DIFFUSION-BASED POSTERIOR SAMPLING: A FEYNMAN–KAC ANALYSIS OF BIAS AND STABILITY

Anonymous authors
Paper under double-blind review

ABSTRACT

Diffusion-based posterior samplers draw from measurement- or reward-conditioned posteriors using pretrained diffusion priors. Though widely used for inverse problems, their output is biased even with exact prior scores. First, we quantify this bias for the DPS algorithm, the base method for a family of samplers, by introducing a surrogate curve of probability densities which interpolates between the true posterior and the standard Gaussian. The ratio of sampler to surrogate densities solves a parabolic PDE, and its Feynman–Kac formula yields an explicit representation of this ratio as the expectation of a path functional. This representation reveals which posterior regions are over- or under-sampled. This framework also provides a design principle for bias reduction that includes STSL, a related sampler.

Second, in hard-constraint problems (e.g. the low measurement noise regime) these samplers are known to be numerically unstable. We find that this instability of discretized DPS arises from an effective guidance update whose strength scales inversely with step size, inevitably violating the forward-Euler stability criterion near the constraint. This finding motivates implicit (backward-Euler) guidance updates. Finally, we show that the standard practice of early guidance-stopping, used to avoid oscillations, introduces a secondary bias which we characterize through our framework.

1 INTRODUCTION

Diffusion and score-based generative models (Sohl-Dickstein et al., 2015; Ho et al., 2020; Song & Ermon, 2019; Song et al., 2021) have become the workhorse of modern generative modeling, powering text-to-image systems (Rombach et al., 2022; Ramesh et al., 2021; Saharia et al., 2022; Dhariwal & Nichol, 2021) and an expanding range of scientific and medical inverse problems (Song et al., 2022). Their flexibility hinges on a single learned object (the score $\nabla \log \rho_t$ of the noised data distribution) which can be repurposed across downstream tasks without retraining.

A canonical such task is sampling from a posterior of the form $\mu_y(x) \propto e^{R_y(x)}\rho_*(x)$, encompassing both classical inverse problems $y = \mathcal{A}(x) + \varepsilon$ and reward-tilted generation (Daras et al., 2024). Even granting access to a perfect score oracle, posterior sampling is computationally intractable in the worst case (Gupta et al., 2024), so practical algorithms rely on heuristic guidance that approximates the time-dependent posterior score. An early and influential such heuristic is *Diffusion Posterior Sampling* (DPS) (Chung et al., 2023): it replaces the intractable conditional score $\nabla_{x_t} \log p(y \mid x_t)$ by the gradient of the reward evaluated at the Tweedie posterior mean $\hat{x}_t(x_t) = \mathbb{E}[X_0 \mid X_t = x_t]$ (Robbins, 1956), yielding a plug-and-play guidance compatible with any pretrained score network. Its simplicity has made DPS the de facto baseline for inverse problems and inspired a line of research targeting its known weaknesses: manifold-constrained gradients (Chung et al., 2022), denoising restoration (Kawar et al., 2022), pseudoinverse-guided diffusion (Song et al., 2023), latent-space extensions (Rout et al., 2023), second-order Tweedie corrections (Rout et al., 2024; Boys et al., 2024), proximal approaches to decrease the gradient computation burden (Rout et al., 2025), filtering and SMC-based reweightings (Dou & Song, 2024; Wu et al., 2023; Moufad et al., 2025), and recent drift-control schemes (Ren et al., 2026; Guo et al., 2026; Anil et al., 2026).

Yet despite this flurry of activity, two basic questions remain open. First, the DPS approximation is biased even for Gaussian-mixture priors with quadratic rewards; but *which* samples does this bias over- or under-represent, and *why* do correctives like STSL (Rout et al., 2024) improve performance?

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.

Existing analyses establish convergence under restrictive assumptions on the prior or measurement operator (Xu & Chi, 2024; Parulekar et al., 2026; Moitra et al., 2026) or treat the algorithm as a black box, leaving its preferred classes unexplained. Second, in the low-temperature regime needed for hard measurement constraints in image inverse problems, standard DPS is numerically unstable. Practitioners routinely fall back on early guidance-stopping and trajectory-dependent step sizes, but the effect of these heuristics on the sampled distribution has never been quantified.

Contributions. We close both gaps with a unified analysis based on the classical Feynman–Kac formula (Karatzas & Shreve, 1991), recently used to correct guided and annealed diffusion sampling via SMC reweighting (Skreta et al., 2025; Akhound-Sadegh et al., 2025), and complementing stochastic-analytic perspectives on guidance (Bruna & Han, 2024; Ren et al., 2026; Guo et al., 2026).

