The remaining term $\mathcal{D}_T(G)$ is controlled by the terminal diffusion level. For the VP model with $\widetilde{p}_T^\lambda = \mathcal{N}(0, \lambda v_T I_d)$ and $\lambda \geq \underline{\lambda} > 0$, $\sup_{G\in\mathcal{G}} \mathcal{D}_T(G) \leq \frac{a_T^2 \mathbb{E}\|X_0\|^2}{2\underline{\lambda} v_T}$. Finally, let P_τ^G and $Q_{\tau,G,\widehat{s}}$ denote the endpoint distributions of the exact and learned reverse processes, respectively. The data-processing inequality (Cover et al., 1991) and Pinsker's inequality (Pinsker, 1964) give $\left\|P_\tau^G - Q_{\tau,G,\widehat{s}}\right\|_{\mathrm{TV}} \leq \sqrt{\frac{1}{2}\left[\mathcal{D}_T(G) + \mathcal{R}_{\mathrm{score},\tau}(G,\widehat{s})\right]}$. Therefore the same score risk used to learn the perturbation law also enters the bound on the discrepancy between the reverse process and its generated endpoint distribution. When the terminal mismatch and score risk vanish, the learned reverse process is consistent with the exact one. All path-space arguments and regularity conditions are collected in Appendix E.

4 Experiments

We evaluate the proposed framework in two stages. First, controlled synthetic experiments isolate the statistical mechanism behind perturbation adaptation. They test whether the perturbation law changes finite-resource score learning, whether raw DSM can identify the preferred law, and whether Bayesian nonparametric inference adapts to different score-learning regimes. We then move to neural diffusion models on real data, where we examine whether the preferred perturbation law continues to depend on the available training resources.

4.1 Synthetic Experiments

Setup. We use a controlled one-dimensional model to isolate how the perturbation law affects finite-resource score learning. The clean distribution is $X_0 = \mu B$, where $B \in \{-1, +1\}$ with equal probability, and at a fixed VP time $X_t^\lambda = a_t X_0 + \sqrt{\lambda v_t}\, Z$, $Z \sim \mathcal{N}(0,1)$. The score learner is a one-parameter fixed-feature model whose coefficient is shared across all values of λ. Hence changing G changes the conditional scores that must be represented by the same finite-capacity learner. Because both the marginal score and the weighted least-squares learner are available analytically in this model, we can compute the finite-resource score risk $\mathfrak{R}_{m,\mathcal{A}_s}(G)$ exactly. We use $m = 2048$ throughout the main synthetic study and all synthetic experiments were run on CPU, with full implementation details in Appendix F.

Finite-resource perturbation regimes. The finite-sample analysis in Section 3.3 predicts that the optimal perturbation law can depend on the score-learning problem. To test this prediction under a controlled comparison, we construct two regimes with the same data model and score learner but different finite-resource optima. We vary the signal-to-noise parameter $c = a_t^2\mu^2/v_t$ while keeping the VP construction fixed. At $c = c_\star = 2.0839$, the exact $m = 2048$ optimum within the normalized perturbation family is $G_{2048,\mathcal{A}_s}^\star \approx 0.65061\,\delta_{0.48824} + 0.34939\,\delta_{1.95298}$, with score risk 0.81085. We refer to this setting as the *non-Gaussian-optimal regime*. As a control, we set $c = c_\star/2 = 1.0420$, for which the finite-resource optimum is the unit-scale Gaussian law $G = \delta_1$. We refer to this setting as the *Gaussian-optimal regime*.

Bayesian nonparametric adaptation. Having established two regimes with different finite-resource optimal perturbation laws, we next test whether BNP adapts its perturbation law accordingly. We use the same BNP prior and optimization protocol in both regimes. As shown in Table 1, in

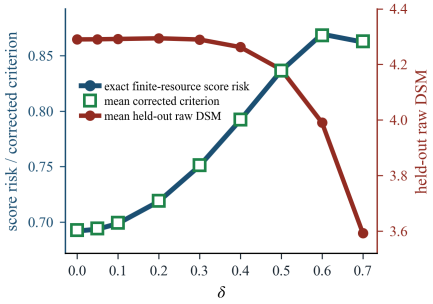

Figure 1: Corrected and raw DSM objectives over the perturbation family G_h.

Table 1: BNP adaptation in two controlled synthetic regimes. The table reports corrected score risk, the relative risk change of BNP against the Gaussian baseline, and the BNP posterior-averaged variance of the latent scale, $\mathbb{E}_{\Pi_{\mathrm{VI}}}[\mathrm{Var}_G(\Lambda)]$, where Π_{VI} denotes the variational posterior over perturbation laws G.

	Corrected score risk		BNP adaptation	
Regime	Gaussian	BNP	Risk change vs. Gaussian	Posterior-avg. scale variance
Non-Gaussian-optimal	0.9342	**0.8281**	**-11.4%**	0.4576
Gaussian-optimal	0.6920	0.6929	**+0.13%**	0.0014

the non-Gaussian-optimal regime, BNP reduces the score risk from 0.9342 for the unit-scale Gaussian baseline to 0.8281, an improvement of 11.4%. This recovers 86.0% of the improvement from the Gaussian baseline to the optimum risk of 0.81085. In the Gaussian-optimal regime, BNP attains a risk of 0.6929, compared with 0.6920 for the Gaussian optimum, corresponding to only 0.13% relative regret. The learned perturbation laws show the same adaptation. The posterior-averaged scale variance $\mathbb{E}_{\Pi_{\mathrm{VI}}}[\mathrm{Var}_G(\Lambda)]$ is 0.4576 in the non-Gaussian-optimal regime but only 0.00142 in the Gaussian-optimal regime. Thus BNP learns a dispersed scale distribution when a non-Gaussian perturbation is beneficial, while remaining close to the unit-scale law when Gaussian perturbation is already optimal.

Raw DSM and corrected risk. We next isolate the role of the correction in Section 2.3. For the symmetric family $G_h = \frac{1}{2}\delta_{1-h} + \frac{1}{2}\delta_{1+h}$, where h determines the separation of the two scales around their fixed mean of one. Figure 1 compares three quantities. The blue curve is the exact finite-resource score risk $\mathfrak{R}_{2048,\mathcal{A}_s}(G_h)$, which provides the reference ordering available in this controlled model. The red curve is mean held-out raw DSM, evaluated on validation samples independent of score fitting. The green markers show the mean corrected criterion $\widehat{\mathcal{L}}_{\mathrm{DSM},\tau}(G_h) - \widehat{\mathcal{C}}_\tau(G_h)$, where the correction is evaluated with the analytically available Bayes denoiser. This isolates the effect of removing the G-dependent DSM variance term. The consistency result in Section 3.1 extends the argument to an estimated denoiser. As h increases, the exact score risk rises while raw DSM decreases. The discrepancy is caused by the G-dependent irreducible term $\mathcal{C}_\tau(G)$, whose reduction dominates the increase in score error and reverses the apparent preference of the raw DSM objective. The corrected criterion instead closely tracks the exact score risk. Moreover, this discrepancy is not confined to the minimizer. Among the 36 pairwise comparisons induced by the nine candidate laws, raw DSM reverses 29 relative to the exact score-risk ordering, whereas the corrected criterion reverses none. This ordering matters for generalized Bayes because posterior mass depends on the relative risks across the candidate perturbation laws, not only on the location of a single minimizer.

4.2 Real-World Experiments

Setup. We evaluate on CIFAR-10 (Krizhevsky et al., 2009) and the High-Resolution Rapid Refresh (HRRR) weather dataset (Dowell et al., 2022). For HRRR, we consider vertically integrated liquid water (VIL) and pressure vertical velocity at level 20 (w20), which represent distinct meteorological fields and provide a complementary test beyond natural images. For every task, we compare the unit-scale Gaussian law, Student-t perturbations with $\nu \in \{3, 5, 10\}$, and the adaptive BNP perturbation model under two training budgets, 25k and 100k optimization steps. All final score models within each task use the same network architecture, optimizer, initialization protocol, and score-training budget, so the final comparison isolates the perturbation law under matched score-training resources. The BNP model is learned using the corrected DSM criterion introduced in Section 2.3, and model selection uses only held-out validation data. During BNP adaptation, the score model and the variational perturbation law are updated in an alternating manner. After adaptation, the learned perturbation law is frozen, the adaptation-stage score model is discarded, and a fresh score model is trained from scratch under the frozen law using the same 25k or 100k score-training budget as the fixed-law baselines. For CIFAR-10, the freshly trained model is evaluated on the untouched test set, and we report corrected test risk. For HRRR, we report per-dimension

Table 2: Corrected score risk across CIFAR-10 and HRRR under two training budgets. Lower is better. Values in **bold** indicate the best result for each task and training regime.

			Fixed perturbation laws				Adaptive
	Regime	Updates	Gaussian	$t_{\nu=3}$	$t_{\nu=5}$	$t_{\nu=10}$	BNP
CIFAR-10	Constrained	25k	0.0504	0.2711	0.1243	0.0627	**0.0421**
	Full	100k	0.0011	**0.0002**	0.0008	0.0016	0.0013
HRRR(VIL)	Constrained	25k	**0.5381**	1.8559	1.1371	0.6576	0.5618
	Full	100k	0.0889	0.1767	0.1178	0.0810	**0.0544**
HRRR(w20)	Constrained	25k	0.8725	1.3507	1.0182	0.8898	**0.8411**
	Full	100k	0.2805	0.4631	0.3443	0.3156	**0.2792**

corrected validation risk, normalized by the 64×64 spatial dimensions. Additional training and evaluation details are given in Appendix G.

Perturbation preferences vary across tasks and training budgets. Table 2 shows that no single perturbation choice is preferred across all tasks and training regimes. On CIFAR-10, BNP attains the lowest corrected risk at 25k updates, 0.0421, compared with 0.0504 for Gaussian and 0.0627 for the best Student-t candidate. This corresponds to a 16.6% reduction relative to the Gaussian baseline. At 100k updates, however, the preference shifts to $t_{\nu=3}$, which attains a corrected risk of 0.0002.