(i) An exact bias formula for DPS. In Section 3, we compare DPS with a surrogate path of densities between the standard Gaussian and the true posterior; their ratio solves a parabolic PDE, whose Feynman–Kac formula expresses the DPS bias exactly as an expectation over Ornstein–Uhlenbeck paths (Theorem 1). The reaction term in this expectation compares conditional covariance with reward curvature, revealing where DPS over- or under-samples, and gives a design principle for bias reduction that includes STSL (Rout et al., 2024) (Section 4).

(ii) The source of DPS instability. In Section 5, we show that in hard-constraint (low-temperature) problems, the instability of discretized DPS arises from an effective guidance update whose strength scales inversely with the step size: the guidance schedule cancels the Euler step, so the forward-Euler stability criterion is inevitably violated near the constraint, and no choice of step size repairs it. This motivates implicit (backward-Euler) guidance updates, which we illustrate in Section I.

(iii) The bias of early guidance-stopping. Practitioners mitigate these oscillations with early guidance-stopping, see Algorithm 2. Using the framework of (i), we show that this introduces a secondary bias, and characterize the resulting output as an explicit reweighting of the prior (Theorem 2).

2 Background and Related Work

Score Based Generative Models. We consider sampling from a density ρ_*. A subscript t denotes a noised distribution: $\rho_t := \big(e^{td}\rho_*(e^t\,\cdot)\big) * \mathcal{N}(0, \sigma_t^2 I)$, with $\sigma_t^2 := 1 - e^{-2t}$, is the marginal of the Ornstein–Uhlenbeck (OU) noising process

$$dX_t = -X_t\,dt + \sqrt{2}\,dB_t, \qquad X_0 \sim \rho_*, \tag{1}$$

which interpolates between ρ_* at $t = 0$ and the standard Gaussian $\gamma = \mathcal{N}(0, I)$ as $t \to \infty$. Score based generative models sample by reversing (1) on a horizon $[0, T]$: the reversed process $\overleftarrow{X}_t := X_{T-t}$, with marginals $\overleftarrow{\rho}_t = \rho_{T-t}$, solves the *denoising* SDE

$$d\overleftarrow{X}_t = \big(\overleftarrow{X}_t + 2\nabla \log \rho_{T-t}(\overleftarrow{X}_t)\big)\,dt + \sqrt{2}\,d\overleftarrow{B}_t, \qquad \overleftarrow{X}_0 \sim \rho_T, \tag{2}$$

where $\overleftarrow{B}$ is a Brownian motion in the filtration of $\overleftarrow{X}$ on $[0, T)$ (Haussmann & Pardoux, 1986; Anderson, 1982). In practice, ρ_T is replaced by γ and the score by a network $s_{\theta_*}(x, t) \approx \nabla \log \rho_t(x)$, trained from samples $X \sim \rho_*$ by denoising score matching (Vincent, 2011) in the continuous-time form of Song et al. (2021), of which (1) is the variance-preserving case: $\theta_* = \operatorname{argmin}_\theta \mathbb{E}_{t,X,\xi}\left[\sigma_t^2 \left\|s_\theta(e^{-t}X + \sigma_t\xi,\, t) + \xi/\sigma_t\right\|^2\right]$, with $\xi \sim \mathcal{N}(0, I)$ independent of X and $t \sim \mathrm{Unif}(0, T]$. This is minimized by $s(\cdot, t) = \nabla \log \rho_t$, and coincides with the DDPM objective $\mathbb{E}\,\|\epsilon_\theta - \xi\|^2$ (Ho et al., 2020) under $\epsilon_\theta := -\sigma_t s_\theta$. By Tweedie's formula, the network also gives the denoiser $\mathbb{E}[X_0 \mid X_t = x] = e^t\big(x + \sigma_t^2 \nabla \log \rho_t(x)\big) \approx e^t\big(x + \sigma_t^2\, s_{\theta_*}(x, t)\big)$. Sampling via this reverse process runs in polynomial time (Rombach et al., 2022; Song & Ermon, 2019; Ramesh et al., 2021; Saharia et al., 2022), which compares favorably to classical approaches such as Langevin dynamics (Vempala & Wibisono, 2019), whose convergence rate is instance-dependent and can be arbitrarily slow, see Section C.1.

Posterior Sampling. A natural application of score-based models is to inverse problems and posterior sampling. The score network characterizes a prior ρ_*, and at test time one tilts the samples by a log-likelihood $R_y(x)$ to target the posterior $\mu_y := e^{R_y}\rho_*/Z$; the subscript records the dependence on the measurement y, e.g. $R_y(x) = -\frac{\alpha}{2}\|\mathcal{A}(x) - y\|^2$ for $y = \mathcal{A}(x) + \varepsilon$ with $\varepsilon \sim \mathcal{N}(0, \alpha^{-1} I)$. The main challenge is that the posterior score $\nabla \log(\mu_y)_t$ (the noised posterior is denoted $\overrightarrow{\mu}_t^*$ in Section 3)

cannot be easily computed from $\nabla \log \rho_t$.[1] A range of approximate algorithms have been proposed to circumvent this. A central theme is the use of the prior scores $\nabla \log \rho_t$ through Tweedie's formula to obtain realistic-looking samples even when posterior sampling. Specifically, $\mathbb{E}[X_0 \mid X_t = x_t]$ is used[2] as a computationally tractable proxy for the initial condition X_0 (i.e., the mean endpoint of the reverse diffusion (2) run to completion from x_t) and is fed into the reward model $R_y(\cdot)$ when modifying the drift at test time, see Section C.2. Although the resulting samples are not formally drawn from μ_y, this heuristic performs well in practice.

Feynman–Kac formulas. The sampling literature often focuses on error bounds for approximate sampling algorithms (Lee et al., 2022; Chen et al., 2023; Vempala & Wibisono, 2019), see Section C.1. These are instantiated as upper bounds on the KL/TV/χ^2 distance between the distribution of the sampling algorithm and the ground truth, and under favorable circumstances can be shown to be polynomially or exponentially small in the parameters of the instance. In posterior sampling, such an error is *known* to be large. As such, a KL bound is often vacuous, unless it is accompanied by strong assumptions about the instance. Rather than focusing on bounding this error, we apply machinery that allows us to explicitly track the Radon–Nikodym derivative of approximate posterior sampling algorithm with respect to the true posterior.

In particular, we will exploit the Feynman–Kac representation. Consider two time-dependent densities evolving under possibly different transport and reaction fields:

$$\begin{cases} \partial_t \pi_t = -\nabla \cdot (v_t \pi_t) + \Delta \pi_t + f_t \pi_t \\ \partial_t \pi'_t = \underbrace{-\nabla \cdot (v'_t \pi'_t)}_{\text{transport}} + \underbrace{\Delta \pi'_t}_{\text{diffusion}} + \underbrace{f'_t \pi'_t}_{\text{reaction}} \end{cases} \tag{3}$$

Without the reaction term, (3) are Fokker–Planck equations, whose solutions are the laws of SDEs with drift v_t (resp. v'_t). The reaction term reweights paths by $\exp\big(\int_0^t f_s(X_s)\, ds\big)$, i.e. killing where $f_t < 0$ and branching where $f_t > 0$, so π_t is an unnormalized measure whose normalization is the law of the weighted process.

Assuming $\pi_t > 0$, the *ratio* of the marginals $g_t := \pi'_t / \pi_t$ satisfies

$$\partial_t g_t \;=\; \Delta g_t \;+\; b_t \cdot \nabla g_t \;-\; c_t\, g_t, \tag{4}$$

for an explicit choice of b_t, c_t given in Lemma 2. Fix $t > 0$ and let $(A_s)_{s \in [0,t]}$ be the diffusion process associated to the stochastic characteristics

$$dA_s = b_{t-s}(A_s)\, ds + \sqrt{2}\, dB_s, \tag{5}$$

which run backward through the coefficients of (4). The Feynman–Kac representation of (4) then reads

$$g_t(x) = \mathbb{E}\left[g_0(A_t) \exp\left(-\int_0^t c_{t-s}(A_s)\, ds \right) \,\middle|\, A_0 = x \right]. \tag{6}$$

See Section B for the precise statement and proof, and see Skreta et al. (2025), Akhound-Sadegh et al. (2025), and Ren et al. (2026) for the use of Feynman–Kac formulae in posterior sampling and annealing with diffusion models.

3 The Bias of DPS via Surrogate Paths

This section introduces a surrogate path framework for analyzing diffusion-based posterior samplers. Given a reward $R_y : \mathbb{R}^d \to \mathbb{R}$ our goal is to sample from the posterior $\mu_y(x) = \exp(R_y(x))\rho_*(x)/Z$, an exponential tilt of the prior where $Z > 0$ is the normalization. To this end, we define a *surrogate path* to be a continuous curve of probability measures (marginal densities) $\overrightarrow{\mu} : [0, \infty) \to \mathcal{P}(\mathbb{R}^d)$ of the form $t \mapsto \overrightarrow{\mu}_t := h_t \rho_t / Z_t$ where the *positive* function $h : [0, \infty) \times \mathbb{R}^d \to (0, \infty)$ is a weight satisfying $h \in C^{1,2}([0, \infty) \times \mathbb{R}^d)$, $h_0(x) = \exp(R_y(x))$ (so that $Z_0 = Z$) and h_∞ identically constant. A valid surrogate thus interpolates between μ_y at time zero and the standard Gaussian γ as

[1] The noising process and the tilt do not commute: $(e^{R_y}\rho_*)_t \neq (e^{R_y})_t \rho_t$.