HRRR exhibits a different pattern. For VIL, Gaussian is preferred at 25k updates with a corrected risk of 0.5381, whereas BNP becomes preferred at 100k updates with a risk of 0.0544. For w20, BNP attains the lowest corrected risk under both budgets, with 0.8411 at 25k and 0.2792 at 100k. Thus, changing the learning task can change the preferred perturbation under the same training budget, while changing the available training budget can change the preference within the same task. The preferred perturbation is therefore determined by the interaction between the learning problem and the available training resources, rather than by either factor alone, which motivates an adaptive treatment of the perturbation law rather than selecting a single fixed baseline in advance.

The CIFAR-10 Student-t baselines provide a finer example of the same dependence. Their ordering reverses from $t_{\nu=10} < t_{\nu=5} < t_{\nu=3}$ at 25k updates to $t_{\nu=3} < t_{\nu=5} < t_{\nu=10}$ at 100k updates. Hence, even after fixing a parametric perturbation family, its preferred tail parameter can remain resource-dependent. Together, these results support the finite-resource target $G^{\star}_{m,\mathcal{A}_s}$: the perturbation law preferred by a score learner can depend jointly on the learning problem and the available training resources. BNP addresses this additional specification by learning the perturbation distribution directly rather than prescribing a single parametric family or tail parameter in advance.

5 Discussion

Across our controlled experiments, BNP adapts across Gaussian and non-Gaussian settings, while the corrected DSM criterion removes the perturbation-dependent irreducible term that can otherwise distort comparisons between candidate laws. On CIFAR-10 and HRRR, the perturbation preferred by corrected score risk changes across learning tasks and training budgets, and the preferred Student-t tail parameter can change even within a fixed parametric family, further demonstrating the importance of adapting the perturbation law to the learning setting. This behavior can be interpreted directly through the latent-scale construction. The learned law G specifies the trajectory-level corruption scales used during training and how frequently training trajectories are assigned to them. A concentrated G emphasizes a narrow range of corruption levels, while a more dispersed G allocates training across multiple noise regimes. The preferred allocation therefore depends on the score learner, the data, and the available training budget. Together with our theoretical results, these findings connect adaptive allocation across perturbation scales to finite-resource score risk and reverse-process accuracy. The current BNP construction is still constrained by the support and approximation family used for G, which may limit its ability to represent the regime-specific optimal perturbation law. Extending the framework to richer perturbation families and broader score-learning settings is a natural direction for future work.

AI USE STATEMENT

In this work, we used generative AI tools to assist with research ideation and execution, retrieval and discovery of related work, drafting and polishing portions of the manuscript, and developing and checking mathematical arguments and proofs. We did not use generative AI tools to generate synthetic datasets, and other disclosure categories not relevant to this work are not applicable. All AI-assisted research ideas, mathematical claims, derivations, proofs, citations, experimental procedures, and manuscript text were reviewed and verified by the authors. Literature suggested through AI-assisted retrieval was checked against the original sources, and mathematical arguments and experimental results were independently inspected for correctness. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

Complete proofs and experimental details are provided in the appendix. Anonymized source code for reproducing the main experiments and evaluations is included in the supplementary material.

REFERENCES

Brian D.O. Anderson. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982. ISSN 0304-4149. doi: https://doi.org/10.1016/0304-4149(82)90051-5. URL https://www.sciencedirect.com/science/article/pii/0304414982900515.

D. F. Andrews and C. L. Mallows. Scale mixtures of normal distributions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(1):99–102, 1974. doi: https://doi.org/10.1111/j.2517-6161.1974.tb00989.x. URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1974.tb00989.x.

Grigory Bartosh, Dmitry Vetrov, and Christian A. Naesseth. Neural flow diffusion models: Learnable forward process for improved diffusion modelling. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 73952–73985. Curran Associates, Inc., 2024. doi: 10.52202/079017-2352. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/871a8ccb9232487366feb5e2d9069915-Paper-Conference.pdf.

Grigory Bartosh, Teodora Pandeva, Sushrut Karmalkar, and Javier Zazo. Forward-learned discrete diffusion: Learning how to noise to denoise faster. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), *International Conference on Learning Representations*, volume 2026, pp. 94995–95006, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/99469f6f2526d9733535ab7b19732e03-Paper-Conference.pdf.

P. G. Bissiri, C. C. Holmes, and S. G. Walker. A general framework for updating belief distributions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 78(5):1103–1130, 11 2016. ISSN 1369-7412. doi: 10.1111/rssb.12158. URL https://doi.org/10.1111/rssb.12158.

Sitan Chen, Sinho Chewi, Jerry Li, Yuanzhi Li, Adil Salim, and Anru Zhang. Sampling Is as Easy as Learning the Score: Theory for Diffusion Models with Minimal Data Assumptions. In *International Conference on Learning Representations*, 2023. URL https://mlanthology.org/iclr/2023/chen2023iclr-sampling/.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 02 2018. ISSN 1368-4221. doi: 10.1111/ectj.12097. URL https://doi.org/10.1111/ectj.12097.

Thomas M Cover, Joy A Thomas, and John Kieffer. *Elements of information theory*, volume 2. wiley New York, 1991.

David C. Dowell, Curtis R. Alexander, Eric P. James, Stephen S. Weygandt, Stanley G. Benjamin, Geoffrey S. Manikin, Benjamin T. Blake, John M. Brown, Joseph B. Olson, Ming Hu, Tatiana G. Smirnova, Terra Ladwig, Jaymes S. Kenyon, Ravan Ahmadov, David D. Turner, Jeffrey D. Duda, and Trevor I. Alcott. The high-resolution rapid refresh (hrrr): An hourly updating convection-allowing forecast model. part i: Motivation and system description. *Weather and Forecasting*, 37(8):1371 – 1395, 2022. doi: 10.1175/WAF-D-21-0151.1. URL https://journals.ametsoc.org/view/journals/wefo/37/8/WAF-D-21-0151.1.xml.

Thomas S. Ferguson. A Bayesian Analysis of Some Nonparametric Problems. *The Annals of Statistics*, 1(2):209 – 230, 1973. doi: 10.1214/aos/1176342360. URL https://doi.org/10.1214/aos/1176342360.

Robert Gruhlke, Valentin Resseguier, and Merveille Talla. Multiplicative diffusion models: Beyond gaussian latents. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), *International Conference on Learning Representations*, volume 2026, pp. 108324–108412, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/b05cd0262afc98ce2536bba9629545ee-Paper-Conference.pdf.

Shivam Gupta, Aditya Parulekar, Eric Price, and Zhiyang Xun. Improved sample complexity bounds for diffusion model training. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 40976–41012. Curran Associates, Inc., 2024. doi: 10.52202/079017-1296. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/480e563034dbb6d1dd622d8eab7d129b-Paper-Conference.pdf.

Yinbin Han, Meisam Razaviyayn, and Renyuan Xu. Neural network-based score estimation in diffusion models: Optimization and generalization. In B. Kim, Y. Yue, S. Chaudhuri, K. Fragkiadaki, M. Khan, and Y. Sun (eds.), *International Conference on Learning Representations*, volume 2024, pp. 42520–42558, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/file/ba3354bcfeae4f166a8bfe75443ac8f7-Paper-Conference.pdf.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf.

Hao-Chien Hsueh, Wen-Hsiao Peng, and Ching-Chun Huang. Warm diffusion: Recipe for blur-noise mixture diffusion models. In Y. Yue, A. Garg, N. Peng, F. Sha, and R. Yu (eds.), *International Conference on Learning Representations*, volume 2025, pp. 13020–13037, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/22a25fc3da528794d52664dacc7bd470-Paper-Conference.pdf.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (eds.), *Advances in Neural Information Processing Systems*, volume 35, pp. 26565–26577. Curran Associates, Inc., 2022. doi: 10.52202/068431-1926. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/a98846e9d9cc01cfb87eb694d946ce6b-Paper-Conference.pdf.

Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. 2009.

Syamantak Kumar, Dheeraj Nagaraj, and Purnamrita Sarkar. Dimension-free score matching and time bootstrapping for diffusion models. In D. Belgrave, C. Zhang, H. Lin, R. Pascanu, P. Koniusz, M. Ghassemi, and N. Chen (eds.), *Advances in Neural Information Processing Systems*, volume 38, Main Conference, pp. 32224–32277. Curran Associates, Inc., 2025. doi: 10.52202/085713-1082. URL https://proceedings.neurips.cc/paper_files/paper/2025/file/2e50c5ed85b16421919d741760e4ada0-Paper-Conference.pdf.

Kushagra Pandey, Jaideep Pathak, Yilun Xu, Stephan Mandt, Michael Pritchard, Arash Vahdat, and Morteza Mardani. Heavy-tailed diffusion models. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=tozlOEN4qp.

Ethan Perez, Florian Strub, Harm De Vries, Vincent Dumoulin, and Aaron Courville. Film: Visual reasoning with a general conditioning layer. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

M.S. Pinsker. *Information and Information Stability of Random Variables and Processes*. Holden-Day series in time series analysis. Holden-Day, 1964. ISBN 9780816268047. URL https://books.google.com/books?id=cyJrD6wa2pYC.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241. Springer, 2015.

Jayaram Sethuraman. A constructive definition of dirichlet priors. *Statistica sinica*, pp. 639–650, 1994.

Dario Shariatian, Umut Simsekli, and Alain Oliviero Durmus. Denoising levy probabilistic models. In Y. Yue, A. Garg, N. Peng, F. Sha, and R. Yu (eds.), *International Conference on Learning Representations*, volume 2025, pp. 96991–97024, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/f12c9da400fec2594a05e37b0507fd00-Paper-Conference.pdf.

Galen R Shorack and Jon A Wellner. *Empirical processes with applications to statistics*. SIAM, 2009.

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. URL https://openreview.net/forum?id=PxTIG12RRHS.

Pascal Vincent. A connection between score matching and denoising autoencoders. *Neural Computation*, 23(7):1661–1674, 07 2011. ISSN 0899-7667. doi: 10.1162/NECO_a_00142. URL https://doi.org/10.1162/NECO_a_00142.

Yuqing Wang, Ye He, and Molei Tao. Evaluating the design space of diffusion-based generative models. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 19307–19352. Curran Associates, Inc., 2024. doi: 10.52202/079017-0610. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/227404a13d20898dec2018ebe368b202-Paper-Conference.pdf.