[2] By Tweedie's formula, $\mathbb{E}[X_0 \mid X_t = x_t] = e^t x_t + (e^t - e^{-t}) \nabla \log \rho_t(x_t)$, which can be efficiently estimated at test-time.

$t \to \infty$. For DPS, the sampler density is the marginal law of an associated SDE, which we recover by dropping the reaction term from the exact evolution of the surrogate. This enables us to express the ratio of surrogate to sampler marginals via Feynman–Kac formula. First, we record the evolution of a general surrogate.

Lemma 1 (Informal, see Lemma 3). *For any finite horizon $T < \infty$ and surrogate path $\overrightarrow{\mu}_t = h_t \rho_t / Z_t$, its time-reversal $\overleftarrow{\mu}_t = \overrightarrow{\mu}_{T-t}$ satisfies*

$$\begin{cases} \partial_t \overleftarrow{\mu}_t = -\nabla \cdot (x \overleftarrow{\mu}_t) - 2\nabla \cdot (\nabla \log \overrightarrow{\mu}_{T-t} \overleftarrow{\mu}_t) + \Delta \overleftarrow{\mu}_t - c[h_{T-t}, \rho_{T-t}] \overleftarrow{\mu}_t \\ \overleftarrow{\mu}_0 = \overrightarrow{\mu}_T \end{cases} \quad \text{(Surrogate Path)}$$

where, denoting the generator of the reverse OU process (2) by $\overleftarrow{L}_t^{OU}$ (see Eq. (45)), the reaction term is

$$c[h_{T-t}, \rho_{T-t}](x) = -\frac{1}{h_{T-t}(x)} \left[(\partial_t + \overleftarrow{L}_t^{OU}) h_{T-t} \right](x) + \frac{d}{dt} \log Z_{T-t}.$$

We suppose access to $\nabla \log \rho_t(x)$ via pretrained score network and also access to the reward and its spatial gradient, computing $\nabla \log \overrightarrow{\mu}_t = \nabla \log h_t + \nabla \log \rho_t$. We now introduce the notion of the *algorithm path* $\overleftarrow{\nu}_t$ with horizon T associated to the surrogate path $\overleftarrow{\mu}_t$, solving the reactionless equation

$$\begin{cases} \partial_t \overleftarrow{\nu}_t = -\nabla \cdot (x \overleftarrow{\nu}_t) - 2\nabla \cdot (\nabla \log \overrightarrow{\mu}_{T-t} \overleftarrow{\nu}_t) + \Delta \overleftarrow{\nu}_t \\ \overleftarrow{\nu}_0 = \gamma. \end{cases} \quad \text{(Algorithm Path)}$$

We find $\overleftarrow{\nu}_t = \mathrm{Law}(Y_t)$ for the associated SDE

$$\begin{cases} dY_t = \left[Y_t + 2\nabla \log \overrightarrow{\mu}_{T-t}(Y_t) \right] dt + \sqrt{2}\, dB_t, \\ Y_0 \sim \gamma. \end{cases} \quad \text{(Algorithm SDE)}$$

The difference between (Surrogate Path) and (Algorithm Path) is simply presence of the reaction term. By a Feynman–Kac argument (see Theorem 1(i) below), we can express their ratio as a weighted expectation over the paths (Algorithm SDE):

$$\frac{\overleftarrow{\mu}_T(x)}{\overleftarrow{\nu}_T(x)} \approx \mathbb{E}_{Y \sim \text{(Algorithm SDE)}} \left[\exp\left(-\int_0^T c[h_{T-t}, \rho_{T-t}](Y_t)\, dt \right) \middle| Y_T = x \right],$$

where the approximation is only the boundary approximation $\overrightarrow{\mu}_T(Y_0)/\gamma(Y_0) \approx 1$ (valid at large T as h_∞ is constant). Thus, the reaction term encodes the gap between the output of the algorithm $\mathrm{Law}(Y_T)$ and the true posterior μ_y. Note that due to this modification in (Algorithm Path), some amount of bias is unavoidable. Indeed in the worst case, any path beginning at the posterior μ_y and ending in a tractable distribution like γ, is generated by an evolution that we cannot compute in polynomial time, see Remark 1 and (Gupta et al., 2024). Nevertheless, some paths inspire useful approximations, that yield good empirical results; see for instance (Bruna & Han, 2024; Parulekar et al., 2026; Ren et al., 2026).