Quan Xiao, Hui Yuan, A F M Saif, Gaowen Liu, Ramana Rao Kompella, Mengdi Wang, and Tianyi Chen. A first-order generative bilevel optimization framework for diffusion models. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 68535–68558. PMLR, 13–19 Jul 2025. URL https://proceedings.mlr.press/v267/xiao25i.html.

EUN BI YOON, Keehun Park, Sungwoong Kim, and Sungbin Lim. Score-based generative models with lévy processes. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (eds.), *Advances in Neural Information Processing Systems*, volume 36, pp. 40694–40707. Curran Associates, Inc., 2023. doi: 10.52202/075280-1772. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/8011b23e1dc3f57e1b6211ccad498919-Paper-Conference.pdf.

Zhaoyu Zhang, Yang Hua, Guanxiong Sun, Hui Wang, and Seán Mcloone. Training diffusion-based generative models with limited data. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.),

Proceedings of the 42nd International Conference on Machine Learning, volume 267 of *Proceedings of Machine Learning Research*, pp. 74950–74965. PMLR, 13–19 Jul 2025. URL https://proceedings.mlr.press/v267/zhang25z.html.

A Population Corrected-Risk Consistency

This section proves Theorem 3.1.

A.1 Common-randomness representation

Let $\mathcal{S} \subset (0, \infty)$ denote the scale domain supporting the candidate perturbation laws in $\mathcal{G}$. For the variance-preserving forward diffusion, conditional on the latent scale $\lambda \in \mathcal{S}$,

$$X_t^\lambda = a_t X_0 + \sqrt{\lambda v_t}\, Z, \qquad v_t = 1 - a_t^2, \qquad Z \sim \mathcal{N}(0, I_d). \tag{1}$$

where Z is independent of X_0.

Let ρ be a probability density on $[\tau, T]$ satisfying

$$0 < \underline{\rho} \le \rho(t) \le \overline{\rho} < \infty, \qquad t \in [\tau, T],$$

and draw $\mathsf{T} \sim \rho$ independently of (X_0, Z). Set $U := (X_0, \mathsf{T}, Z)$. For every candidate $\lambda \in \mathcal{S}$, the perturbed observation X_{T}^λ is therefore a measurable function of the same random variable U.

Define

$$\omega_{\mathrm{DSM}}(t, \lambda) := \frac{\lambda g(t)^2}{2\rho(t)}.$$

Then, for every integrable function h,

$$\mathbb{E}_{\mathsf{T}}\left[\omega_{\mathrm{DSM}}(\mathsf{T}, \lambda) h(\mathsf{T})\right] = \frac{1}{2} \int_\tau^T \lambda g(t)^2 h(t)\, dt.$$

For $s \in \mathcal{F}$, define

$$\ell_{\lambda,s}^{\mathrm{DSM}}(U) := \omega_{\mathrm{DSM}}(\mathsf{T}, \lambda) \left\| s(X_{\mathsf{T}}^\lambda, \mathsf{T}, \lambda) - Y_{\mathsf{T}}^\lambda \right\|^2.$$

Let $\mathbb{P}_U$ denote the law of U and define the empirical measure $\mathbb{P}_n := n^{-1} \sum_{i=1}^n \delta_{U_i}$. Then

$$\mathcal{L}_{\mathrm{DSM},\tau}(G, s) = \int \mathbb{P}_U \ell_{\lambda,s}^{\mathrm{DSM}}\, G(d\lambda),$$

$$\widehat{\mathcal{L}}_{\mathrm{DSM},\tau,n}(G, s) = \int \mathbb{P}_n \ell_{\lambda,s}^{\mathrm{DSM}}\, G(d\lambda).$$

A.2 Conditional-variance representation of the correction

Under equation 1, the Gaussian transition score is

$$Y_t^\lambda = -\frac{X_t^\lambda - a_t X_0}{\lambda v_t}. \tag{2}$$

Let $r_t^\lambda(x) := \mathbb{E}[X_0 \mid X_t^\lambda = x]$ denote the Bayes denoiser.

Using Fisher's identity, $s_t^\lambda(X_t^\lambda) = \mathbb{E}[Y_t^\lambda \mid X_t^\lambda]$. From equation 2,

$$Y_t^\lambda - s_t^\lambda(X_t^\lambda) = \frac{a_t}{\lambda v_t} \left[X_0 - r_t^\lambda(X_t^\lambda)\right].$$

Define

$$\omega_{\mathrm{corr}}(t, \lambda) := \frac{a_t^2 g(t)^2}{2\rho(t)\lambda v_t^2}.$$

Then the scale-specific correction is

$$C_\tau(\lambda) := \mathbb{E}\left[\omega_{\mathrm{corr}}(\mathsf{T}, \lambda) \left\| X_0 - r_{\mathsf{T}}^\lambda(X_{\mathsf{T}}^\lambda) \right\|^2\right],$$

and therefore

$$\mathcal{C}_\tau(G) = \int C_\tau(\lambda)\, G(d\lambda).$$

A.3 CROSS-FITTED ESTIMATION OF THE CORRECTION TERM

Partition $\{1, \dots, n\}$ into a fixed number K of folds $I_1, \dots, I_K$. For each fold k, let $\widehat{r}^{(-k)}$ be a denoiser trained without observations in I_k.

For a generic denoiser r, define

$$\ell^{\mathrm{corr}}_{\lambda,r}(U) := \omega_{\mathrm{corr}}(\mathsf{T}, \lambda) \left\| X_0 - r(X^\lambda_\mathsf{T}, \mathsf{T}, \lambda) \right\|^2 .$$

For $i \in I_k$, evaluate $\ell^{\mathrm{corr}}_{\lambda,\widehat{r}^{(-k)}}(U_i)$.

Define the scale-specific cross-fitted correction estimator

$$\widehat{C}_{\tau,n}(\lambda) := \frac{1}{n} \sum_{k=1}^{K} \sum_{i \in I_k} \ell^{\mathrm{corr}}_{\lambda,\widehat{r}^{(-k)}}(U_i),$$

and hence

$$\widehat{\mathcal{C}}_{\tau,n}(G) := \int \widehat{C}_{\tau,n}(\lambda)\, G(d\lambda).$$

Assumption A.1 (Uniform law for the DSM loss)**.** *The class $\mathscr{L}_{\mathrm{DSM}} := \{\ell^{\mathrm{DSM}}_{\lambda,s} : \lambda \in \mathcal{S},\, s \in \mathcal{F}\}$ is $\mathbb{P}_U$-Glivenko–Cantelli. Equivalently,*

$$\Delta^{\mathrm{DSM}}_n := \sup_{\lambda \in \mathcal{S}} \sup_{s \in \mathcal{F}} \left| (\mathbb{P}_n - \mathbb{P}_U) \ell^{\mathrm{DSM}}_{\lambda,s} \right| \xrightarrow{\mathbb{P}} 0.$$

Assumption A.2 (Uniform law for the correction loss)**.** *There exists a deterministic class $\mathcal{M}$ such that, with probability tending to one, $\widehat{r}^{(-k)} \in \mathcal{M}$ for every k, and the class $\mathscr{L}_{\mathrm{corr}} := \{\ell^{\mathrm{corr}}_{\lambda,r} : \lambda \in \mathcal{S},\, r \in \mathcal{M}\}$ is $\mathbb{P}_U$-Glivenko–Cantelli. In particular, with $\mathbb{P}_{I_k}$ denoting the empirical measure on fold I_k,*

$$\Delta^{\mathrm{corr}}_n := \max_{1 \le k \le K} \sup_{\lambda \in \mathcal{S}} \sup_{r \in \mathcal{M}} \left| (\mathbb{P}_{I_k} - \mathbb{P}_U) \ell^{\mathrm{corr}}_{\lambda,r} \right| \xrightarrow{\mathbb{P}} 0.$$

Assumption A.3 (Uniform cross-fitted denoiser consistency)**.** *For the Bayes denoiser r^λ_t, define*

$$\varepsilon^{\mathrm{den}}_n := \max_{1 \le k \le K} \sup_{\lambda \in \mathcal{S}} \mathbb{E}\left[\omega_{\mathrm{corr}}(\mathsf{T}, \lambda) \left\| \widehat{r}^{(-k)}(X^\lambda_\mathsf{T}, \mathsf{T}, \lambda) - r^\lambda_\mathsf{T}(X^\lambda_\mathsf{T}) \right\|^2 \right] \xrightarrow{\mathbb{P}} 0.$$

Lemma A.4 (Consistency of the cross-fitted correction)**.** *Under Assumptions A.2 and A.3,*

$$\sup_{G \in \mathcal{G}} \left| \widehat{\mathcal{C}}_{\tau,n}(G) - \mathcal{C}_\tau(G) \right| \le \Delta^{\mathrm{corr}}_n + \varepsilon^{\mathrm{den}}_n + o_{\mathbb{P}}(1),$$

and therefore

$$\sup_{G \in \mathcal{G}} \left| \widehat{\mathcal{C}}_{\tau,n}(G) - \mathcal{C}_\tau(G) \right| \xrightarrow{\mathbb{P}} 0.$$

Proof. Fix a fold k and scale λ. Conditional on the data outside I_k, $\widehat{r}^{(-k)}$ is fixed and independent of the validation observations in I_k.

Let $M := \widehat{r}^{(-k)}(X^\lambda_\mathsf{T}, \mathsf{T}, \lambda)$ and $M^\star := r^\lambda_\mathsf{T}(X^\lambda_\mathsf{T})$. Then $X_0 - M = (X_0 - M^\star) + (M^\star - M)$, and hence

$$\|X_0 - M\|^2 = \|X_0 - M^\star\|^2 + \|M - M^\star\|^2 + 2(X_0 - M^\star)^\top (M^\star - M).$$

By the definition of conditional expectation,

$$\mathbb{E}\left[X_0 - M^\star \mid X^\lambda_\mathsf{T}, \mathsf{T}, \lambda \right] = 0.$$

Since $\omega_{\mathrm{corr}}(\mathsf{T}, \lambda)(M^\star - M)$ is measurable with respect to $(X^\lambda_\mathsf{T}, \mathsf{T}, \lambda)$, the cross term has expectation zero. Therefore

$$\mathbb{P}_U \ell^{\mathrm{corr}}_{\lambda,\widehat{r}^{(-k)}} = C_\tau(\lambda) + \mathbb{E}\left[\omega_{\mathrm{corr}}(\mathsf{T}, \lambda) \|M - M^\star\|^2 \right].$$

The second term is bounded uniformly over (k, λ) by $\varepsilon^{\mathrm{den}}_n$.

Moreover, on the event $\widehat{r}^{(-k)} \in \mathcal{M}$, Assumption A.2 gives

$$\sup_{\lambda \in \mathcal{S}} \left| \mathbb{P}_{I_k} \ell^{\mathrm{corr}}_{\lambda, \widehat{r}^{(-k)}} - \mathbb{P}_U \ell^{\mathrm{corr}}_{\lambda, \widehat{r}^{(-k)}} \right| \le \Delta^{\mathrm{corr}}_n.$$

Since K is fixed, averaging over folds yields

$$\sup_{\lambda \in \mathcal{S}} \left| \widehat{C}_{\tau,n}(\lambda) - C_\tau(\lambda) \right| \le \Delta^{\mathrm{corr}}_n + \varepsilon^{\mathrm{den}}_n + o_{\mathbb{P}}(1).$$

Finally, for every $G \in \mathcal{G}$,

$$\left| \widehat{\mathcal{C}}_{\tau,n}(G) - \mathcal{C}_\tau(G) \right| = \left| \int \left[\widehat{C}_{\tau,n}(\lambda) - C_\tau(\lambda) \right] G(d\lambda) \right| \le \sup_{\lambda \in \mathcal{S}} \left| \widehat{C}_{\tau,n}(\lambda) - C_\tau(\lambda) \right|.$$

Taking the supremum over $G \in \mathcal{G}$ proves the claim. □

A.4 Uniform consistency of the profiled criterion

Recall

$$\mathfrak{R}^{\mathrm{prof}}(G) = \inf_{s \in \mathcal{F}} \mathcal{L}_{\mathrm{DSM},\tau}(G, s) - \mathcal{C}_\tau(G)$$

and

$$\widehat{\mathfrak{R}}^{\mathrm{prof}}_n(G) = \inf_{s \in \mathcal{F}} \widehat{\mathcal{L}}_{\mathrm{DSM},\tau,n}(G, s) - \widehat{\mathcal{C}}_{\tau,n}(G).$$

Proof of Theorem 3.1. For any two real-valued functions F_1 and F_2 on a common domain,

$$|\inf F_1 - \inf F_2| \le \sup |F_1 - F_2|.$$

Hence, for any fixed G,

$$\left| \inf_{s \in \mathcal{F}} \widehat{\mathcal{L}}_{\mathrm{DSM},\tau,n}(G, s) - \inf_{s \in \mathcal{F}} \mathcal{L}_{\mathrm{DSM},\tau}(G, s) \right| \le \sup_{s \in \mathcal{F}} \left| \widehat{\mathcal{L}}_{\mathrm{DSM},\tau,n}(G, s) - \mathcal{L}_{\mathrm{DSM},\tau}(G, s) \right|.$$

Furthermore,

$$\begin{aligned} \sup_{s \in \mathcal{F}} \left| \widehat{\mathcal{L}}_{\mathrm{DSM},\tau,n}(G, s) - \mathcal{L}_{\mathrm{DSM},\tau}(G, s) \right| &= \sup_{s \in \mathcal{F}} \left| \int (\mathbb{P}_n - \mathbb{P}_U) \ell^{\mathrm{DSM}}_{\lambda,s} \, G(d\lambda) \right| \\ &\le \sup_{\lambda \in \mathcal{S}} \sup_{s \in \mathcal{F}} \left| (\mathbb{P}_n - \mathbb{P}_U) \ell^{\mathrm{DSM}}_{\lambda,s} \right| = \Delta^{\mathrm{DSM}}_n. \end{aligned}$$

Therefore,

$$\sup_{G \in \mathcal{G}} \left| \widehat{\mathfrak{R}}^{\mathrm{prof}}_n(G) - \mathfrak{R}^{\mathrm{prof}}(G) \right| \le \Delta^{\mathrm{DSM}}_n + \sup_{G \in \mathcal{G}} \left| \widehat{\mathcal{C}}_{\tau,n}(G) - \mathcal{C}_\tau(G) \right|.$$

Applying Lemma A.4 gives

$$\sup_{G \in \mathcal{G}} \left| \widehat{\mathfrak{R}}^{\mathrm{prof}}_n(G) - \mathfrak{R}^{\mathrm{prof}}(G) \right| \le \Delta^{\mathrm{DSM}}_n + \Delta^{\mathrm{corr}}_n + \varepsilon^{\mathrm{den}}_n + o_{\mathbb{P}}(1).$$

Assumptions A.1–A.3 imply that every term on the right-hand side converges to zero in probability, proving

$$\sup_{G \in \mathcal{G}} \left| \widehat{\mathfrak{R}}^{\mathrm{prof}}_n(G) - \mathfrak{R}^{\mathrm{prof}}(G) \right| \xrightarrow{\mathbb{P}} 0.$$

□

B Cross-Fitted Estimation of the Algorithm-Dependent Risk

This section proves Theorem 3.2.

Let $I_1, \dots, I_K$ be a fixed balanced K-fold partition of the observations, and write $m_n := n - |I_k|$ for the training-sample size in each fold. For a perturbation law G, let

$$\widehat{s}^{(-k)}_G := \mathcal{A}_s(D_{-k}; G), \qquad \widehat{r}^{(-k)}_G := \mathcal{A}_r(D_{-k}; G)$$

denote the score estimator and denoiser trained without fold I_k. Recall the algorithm-dependent population risk

$$\mathfrak{R}_{m_n, \mathcal{A}_s}(G) := \mathbb{E}_{D_{m_n}, \mathcal{A}_s} \left[\mathcal{R}_{\mathrm{score},\tau} \left(G, \mathcal{A}_s(D_{m_n}; G) \right) \right]. \tag{3}$$

B.1 HELD-OUT DSM AND CORRECTION LOSSES

For fold k, define the held-out DSM loss by

$$\widehat{\mathcal{L}}^{(k)}_{\mathrm{DSM},\tau}\left(G, \widehat{s}^{(-k)}_G\right) := \int G(d\lambda)\, \mathbb{P}_{I_k} \ell^{\mathrm{DSM}}_{\lambda, \widehat{s}^{(-k)}_G},$$

and the held-out correction by

$$\widehat{\mathcal{C}}^{(k)}_{\tau}\left(G, \widehat{r}^{(-k)}_G\right) := \int G(d\lambda)\, \mathbb{P}_{I_k} \ell^{\mathrm{corr}}_{\lambda, \widehat{r}^{(-k)}_G}.$$

The cross-fitted corrected criterion is

$$\widehat{\mathfrak{R}}^{\mathrm{CF}}_{n,\mathcal{A}}(G) := \frac{1}{K}\sum_{k=1}^{K}\left[\widehat{\mathcal{L}}^{(k)}_{\mathrm{DSM},\tau}\left(G, \widehat{s}^{(-k)}_G\right) - \widehat{\mathcal{C}}^{(k)}_{\tau}\left(G, \widehat{r}^{(-k)}_G\right)\right]. \tag{4}$$

B.2 CONDITIONAL EXPECTATION OF THE CORRECTED LOSS

Conditional on D_{-k}, the fitted functions $\widehat{s}^{(-k)}_G$ and $\widehat{r}^{(-k)}_G$ are independent of the held-out observations in I_k. Therefore,

$$\mathbb{E}\left[\widehat{\mathcal{L}}^{(k)}_{\mathrm{DSM},\tau}\left(G, \widehat{s}^{(-k)}_G\right) \,\middle|\, D_{-k}\right] = \mathcal{R}_{\mathrm{score},\tau}\left(G, \widehat{s}^{(-k)}_G\right) + \mathcal{C}_\tau(G). \tag{5}$$

For a generic denoiser $\widehat{r}$, define its weighted excess denoising risk by

$$\mathcal{E}_{\mathrm{den}}(G, \widehat{r}) := \int G(d\lambda)\, \mathbb{E}\Big[\omega_{\mathrm{corr}}(\mathsf{T}, \lambda)\big\|\widehat{r}(X^\lambda_{\mathsf{T}}, \mathsf{T}, \lambda) - r^\lambda_{\mathsf{T}}(X^\lambda_{\mathsf{T}})\big\|^2\Big].$$

The conditional-mean decomposition from Section A.3 then gives

$$\mathbb{E}\left[\widehat{\mathcal{C}}^{(k)}_{\tau}\left(G, \widehat{r}^{(-k)}_G\right) \,\middle|\, D_{-k}\right] = \mathcal{C}_\tau(G) + \mathcal{E}_{\mathrm{den}}\left(G, \widehat{r}^{(-k)}_G\right). \tag{6}$$

Subtracting the two identities yields

$$\mathbb{E}\left[\widehat{\mathcal{L}}^{(k)}_{\mathrm{DSM},\tau} - \widehat{\mathcal{C}}^{(k)}_{\tau} \,\middle|\, D_{-k}\right] = \mathcal{R}_{\mathrm{score},\tau}\left(G, \widehat{s}^{(-k)}_G\right) - \mathcal{E}_{\mathrm{den}}\left(G, \widehat{r}^{(-k)}_G\right). \tag{7}$$

B.3 UNIFORM ERROR DECOMPOSITION

Define the held-out DSM deviation

$$\Delta^{\mathrm{score}}_{n,\mathrm{val}} := \max_{1\le k\le K}\sup_{G\in\mathcal{G}}\left|\widehat{\mathcal{L}}^{(k)}_{\mathrm{DSM},\tau}\left(G, \widehat{s}^{(-k)}_G\right) - \mathcal{L}_{\mathrm{DSM},\tau}\left(G, \widehat{s}^{(-k)}_G\right)\right|,$$

and the held-out correction deviation

$$\Delta^{\mathrm{corr}}_{n,\mathrm{val}} := \max_{1\le k\le K}\sup_{G\in\mathcal{G}}\left|\widehat{\mathcal{C}}^{(k)}_{\tau}\left(G, \widehat{r}^{(-k)}_G\right) - \mathbb{E}\left[\widehat{\mathcal{C}}^{(k)}_{\tau}\left(G, \widehat{r}^{(-k)}_G\right) \mid D_{-k}\right]\right|.$$