Canonical Interpolation. A canonical choice of (Surrogate Path) $\overrightarrow{\mu}^*$ is induced by the Doob h-transform (Doob, 1984),

$$\overrightarrow{\mu}^*_t(x) := \frac{h^*_t(x)\rho_t(x)}{Z_t}, \quad h^*_t(x) := \mathbb{E}\left[e^{R_y(X_0)} \;\middle|\; X_t = x \right], \tag{7}$$

where X_t is the OU process (1). This is the unique tilt h^*_t satisfying $h^*_0(x) = e^{R_y(x)}$ that also enjoys null reaction term $c[h^*_t, \rho_t] \equiv 0$ in (Surrogate Path). This zero reaction term implies that $\overleftarrow{\mu}^*_t = \overrightarrow{\mu}^*_{T-t}$ is exactly the marginal law of the denoising SDE (Algorithm SDE) with $\overrightarrow{\mu}_{T-t} = \overrightarrow{\mu}^*_{T-t}$. We call this induced path $\overrightarrow{\mu}^*$ the *canonical interpolation* as it is equal (by Bayes' rule) to the conditional law, $\overrightarrow{\mu}^*_t = \mathrm{Law}(X_t | \mathcal{A}(X_0) + \varepsilon = y)$ in the measurement case, and retains the correct endpoints $\overrightarrow{\mu}^*_0 = \mu_y$ and $\overrightarrow{\mu}^*_\infty = \gamma$. To simulate $\overrightarrow{\mu}^*$ via (Algorithm SDE) requires access to the score $\nabla \log \overrightarrow{\mu}^*_t = \nabla \log \rho_t + \nabla \log h^*_t$. While the score network supplies the first term, the latter requires us to solve for an optimal drift control (Fleming & Soner, 2006); this is expensive and must be done for each new choice of reward. For approximations to this optimal drift, see (Ren et al., 2026) for an online technique and (Guo et al., 2026) for an offline training method.

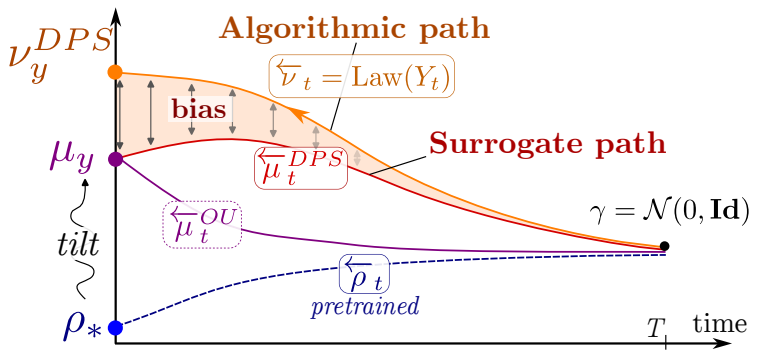


Figure 1: Paths between the endpoints ρ_* or μ_y (left) and γ (right). Blue, dotted: the forward OU path ρ_t and its reversal $\overleftarrow{\rho}_t$. Violet: the canonical interpolation $\overrightarrow{\mu}_t^*$ (OU started at μ_y), whose reversal is intractable at inference time. Red: the DPS surrogate $\overrightarrow{\mu}_t^{\mathrm{DPS}} = e^{R_y(\hat{x}_t)}\rho_t/Z_t$, with the same endpoints. Orange: the DPS algorithm path $\overleftarrow{\nu}_t^{\mathrm{DPS}}$ which ends at the biased ν_y^{DPS}.

Design Problem: identify a surrogate path $t \mapsto \overrightarrow{\mu}_t$ whose algorithm SDE is readily computable for any choice of differentiable reward, and yet also enjoys a small reaction term $c[h_t, \rho_t]$.