Let

$$\varepsilon^{\mathrm{den}}_{m_n} := \max_{1\le k\le K}\sup_{G\in\mathcal{G}}\mathcal{E}_{\mathrm{den}}\left(G, \widehat{r}^{(-k)}_G\right),$$

and define the trained-score risk fluctuation

$$\Delta^{\mathrm{alg}}_{m_n} := \sup_{G\in\mathcal{G}}\left|\frac{1}{K}\sum_{k=1}^{K}\mathcal{R}_{\mathrm{score},\tau}\left(G, \widehat{s}^{(-k)}_G\right) - \mathfrak{R}_{m_n,\mathcal{A}_s}(G)\right|.$$

Lemma B.1 (Uniform error decomposition)**.** *For every n,*

$$\sup_{G\in\mathcal{G}}\left|\widehat{\mathfrak{R}}^{\mathrm{CF}}_{n,\mathcal{A}}(G) - \mathfrak{R}_{m_n,\mathcal{A}_s}(G)\right| \le \Delta^{\mathrm{score}}_{n,\mathrm{val}} + \Delta^{\mathrm{corr}}_{n,\mathrm{val}} + \varepsilon^{\mathrm{den}}_{m_n} + \Delta^{\mathrm{alg}}_{m_n}. \tag{8}$$

Proof. For each fold k, add and subtract $\mathcal{R}_{\mathrm{score},\tau}(G, \widehat{s}_G^{(-k)})$ and the conditional population correction in equation 6. Using equation 5 and equation 6, we obtain

$$\left|\widehat{\mathcal{L}}_{\mathrm{DSM},\tau}^{(k)} - \widehat{\mathcal{C}}_\tau^{(k)} - \mathcal{R}_{\mathrm{score},\tau}\left(G, \widehat{s}_G^{(-k)}\right)\right| \le \Delta_{n,\mathrm{val}}^{\mathrm{score}} + \Delta_{n,\mathrm{val}}^{\mathrm{corr}} + \mathcal{E}_{\mathrm{den}}\left(G, \widehat{r}_G^{(-k)}\right).$$

Averaging over the K folds and taking the supremum over G gives

$$\sup_{G\in\mathcal{G}}\left|\widehat{\mathfrak{R}}_{n,\mathcal{A}}^{\mathrm{CF}}(G) - \frac{1}{K}\sum_{k=1}^{K}\mathcal{R}_{\mathrm{score},\tau}\left(G, \widehat{s}_G^{(-k)}\right)\right| \le \Delta_{n,\mathrm{val}}^{\mathrm{score}} + \Delta_{n,\mathrm{val}}^{\mathrm{corr}} + \varepsilon_{m_n}^{\mathrm{den}}.$$

Applying the triangle inequality and the definition of $\Delta_{m_n}^{\mathrm{alg}}$ proves equation 8. □

Proof of Theorem 3.2. Under the assumptions of Theorem 3.2,

$$\Delta_{n,\mathrm{val}}^{\mathrm{score}} = o_{\mathbb{P}}(1), \qquad \Delta_{n,\mathrm{val}}^{\mathrm{corr}} = o_{\mathbb{P}}(1),$$

and

$$\varepsilon_{m_n}^{\mathrm{den}} = o_{\mathbb{P}}(1), \qquad \Delta_{m_n}^{\mathrm{alg}} = o_{\mathbb{P}}(1).$$

Lemma B.1 therefore implies

$$\sup_{G\in\mathcal{G}}\left|\widehat{\mathfrak{R}}_{n,\mathcal{A}}^{\mathrm{CF}}(G) - \mathfrak{R}_{m_n,\mathcal{A}_s}(G)\right| \xrightarrow{\mathbb{P}} 0.$$

□

C Concentration Around the Sample-Size-Dependent Optimum

This section proves Theorem 3.3.

Define the optimal algorithm-dependent risk by $\mathfrak{R}_{m_n,\mathcal{A}_s}^\star := \inf_{G\in\mathcal{G}}\mathfrak{R}_{m_n,\mathcal{A}_s}(G)$, with minimizer set $\mathcal{G}_{m_n,\mathcal{A}_s}^\star := \left\{G \in \mathcal{G} : \mathfrak{R}_{m_n,\mathcal{A}_s}(G) = \mathfrak{R}_{m_n,\mathcal{A}_s}^\star\right\}$. For a fixed $\varepsilon > 0$, define

$$A_{n,\varepsilon} := \left\{G \in \mathcal{G} : \mathrm{dist}\left(G, \mathcal{G}_{m_n,\mathcal{A}_s}^\star\right) \ge \varepsilon\right\},$$

and the corresponding risk gap

$$\Delta_{n,\varepsilon} := \inf_{G\in A_{n,\varepsilon}}\mathfrak{R}_{m_n,\mathcal{A}_s}(G) - \mathfrak{R}_{m_n,\mathcal{A}_s}^\star.$$

Let

$$\eta_n := \sup_{G\in\mathcal{G}}\left|\widehat{\mathfrak{R}}_{n,\mathcal{A}}^{\mathrm{CF}}(G) - \mathfrak{R}_{m_n,\mathcal{A}_s}(G)\right|.$$

For a positive sequence ξ_n, define the near-optimal set

$$B_{n,\xi_n} := \left\{G \in \mathcal{G} : \mathfrak{R}_{m_n,\mathcal{A}_s}(G) \le \mathfrak{R}_{m_n,\mathcal{A}_s}^\star + \xi_n\right\}.$$

C.1 Posterior bound

Recall the generalized posterior

$$\Pi_n(dG) = \frac{\exp\left\{-\kappa_n\widehat{\mathfrak{R}}_{n,\mathcal{A}}^{\mathrm{CF}}(G)\right\}\Pi_0(dG)}{\int_{\mathcal{G}}\exp\left\{-\kappa_n\widehat{\mathfrak{R}}_{n,\mathcal{A}}^{\mathrm{CF}}(G')\right\}\Pi_0(dG')}.$$

On the event $\eta_n < \frac{\Delta_{n,\varepsilon}-\xi_n}{2}$, every $G \in A_{n,\varepsilon}$ satisfies

$$\widehat{\mathfrak{R}}_{n,\mathcal{A}}^{\mathrm{CF}}(G) \ge \mathfrak{R}_{m_n,\mathcal{A}_s}^\star + \Delta_{n,\varepsilon} - \eta_n.$$

Hence

$$\int_{A_{n,\varepsilon}}\exp\left\{-\kappa_n\widehat{\mathfrak{R}}_{n,\mathcal{A}}^{\mathrm{CF}}(G)\right\}\Pi_0(dG) \le \exp\left\{-\kappa_n\left(\mathfrak{R}_{m_n,\mathcal{A}_s}^\star + \Delta_{n,\varepsilon} - \eta_n\right)\right\}.$$

On the other hand, every $G \in B_{n,\xi_n}$ satisfies $\widehat{\mathfrak{R}}_{n,\mathcal{A}}^{\mathrm{CF}}(G) \le \mathfrak{R}_{m_n,\mathcal{A}_s}^\star + \xi_n + \eta_n$. Therefore the normalizing constant is bounded below by

$$\Pi_0(B_{n,\xi_n})\exp\left\{-\kappa_n\left(\mathfrak{R}_{m_n,\mathcal{A}_s}^\star + \xi_n + \eta_n\right)\right\}.$$

Combining the numerator and denominator bounds gives

$$\Pi_n(A_{n,\varepsilon}) \le \Pi_0(B_{n,\xi_n})^{-1}\exp\left\{-\kappa_n\left[\Delta_{n,\varepsilon} - \xi_n - 2\eta_n\right]\right\}. \tag{9}$$

C.2 Proof of posterior concentration

Proof of Theorem 3.3. By assumption, $\eta_n = o_{\mathbb{P}}(\Delta_{n,\varepsilon})$, $\xi_n = o(\Delta_{n,\varepsilon})$. Hence $\eta_n < (\Delta_{n,\varepsilon} - \xi_n)/2$ with probability tending to one. Taking logarithms in equation 9 gives

$$\log \Pi_n(A_{n,\varepsilon}) \leq -\kappa_n \left[\Delta_{n,\varepsilon} - \xi_n - 2\eta_n\right] - \log \Pi_0(B_{n,\xi_n}).$$

The assumptions $\kappa_n \Delta_{n,\varepsilon} \to \infty$, $-\log \Pi_0(B_{n,\xi_n}) = o(\kappa_n \Delta_{n,\varepsilon})$ imply that the right-hand side converges to $-\infty$ in probability. Therefore

$$\Pi_n\left(G : \operatorname{dist}\left(G, \mathcal{G}^{\star}_{m_n,\mathcal{A}_s}\right) \geq \varepsilon\right) \xrightarrow{\mathbb{P}} 0.$$

□

D Finite-Sample Score Estimation

This section proves Theorem 3.4.

Let $\Lambda \sim G$ and $\mathsf{T} \sim \rho$ independently, and define

$$\mathsf{Z} := (X^{\Lambda}_{\mathsf{T}}, \mathsf{T}, \Lambda), \qquad \varpi_{\mathsf{T},\Lambda} := \frac{\Lambda g(\mathsf{T})^2}{2\rho(\mathsf{T})}.$$

The exact score is

$$S^{\star}(\mathsf{Z}) := s^{\Lambda}_{\mathsf{T}}(X^{\Lambda}_{\mathsf{T}}),$$

and the DSM target admits the regression decomposition

$$Y = S^{\star}(\mathsf{Z}) + \xi, \qquad \mathbb{E}[\xi \mid \mathsf{Z}] = 0.$$

By construction,

$$\mathcal{C}_{\tau}(G) = \mathbb{E}\left[\varpi_{\mathsf{T},\Lambda} \|\xi\|^2\right].$$

Consider the d_ϕ-dimensional score class

$$\mathcal{F}_{d_\phi} := \left\{ s_\Theta(\mathsf{Z}) = \Theta\phi(\mathsf{Z}) : \Theta \in \mathbb{R}^{d \times d_\phi} \right\},$$

and define the weighted population Gram matrix

$$\Sigma_G := \mathbb{E}\left[\varpi_{\mathsf{T},\Lambda} \phi(\mathsf{Z})\phi(\mathsf{Z})^{\top}\right].$$

Throughout this section, assume $\Sigma_G \succ 0$.

Let

$$\Theta^{\star}_G \in \arg\min_{\Theta} \mathbb{E}\left[\varpi_{\mathsf{T},\Lambda} \left\|\Theta\phi(\mathsf{Z}) - S^{\star}(\mathsf{Z})\right\|^2\right],$$

and define the corresponding approximation error by

$$\mathcal{E}_{\mathrm{app},d_\phi}(G) := \mathbb{E}\left[\varpi_{\mathsf{T},\Lambda} \left\|\Theta^{\star}_G \phi(\mathsf{Z}) - S^{\star}(\mathsf{Z})\right\|^2\right].$$

Given m i.i.d. training samples $(\mathsf{Z}_i, Y_i, \varpi_i)_{i=1}^m$, let

$$\widehat{\Theta}_G \in \arg\min_{\Theta} \frac{1}{m} \sum_{i=1}^{m} \varpi_i \left\|\Theta\phi_i - Y_i\right\|^2, \qquad \phi_i := \phi(\mathsf{Z}_i),$$

and define

$$\widehat{\Sigma}_G := \frac{1}{m} \sum_{i=1}^{m} \varpi_i \phi_i \phi_i^{\top}.$$

D.1 Population projection

Define the approximation residual

$$A(\mathsf{Z}) := \Theta_G^\star \phi(\mathsf{Z}) - S^\star(\mathsf{Z}).$$

The population first-order condition gives

$$\mathbb{E}\left[\varpi_{\mathsf{T},\Lambda} A(\mathsf{Z})\phi(\mathsf{Z})^\top\right] = 0.$$

Let

$$R := Y - \Theta_G^\star \phi(\mathsf{Z}) = \xi - A(\mathsf{Z}).$$

Since $\mathbb{E}[\xi \mid \mathsf{Z}] = 0$,

$$\mathbb{E}\left[\varpi_{\mathsf{T},\Lambda}\xi\phi(\mathsf{Z})^\top\right] = 0,$$

and therefore

$$\mathbb{E}\left[\varpi_{\mathsf{T},\Lambda} R\phi(\mathsf{Z})^\top\right] = 0.$$

Moreover,

$$\mathbb{E}\left[\varpi_{\mathsf{T},\Lambda}\|R\|^2\right] = \mathbb{E}\left[\varpi_{\mathsf{T},\Lambda}\|\xi\|^2\right] + \mathbb{E}\left[\varpi_{\mathsf{T},\Lambda}\|A(\mathsf{Z})\|^2\right],$$

because the cross term vanishes by conditional centering. Hence

$$\mathbb{E}\left[\varpi_{\mathsf{T},\Lambda}\|R\|^2\right] = \mathcal{C}_\tau(G) + \mathcal{E}_{\mathrm{app},d_\phi}(G). \tag{10}$$

D.2 Weighted least-squares error

Define

$$\Gamma_m := \frac{1}{m}\sum_{i=1}^m \varpi_i R_i \phi_i^\top \in \mathbb{R}^{d\times d_\phi}.$$

The empirical normal equation is

$$\frac{1}{m}\sum_{i=1}^m \varpi_i (Y_i - \widehat{\Theta}_G\phi_i)\phi_i^\top = 0.$$

Since $Y_i = \Theta_G^\star\phi_i + R_i$, $(\widehat{\Theta}_G - \Theta_G^\star)\widehat{\Sigma}_G = \Gamma_m$. Thus, whenever $\widehat{\Sigma}_G$ is invertible,

$$\widehat{\Theta}_G - \Theta_G^\star = \Gamma_m \widehat{\Sigma}_G^{-1}.$$

Because $\Theta_G^\star$ is the weighted L^2 projection of $S^\star$ onto $\mathcal{F}_{d_\phi}$,

$$\mathcal{R}_{\mathrm{score},\tau}(G, s_{\widehat{\Theta}_G}) = \mathcal{E}_{\mathrm{app},d_\phi}(G) + \mathrm{tr}\left[(\widehat{\Theta}_G - \Theta_G^\star)\Sigma_G(\widehat{\Theta}_G - \Theta_G^\star)^\top\right]. \tag{11}$$

D.3 Concentration of the weighted Gram matrix

Assume the weighted leverage condition

$$\varpi_{\mathsf{T},\Lambda}\phi(\mathsf{Z})^\top\Sigma_G^{-1}\phi(\mathsf{Z}) \le \ell_G d_\phi \qquad \text{almost surely.} \tag{12}$$

Define

$$M_i := \varpi_i \Sigma_G^{-1/2}\phi_i\phi_i^\top\Sigma_G^{-1/2}.$$

Then

$$M_i \succeq 0, \qquad \mathbb{E}[M_i] = I_{d_\phi}, \qquad \lambda_{\max}(M_i) \le \ell_G d_\phi.$$

By the matrix Chernoff inequality, there exists a universal constant $c_0 > 0$ such that

$$\mathbb{P}\left(\widehat{\Sigma}_G \not\succeq \frac{1}{2}\Sigma_G\right) \le d_\phi \exp\left\{-c_0\frac{m}{\ell_G d_\phi}\right\}. \tag{13}$$

Let

$$\mathscr{G}_G := \left\{\widehat{\Sigma}_G \succeq \frac{1}{2}\Sigma_G\right\}.$$

On $\mathscr{G}_G$, define

$$A_G := \Sigma_G^{-1/2}\widehat{\Sigma}_G\Sigma_G^{-1/2}.$$

Then $A_G \succeq \frac{1}{2}I_{d_\phi}$ and hence $A_G^{-2} \preceq 4I_{d_\phi}$. By congruence,

$$\widehat{\Sigma}_G^{-1}\Sigma_G\widehat{\Sigma}_G^{-1} \preceq 4\Sigma_G^{-1}. \tag{14}$$

D.4 Control of the stochastic term

The summands in Γ_m have mean zero. Independence therefore gives

$$\mathbb{E}\,\mathrm{tr}\left[\Gamma_m \Sigma_G^{-1} \Gamma_m^\top\right] = \frac{1}{m}\mathbb{E}\left[\varpi_{\mathsf{T},\Lambda}^2 \|R\|^2 \phi(\mathsf{Z})^\top \Sigma_G^{-1} \phi(\mathsf{Z})\right].$$

Using equation 12 and equation 10,

$$\mathbb{E}\,\mathrm{tr}\left[\Gamma_m \Sigma_G^{-1} \Gamma_m^\top\right] \le \frac{\ell_G d_\phi}{m}\left[\mathcal{E}_{\mathrm{app},d_\phi}(G) + \mathcal{C}_\tau(G)\right]. \tag{15}$$

Therefore, by Markov's inequality, for every $\gamma \in (0,1)$, with probability at least $1-\gamma$,

$$\mathrm{tr}\left[\Gamma_m \Sigma_G^{-1} \Gamma_m^\top\right] \le \frac{\ell_G d_\phi}{m\gamma}\left[\mathcal{E}_{\mathrm{app},d_\phi}(G) + \mathcal{C}_\tau(G)\right]. \tag{16}$$

D.5 Proof of the finite-sample score-risk bound

Proof of Theorem 3.4. On $\mathscr{G}_G$, equation 14 implies

$$\mathrm{tr}\left[(\widehat{\Theta}_G - \Theta_G^\star)\Sigma_G(\widehat{\Theta}_G - \Theta_G^\star)^\top\right] = \mathrm{tr}\left[\Gamma_m \widehat{\Sigma}_G^{-1}\Sigma_G\widehat{\Sigma}_G^{-1}\Gamma_m^\top\right] \le 4\,\mathrm{tr}\left[\Gamma_m\Sigma_G^{-1}\Gamma_m^\top\right].$$

Intersecting $\mathscr{G}_G$ with the event in equation 16 gives

$$\mathrm{tr}\left[(\widehat{\Theta}_G - \Theta_G^\star)\Sigma_G(\widehat{\Theta}_G - \Theta_G^\star)^\top\right] \le \frac{4\ell_G d_\phi}{m\gamma}\left[\mathcal{E}_{\mathrm{app},d_\phi}(G) + \mathcal{C}_\tau(G)\right].$$

Substituting this bound into equation 11 yields

$$\mathcal{R}_{\mathrm{score},\tau}(G, s_{\widehat{\Theta}_G}) \le \ \mathcal{E}_{\mathrm{app},d_\phi}(G) + \frac{4\ell_G d_\phi}{m\gamma}\left[\mathcal{E}_{\mathrm{app},d_\phi}(G) + \mathcal{C}_\tau(G)\right]. \tag{17}$$

By equation 13,

$$\mathbb{P}(\mathscr{G}_G^c) \le d_\phi \exp\left\{-c_0 \frac{m}{\ell_G d_\phi}\right\}.$$

A union bound therefore shows that equation 17 holds with probability at least $1 - \gamma - d_\phi \exp\left\{-c_0 \frac{m}{\ell_G d_\phi}\right\}$, as stated in Theorem 3.4. □

If the exact score belongs to $\mathcal{F}_{d_\phi}$, then $\mathcal{E}_{\mathrm{app},d_\phi}(G) = 0$. Hence

$$\mathcal{R}_{\mathrm{score},\tau}(G, s_{\widehat{\Theta}_G}) \le \frac{4\ell_G d_\phi}{m\gamma}\mathcal{C}_\tau(G).$$

E Reverse-Process Guarantees

This section proves Theorem 3.5, Lemma E.2, and Corollary E.3.

E.1 Reverse-time diffusion and regularity conditions

Fix $\tau \in (0,T)$ and a latent scale λ. Recall the conditional forward diffusion

$$dX_t^\lambda = f(X_t^\lambda, t)\,dt + \sqrt{\lambda}\,g(t)\,dW_t.$$

Let $u := T - t$ and define $\bar{X}_u^\lambda := X_{T-u}^\lambda$ for $u \in [0, T-\tau]$.