Diffusion Posterior Sampling. We show below that the DPS algorithm (Chung et al., 2023) can be interpreted as stemming from the following (Surrogate Path) (red line in Fig. 1).

$$\overrightarrow{\mu}_t^{\mathrm{DPS}}(x) = \frac{1}{Z_t} e^{R_y(\hat{x}_t(x))}\rho_t(x), \qquad \text{(DPS Surrogate Path)}$$

which retains the correct endpoints $\overrightarrow{\mu}_0^{\mathrm{DPS}} = \mu_y$, $\overrightarrow{\mu}_\infty^{\mathrm{DPS}} = \gamma$. Crucially, the score $\nabla \log \overrightarrow{\mu}_t^{\mathrm{DPS}}$ can be written in terms of

$$\hat{x}_t(x) \coloneqq \mathbb{E}[X_0 \mid X_t = x], \qquad \Sigma_t(x) \coloneqq \mathrm{Cov}(X_0 \mid X_t = x),$$

both of which are computable from the diffusion score oracle s_{θ_*} and its Jacobian ∇s_{θ_*} via Tweedie's formula (Robbins, 1956), see Section E.1. Heuristically, the (DPS Surrogate Path) is obtained from the canonical interpolation by swapping the conditional expectation *inside* the exponential (7) for

$$h_t^{\mathrm{DPS}}(x) = e^{R_y(\mathbb{E}[X_0 | X_t = x])} \neq h_t^*(x) = \mathbb{E}[e^{R_y(X_0)} \mid X_t = x].$$

We can identify the true reversal of the (DPS Surrogate Path),

$$\partial_t \overleftarrow{\mu}_t^{\mathrm{DPS}} = \Delta \overleftarrow{\mu}_t^{\mathrm{DPS}} - 2\nabla\cdot\left(\nabla \log \overrightarrow{\mu}_{T-t}^{\mathrm{DPS}} \overleftarrow{\mu}_t^{\mathrm{DPS}}\right) - \nabla\cdot(x \overleftarrow{\mu}_t^{\mathrm{DPS}}) - c_{\mathrm{DPS}}(T-t, x)\overleftarrow{\mu}_t^{\mathrm{DPS}}, \qquad \text{(DPS Surrogate PDE)}$$

with an explicit reaction coefficient

$$c_{\mathrm{DPS}}(t,x) = -\frac{\mathrm{tr}\left(\Sigma_t(x)(D^2 R_y)(\hat{x}_t(x))\Sigma_t(x)\right) + \left\|\Sigma_t(x)\nabla R_y(\hat{x}_t(x))\right\|^2}{(e^t - e^{-t})^2} - \frac{d}{dt}\log Z_t. \qquad (8)$$

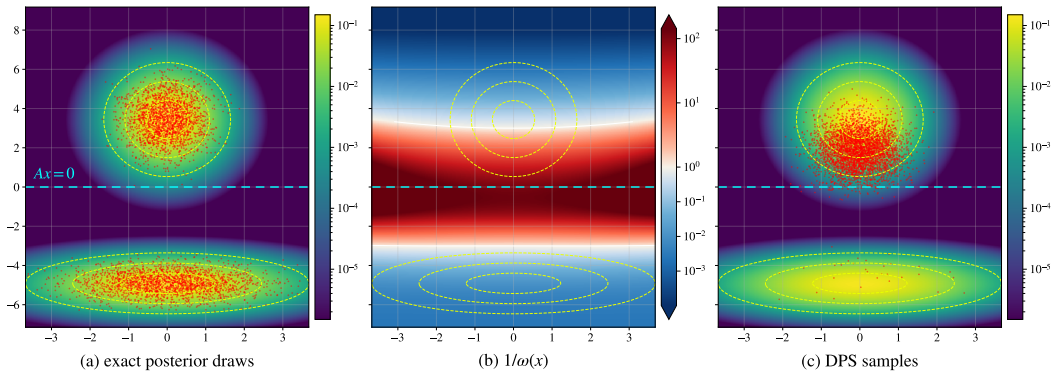
The DPS algorithmic path. The difficulty with implementing Equation (DPS Surrogate PDE) as an SDE is the reaction term. We can construct an alternate PDE that has the same transport and diffusion term but no reaction term

$$\partial_t \overleftarrow{\nu}_t^{\mathrm{DPS}} = \Delta \overleftarrow{\nu}_t^{\mathrm{DPS}} - 2\nabla\cdot\left(\nabla \log \overrightarrow{\mu}_{T-t}^{\mathrm{DPS}} \overleftarrow{\nu}_t^{\mathrm{DPS}}\right) - \nabla\cdot(x \overleftarrow{\nu}_t^{\mathrm{DPS}}). \qquad \text{(DPS Path)}$$

Using the identities of Section E.1, for the score $\nabla \log \overrightarrow{\mu}_t^{\mathrm{DPS}}$, we get the (Algorithm SDE) approximated by the DPS algorithm as

$$\begin{cases} dY_t = \left(Y_t + 2\nabla \log \rho_{T-t}(Y_t) + \frac{2}{e^{T-t} - e^{-(T-t)}} \Sigma_{T-t}(Y_t)\nabla R_y(\hat{x}_{T-t}(Y_t))\right) dt + \sqrt{2}\, dB_t, \\ Y_0 \sim \gamma, \end{cases} \qquad \text{(DPS SDE)}$$

Applying the Feynman–Kac formula (6) to the quotient $\frac{\overleftarrow{\nu}_T^{\mathrm{DPS}}}{\overleftarrow{\mu}_T^{\mathrm{DPS}}}$, we obtain the following characterization of the bias of the DPS algorithm.



(a) exact posterior draws (b) $1/\omega(x)$ (c) DPS samples

Figure 2: Two-mode Gaussian mixture prior; constraint $\{x_2 = 0\}$ is the dashed line; background in all panels is the exact posterior density μ_y. **(a)** Draws from the true posterior μ_y (red). Mode shares $\approx 50\%/50\%$. **(b)** The ratio $1/\omega = \nu_y^{\mathrm{DPS}}/\mu_y$, log scale, from a finite-difference solve of (19). Red: $1/\omega > 1$, DPS over-samples. Blue: $1/\omega < 1$, DPS under-samples. White contour: $1/\omega = 1$. **(c)** Draws generated by (DPS SDE) (red). Mode shares $\approx 99\%/1\%$. Two effects are visible in (c) and predicted by (b): the mode nearer the constraint receives 10^2 times more samples than the other, and inside each mode the samples are shifted toward $x_2 = 0$ relative to (a). Shares are computed from all N samples. Setup and solver: Section A.

Theorem 1. *The terminal law $\nu_y^{\mathrm{DPS}} := \overleftarrow{\nu}_T^{\mathrm{DPS}}$ of the DPS-SDE (DPS SDE) differs from the true posterior μ_y by a pointwise multiplicative weight:*

$$\mu_y(x) = \omega(x)\, \nu_y^{\mathrm{DPS}}(x). \tag{9}$$

The weight ω admits two equivalent Feynman–Kac representations in terms of c_{DPS} defined in (8):

(i) ***Backward path*** *(condition on the DPS denoising process arriving at $Y_T = x$):*

$$\omega(x) = \mathbb{E}_{Y\sim(\text{DPS SDE})}\left[\frac{\overrightarrow{\mu}_T^{\mathrm{DPS}}(Y_0)}{\gamma(Y_0)}\,\exp\Big(-\int_0^T c_{\mathrm{DPS}}(T-s, Y_s)\,ds\Big)\;\Big|\; Y_T = x\right]. \tag{10}$$

(ii) ***Forward path*** *(condition on the OU process (1) starting at $X_0 = x$):*

$$\frac{1}{\omega(x)} = \mathbb{E}_{X\sim\mathrm{OU}}\left[\frac{\gamma(X_T)}{\overrightarrow{\mu}_T^{\mathrm{DPS}}(X_T)}\,\exp\Big(\int_0^T c_{\mathrm{DPS}}(s, X_s)\,ds\Big)\;\Big|\; X_0 = x\right]. \tag{11}$$

Both path functionals are expressible in terms of quantities obtainable from the score oracle and its Jacobian via Tweedie's formula. Importance-weighting DPS samples by ω recovers μ_y exactly.

See Section E for the proof of Theorem 1.

Discussion of Theorem 1. By (9), $\nu_y^{\mathrm{DPS}} = \omega^{-1}\mu_y$: DPS over-samples where $\omega < 1$ and under-samples where $\omega > 1$. For a linear-Gaussian reward $R_y(x) = -\frac{\alpha}{2}\|\mathcal{A}x - y\|^2$, (8) and (34) give, with $J_s := \nabla\hat{x}_s$ and $r_s := \mathcal{A}\hat{x}_s - y$,

$$c_{\mathrm{DPS}}(s, x) = \alpha\|\mathcal{A}J_s\|_F^2 - \alpha^2\|J_s\mathcal{A}^\top r_s\|^2 - \tfrac{d}{ds}\log Z_s, \tag{12}$$

which is largest when the Tweedie mean satisfies the constraint and decreases with the residual. In (11), c_{DPS} is integrated along the forward OU path from x. Since $\|J_s\|$ decays for $s \gtrsim 1$, small s dominates, where $X_s \approx x$ and $r_s \approx r := \mathcal{A}x - y$; substituting,

$$\frac{1}{\omega(x)} \approx C\, e^{-\alpha^2 r^\top K(x)\, r}, \qquad K(x) := \int_0^T (\mathcal{A}J_s)(\mathcal{A}J_s)^\top ds \succeq 0. \tag{13}$$

DPS therefore samples $e^{R_y}\rho_*$ with the likelihood precision αI replaced by $\alpha I + 2\alpha^2 K(x)$: the measurement is treated as less noisy than it is, most strongly along measurement directions tangent to the data manifold, where $J_s \to P_{T\mathcal{M}}$ (Section C.2.1).