Since the diffusion coefficient is state independent, the exact conditional reverse process satisfies

$$d\bar{X}_u^\lambda = b_\lambda^{\mathrm{rev}}(\bar{X}_u^\lambda, u)\,du + \sqrt{\lambda}\,g(T-u)\,d\overline{W}_u,$$

where

$$b_\lambda^{\mathrm{rev}}(x,u) := -f(x, T-u) + \lambda g(T-u)^2 s_{T-u}^\lambda(x).$$

Replacing the exact score by $\widehat{s}_t^{\lambda}$ defines the learned reverse process

$$d\widehat{\bar{X}}_u^{\lambda} = \widehat{b}_{\lambda}^{\mathrm{rev}}(\widehat{\bar{X}}_u^{\lambda}, u)\, du + \sqrt{\lambda}\, g(T-u)\, d\overline{W}_u,$$

with

$$\widehat{b}_{\lambda}^{\mathrm{rev}}(x,u) := -f(x, T-u) + \lambda g(T-u)^2 \widehat{s}_{T-u}^{\lambda}(x).$$

Let $\mathbf{P}_{\lambda}^{\mathrm{rev}}$ and $\mathbf{Q}_{\lambda,\widehat{s}}^{\mathrm{rev}}$ denote the corresponding path measures on $C([0, T-\tau], \mathbb{R}^d)$. The exact reverse process is initialized from p_T^{λ}, while the learned reverse process is initialized from $\widetilde{p}_T^{\lambda}$.

We use the following regularity conditions.

Assumption E.1 (Girsanov regularity)**.** *For G-almost every λ:*

1. *the exact and learned reverse SDEs admit weak solutions;*
2. *$p_T^{\lambda} \ll \widetilde{p}_T^{\lambda}$;*
3. *the standardized drift difference $h_u^{\lambda}(x) := \sqrt{\lambda}\, g(T-u)\left[s_{T-u}^{\lambda}(x) - \widehat{s}_{T-u}^{\lambda}(x)\right]$ satisfies a sufficient Girsanov integrability condition. In particular, it is sufficient that*

$$\mathbb{E}_{\mathbf{P}_{\lambda}^{\mathrm{rev}}}\left[\exp\left\{\frac{1}{2}\int_0^{T-\tau} \|h_u^{\lambda}(\bar{X}_u^{\lambda})\|^2\, du\right\}\right] < \infty.$$

Define the joint reverse path measures by

$$\mathbf{P}_G^{\mathrm{rev}}(d\lambda, d\omega) := G(d\lambda)\, \mathbf{P}_{\lambda}^{\mathrm{rev}}(d\omega),$$

and

$$\mathbf{Q}_{G,\widehat{s}}^{\mathrm{rev}}(d\lambda, d\omega) := G(d\lambda)\, \mathbf{Q}_{\lambda,\widehat{s}}^{\mathrm{rev}}(d\omega).$$

E.2 Reverse-path relative entropy

Recall the terminal mismatch

$$\mathcal{D}_T(G) := \int \mathrm{KL}\left(p_T^{\lambda} \,\|\, \widetilde{p}_T^{\lambda}\right) G(d\lambda),$$

and the weighted conditional score risk

$$\mathcal{R}_{\mathrm{score},\tau}(G, \widehat{s}) := \frac{1}{2}\int G(d\lambda) \int_{\tau}^{T} \lambda g(t)^2 \mathbb{E}_{X_t^{\lambda}\sim p_t^{\lambda}}\left[\left\|\widehat{s}_t^{\lambda}(X_t^{\lambda}) - s_t^{\lambda}(X_t^{\lambda})\right\|^2\right] dt.$$

Proof of Theorem 3.5. Fix λ. The exact and learned reverse processes share the diffusion coefficient

$$\sigma_{\lambda}(u) = \sqrt{\lambda}\, g(T-u) I_d.$$

Their drift difference satisfies

$$b_{\lambda}^{\mathrm{rev}}(x,u) - \widehat{b}_{\lambda}^{\mathrm{rev}}(x,u) = \lambda g(T-u)^2\left[s_{T-u}^{\lambda}(x) - \widehat{s}_{T-u}^{\lambda}(x)\right].$$

Hence

$$\sigma_{\lambda}(u)^{-1}\left(b_{\lambda}^{\mathrm{rev}} - \widehat{b}_{\lambda}^{\mathrm{rev}}\right) = h_u^{\lambda}.$$

By Girsanov's theorem and the chain rule for relative entropy when the initial distributions differ,

$$\mathrm{KL}\left(\mathbf{P}_{\lambda}^{\mathrm{rev}} \,\|\, \mathbf{Q}_{\lambda,\widehat{s}}^{\mathrm{rev}}\right) = \mathrm{KL}\left(p_T^{\lambda} \,\|\, \widetilde{p}_T^{\lambda}\right) + \frac{1}{2}\mathbb{E}_{\mathbf{P}_{\lambda}^{\mathrm{rev}}}\left[\int_0^{T-\tau} \|h_u^{\lambda}(\bar{X}_u^{\lambda})\|^2\, du\right].$$

Under the exact reverse law, $\bar{X}_u^{\lambda}$ has marginal density p_{T-u}^{λ}. Changing variables $t = T-u$ therefore gives

$$\frac{1}{2}\mathbb{E}_{\mathbf{P}_{\lambda}^{\mathrm{rev}}}\left[\int_0^{T-\tau} \|h_u^{\lambda}(\bar{X}_u^{\lambda})\|^2\, du\right] = \frac{1}{2}\int_{\tau}^{T} \lambda g(t)^2 \mathbb{E}_{X_t^{\lambda}\sim p_t^{\lambda}}\left[\left\|\widehat{s}_t^{\lambda}(X_t^{\lambda}) - s_t^{\lambda}(X_t^{\lambda})\right\|^2\right] dt.$$

Since both joint path measures have the same marginal distribution G for λ, the chain rule for relative entropy gives

$$\mathrm{KL}\left(\mathbf{P}_G^{\mathrm{rev}} \,\|\, \mathbf{Q}_{G,\widehat{s}}^{\mathrm{rev}}\right) = \int \mathrm{KL}\left(\mathbf{P}_\lambda^{\mathrm{rev}} \,\|\, \mathbf{Q}_{\lambda,\widehat{s}}^{\mathrm{rev}}\right) G(d\lambda).$$

Combining the preceding identities proves

$$\mathrm{KL}\left(\mathbf{P}_G^{\mathrm{rev}} \,\|\, \mathbf{Q}_{G,\widehat{s}}^{\mathrm{rev}}\right) = \mathcal{D}_T(G) + \mathcal{R}_{\mathrm{score},\tau}(G,\widehat{s}). \tag{18}$$

□

If $\widetilde{p}_T^\lambda = p_T^\lambda$ for G-almost every λ, then $\mathcal{D}_T(G) = 0$, and hence

$$\mathrm{KL}\left(\mathbf{P}_G^{\mathrm{rev}} \,\|\, \mathbf{Q}_{G,\widehat{s}}^{\mathrm{rev}}\right) = \mathcal{R}_{\mathrm{score},\tau}(G,\widehat{s}).$$

E.3 Terminal mismatch for the VP diffusion

For the VP diffusion,

$$X_T^\lambda = a_T X_0 + \sqrt{\lambda v_T}\, Z, \qquad Z \sim \mathcal{N}(0, I_d),$$

and hence

$$p_T^\lambda = \int \mathcal{N}\left(a_T x_0, \lambda v_T I_d\right) P_0(dx_0).$$

We initialize the learned reverse process from $\widetilde{p}_T^\lambda := \mathcal{N}\left(0, \lambda v_T I_d\right)$.

Lemma E.2 (Terminal mismatch)**.** *If every admissible scale satisfies $\lambda \geq \underline{\lambda} > 0$, then*

$$\sup_{G \in \mathcal{G}} \mathcal{D}_T(G) \leq \frac{a_T^2 \mathbb{E}\|X_0\|^2}{2\underline{\lambda} v_T}.$$

Proof. By convexity of relative entropy in its first argument,

$$\begin{aligned}\mathrm{KL}\left(p_T^\lambda \,\|\, \widetilde{p}_T^\lambda\right) &\leq \mathbb{E}_{X_0}\left[\mathrm{KL}\left(\mathcal{N}(a_T X_0, \lambda v_T I_d) \,\|\, \mathcal{N}(0, \lambda v_T I_d)\right)\right] \\ &= \frac{a_T^2 \mathbb{E}\|X_0\|^2}{2\lambda v_T}.\end{aligned}$$

Therefore, if $\lambda \geq \underline{\lambda}$,

$$\mathcal{D}_T(G) \leq \frac{a_T^2 \mathbb{E}\|X_0\|^2}{2\underline{\lambda} v_T}.$$

Taking the supremum over $G \in \mathcal{G}$ proves the result. □

Thus, whenever the terminal signal coefficient a_T is small, the initialization error is uniformly small over the admissible latent-scale distributions.

E.4 Endpoint consistency

Let P_τ^G and $Q_{\tau,G,\widehat{s}}$ denote the endpoint distributions of the exact and learned reverse processes at forward time τ.

Corollary E.3 (Endpoint consistency)**.** *Under the assumptions of Theorem 3.5,*

$$\left\|P_\tau^G - Q_{\tau,G,\widehat{s}}\right\|_{\mathrm{TV}} \leq \sqrt{\frac{1}{2}\left[\mathcal{D}_T(G) + \mathcal{R}_{\mathrm{score},\tau}(G,\widehat{s})\right]}.$$

Proof. The endpoint at time τ is a measurable function of the joint reverse trajectory $(\lambda, \bar{X}_{0:T-\tau}^\lambda)$. The data-processing inequality gives

$$\mathrm{KL}\left(P_\tau^G \,\|\, Q_{\tau,G,\widehat{s}}\right) \leq \mathrm{KL}\left(\mathbf{P}_G^{\mathrm{rev}} \,\|\, \mathbf{Q}_{G,\widehat{s}}^{\mathrm{rev}}\right).$$

Applying Theorem 3.5 and Pinsker's inequality yields the result. □

Consequently, for any sequence $(G_n, \widehat{s}_n)$ satisfying

$$\mathcal{D}_T(G_n) \to 0, \qquad \mathcal{R}_{\mathrm{score},\tau}(G_n, \widehat{s}_n) \to 0,$$

we have

$$\left\| P_\tau^{G_n} - Q_{\tau,G_n,\widehat{s}_n} \right\|_{\mathrm{TV}} \to 0.$$

If the regularity conditions extend to the full interval $[0, T]$ so that $\tau = 0$ is admissible, then $P_0^{G_n} = P_0$ for every n, because the forward process starts from $X_0 \sim P_0$ independently of the latent scale. Therefore,

$$\|P_0 - Q_{0,G_n,\widehat{s}_n}\|_{\mathrm{TV}} \to 0.$$

F Synthetic Experiment Details

We use the one-dimensional binary model described in Section 4.1, with $X_0 = \mu B$, $\mathbb{P}(B = 1) = \mathbb{P}(B = -1) = 1/2$. The VP process is evaluated at the fixed time $t_0 = 0.1$ with $\beta = 10$, giving $a_{t_0} = 0.6065$ and $v_{t_0} = 0.6321$. For a latent scale λ, $X_{t_0}^\lambda = a_{t_0} X_0 + \sqrt{\lambda v_{t_0}} Z$, with $Z \sim \mathcal{N}(0, 1)$. The score learner is the one-parameter fixed-feature model

$$s_\theta^\lambda(x) = -\frac{x}{\lambda v_{t_0}} + \theta \frac{a_{t_0}\mu}{\lambda v_{t_0}} \operatorname{sign}(x),$$

where the same coefficient θ is shared across all scales. We use $m = 2048$ training samples throughout the main synthetic study.

The shared coefficient is fitted by closed-form weighted least squares, so no iterative optimizer is required for the score learner. The corresponding finite-resource score risk $\mathfrak{R}_{m,\mathcal{A}_s}(G)$ is evaluated using the analytic score and deterministic numerical integration. We vary the score-learning geometry through $c = a_{t_0}^2 \mu^2 / v_{t_0}$. The non-Gaussian-optimal regime uses $c = c_\star = 2.0839$, whereas the Gaussian-optimal control uses $c = c_\star/2 = 1.0420$. The finite-resource optima reported in Section 4.1 are computed over the same normalized perturbation family used by the BNP model.

To evaluate cross-law ranking, we use

$$G_h = \frac{1}{2}\delta_{1-h} + \frac{1}{2}\delta_{1+h}, \qquad h \in \{0, .05, .10, .20, .30, .40, .50, .60, .70\}.$$

Here h determines the separation of the two scales around their fixed mean of one. Raw DSM and corrected risk are evaluated on held-out samples independent of score fitting. Because the Bayes denoiser is available analytically in this model, the correction is evaluated without denoiser approximation error. Common random numbers are used across candidate perturbation laws within each replication.

We use a $K = 4$ truncated stick-breaking representation with concentration parameter $\alpha = 1$. Raw atoms lie in $[0.5, 2]$ and are normalized by their mixture mean so that $\mathbb{E}_G[\Lambda] = 1$. The same prior and optimization protocol are used in both synthetic regimes. The variational generalized-Bayes objective uses $\kappa = 2048$ and the independently validated differentiable finite-resource risk surrogate described in Section 4.1. We optimize the variational parameters with Adam for 5000 updates using a learning rate of 3×10^{-3} and 32 variational samples per update. The reported BNP results aggregate three independent seeds $\{20260905, 20260906, 20260907\}$.

The primary generalized-Bayes temperature is $\eta = 1$. We additionally evaluate the preregistered values $\eta \in \{0.25, 1, 4\}$. The qualitative adaptation pattern is unchanged across these settings, and all main-text results use the prespecified value $\eta = 1$.

G Real-World Experiment Details

We evaluate on CIFAR-10 and on fields from the High-Resolution Rapid Refresh (HRRR) weather model. For CIFAR-10, the official training set is split into 45k training and 5k validation examples, while the official 10k test set is kept untouched during perturbation-law selection. For HRRR, we use hourly forecast-zero fields from 1 January 2021 through 31 December 2023. We consider vertically integrated liquid water (VIL) and pressure vertical velocity at level 20 (w20; VVEL, Pa/s).

The fields are downsampled to single-channel 64×64 arrays and split chronologically into 18,396 training, 3,942 validation, and 3,942 test time points. VIL is transformed by $\log(1 + x)$ before standardization, while w20 is kept signed.

For each task, we compare the unit-scale Gaussian law $G = \delta_1$, Student-t perturbations with $\nu \in \{3, 5, 10\}$, and the adaptive BNP model. The Student-t perturbations use the Gaussian scale-mixture representation (Andrews & Mallows, 1974),

$$\Lambda \sim \mathrm{InvGamma}\left(\frac{\nu}{2}, \frac{\nu - 2}{2}\right),$$

which gives $\mathbb{E}[\Lambda] = 1$. The BNP model uses the normalized latent-scale construction of Section 2.2 with $K = 4$, concentration parameter $\alpha = 1$, and raw atoms supported on $[0.5, 2]$. All perturbation laws therefore use the same mean-scale normalization.

All experiments use the VP forward process with $\beta = 10$ and $T = 1$. For CIFAR-10, we use an x_0-prediction U-Net (Ronneberger et al., 2015) with 64 base channels. For HRRR, we use a single-channel x_0-prediction U-Net with 48 base channels. Both networks condition on diffusion time and $\log \lambda$ through FiLM (Perez et al., 2018). Models are trained with Adam at learning rate 2×10^{-4}. The two resource regimes correspond to 25k and 100k score-training updates. Within each task and resource regime, the network architecture, optimizer, and score-training budget are held fixed across perturbation laws.

During BNP adaptation, the score model and the variational perturbation law are updated alternately using the corrected validation criterion. After adaptation, the learned perturbation law is frozen, the adaptation-stage score model is discarded, and a fresh score model is trained from scratch under the frozen law using the same 25k or 100k score-training budget as the fixed-law baselines.

Perturbation laws are compared using the corrected DSM criterion introduced in Section 2.3; raw DSM is not used for model selection. For CIFAR-10, perturbation-law selection is performed using the held-out validation set, and the final selected models are evaluated on the untouched official test set. For HRRR, we report per-dimension corrected validation risk, obtained by normalizing the image-level risk by the 64×64 spatial dimensions. No generation metric or downstream metric is used for perturbation-law selection.

Score models are trained with $t \sim \mathrm{Unif}(0, 1)$ using the unweighted x_0-prediction loss. Corrected-risk evaluation uses $q_{\mathrm{eval}}(t) \propto 1/t$ on $[\tau, T]$, with $\tau = \log(1 + 1/10)/10$, together with the corresponding importance weight. The correction model is estimated using two independently initialized nuisance denoisers trained for 100k updates on the training split under the fixed reference proposal used in the supplementary code. The CIFAR-10 and HRRR batch sizes are 128 and 96, respectively.

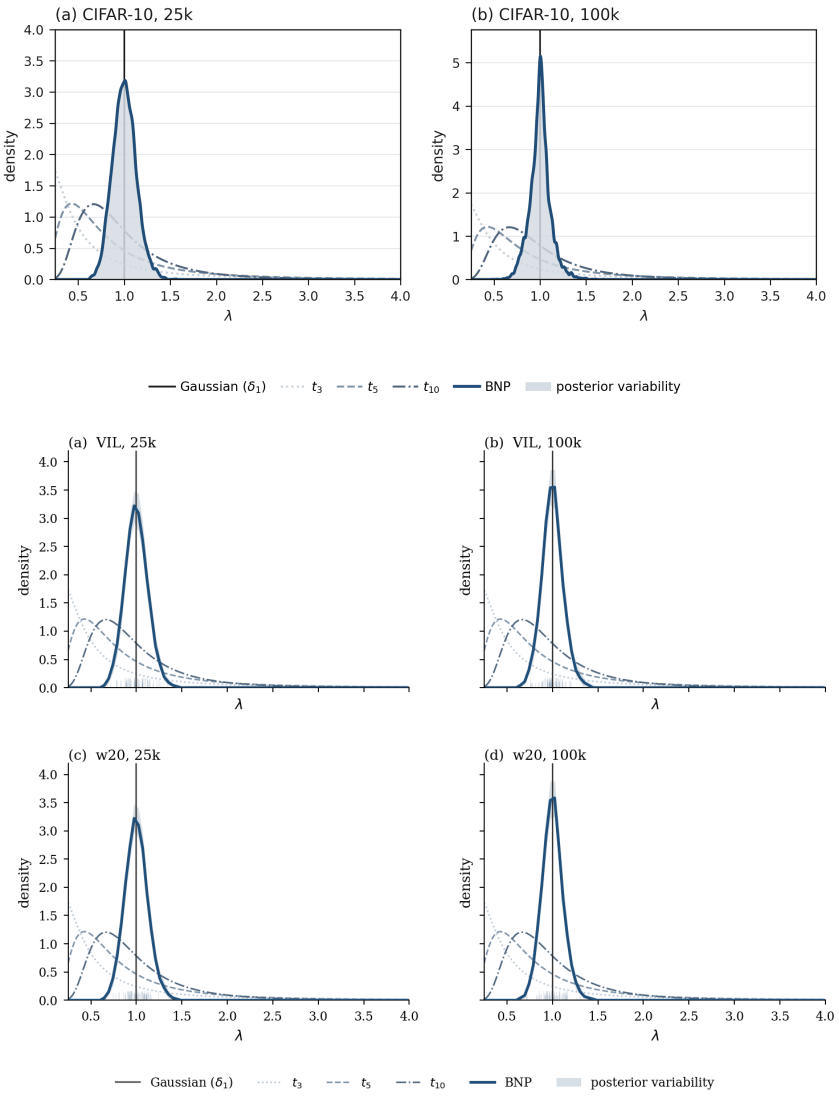

Figure 2: Learned posterior-predictive perturbation-scale distributions under 25k and 100k training budgets. Top: CIFAR-10. Bottom: HRRR VIL and w20.