In Fig. 2, $\mathcal{A} = e_2^\top$, $y = 0$, $\alpha = 0.4$, so K is the scalar $\kappa(x) = \int_0^T \|\partial_{x_2} \hat{x}_s(x)\|^2 ds$ and (13) reads $\nu_y^{\mathrm{DPS}} \propto e^{\beta R_y} \rho_*$ with $\beta = 1 + 2\alpha\kappa(x) > 1$. This explains both panels: (1) the mode whose center is further from $x_2 = 0$ (lower, $x_2 \approx -4.9$) has smaller R_y and loses mass to the nearer one (upper, $x_2 \approx 3.4$), giving the 10^2 ratio in shares; (2) within each mode, samples shift toward $x_2 = 0$, giving the skew in (c) relative to (a).

4 Bias Reduction

We see in Theorem 1 that the ratio between the density of the DPS sampler and the true posterior can be expressed (approximately) as: $\frac{1}{\omega(x)} \approx \mathbb{E}_{X \sim \mathrm{OU}} \left[e^{\int_0^T c_{\mathrm{DPS}}(s, X_s)\, ds} \,\middle|\, X_0 = x \right]$.

This gives a clear design goal: it is beneficial to design paths that result in small variations in c_{DPS} over the trajectories. This would correspond to a smaller reaction term in Equation (Surrogate Path), and a smaller bias when we implement the corresponding algorithm. For instance, we can add an extra potential drift $-\vartheta \nabla U$ to the SDE (Algorithm SDE) that drives trajectories to regions where c_{DPS} has small oscillations. In terms of the algorithm (Algorithm Path), this amounts to solving

$$\partial_t \overleftarrow{\nu}_t = \underbrace{\Delta \overleftarrow{\nu}_t - \nabla \cdot (x \overleftarrow{\nu}_t) - 2\nabla \cdot (\nabla \log \overrightarrow{\mu}_{T-t}^{\mathrm{DPS}} \overleftarrow{\nu}_t)}_{\text{Equation (DPS Path)}} + \underbrace{\vartheta \nabla \cdot (\nabla U \overleftarrow{\nu}_t)}_{\text{extra guidance}},$$

where the drift intensity $\vartheta \geq 0$ is a hyperparameter. As we see below, such a change in drift can readily be matched with a corresponding change in reaction term that reinterprets the (Surrogate Path) with the updated drift and a modified reaction term.

Interpreting the drift as a reaction term. Just as diffusion can be recast as a drift involving the score, $\Delta \rho_t = \nabla \cdot (\nabla \log \rho_t \, \rho_t)$, the additional drift generated by ∇U can be recast as a reaction term through the tautological identity:

$$c_U \;=\; \frac{\nabla \cdot (\nabla U \, \overleftarrow{\mu}_t^{\mathrm{DPS}})}{\overleftarrow{\mu}_t^{\mathrm{DPS}}} \;=\; \Delta U + \nabla U \cdot \nabla \log \overrightarrow{\mu}_{T-t}^{\mathrm{DPS}}.$$

In other words, we can rewrite the (DPS Surrogate PDE) as:

$$\begin{aligned} \partial_t \overleftarrow{\mu}_t^{\mathrm{DPS}} = \Delta \overleftarrow{\mu}_t^{\mathrm{DPS}} - 2\,\nabla \cdot \big(\nabla \log \overrightarrow{\mu}_{T-t}^{\mathrm{DPS}} \overleftarrow{\mu}_t^{\mathrm{DPS}}\big) - \nabla \cdot \big(x \overleftarrow{\mu}_t^{\mathrm{DPS}}\big) \\ + \vartheta\, \nabla \cdot \big(\nabla U \overleftarrow{\mu}_t^{\mathrm{DPS}}\big) - \big(c_{\mathrm{DPS}}(T-t, x) + \vartheta\, c_U\big) \overleftarrow{\mu}_t^{\mathrm{DPS}} \end{aligned}$$

In the language of Theorem 1 (with the DPS SDE augmented by the drift $-\vartheta \nabla U$), this modifies the reaction term to $c_{\mathrm{eff}} = c_{\mathrm{DPS}} + \vartheta\, c_U$. As a consequence, excessively large ϑ is counterproductive: the reaction term becomes dominated by $\vartheta\, c_U$ and the original bias structure is lost.

Remark 1. There exists in principle a potential U^* satisfying $c_{\mathrm{DPS}} + c_{U^*} = 0$ (taking $\vartheta = 1$), which would eliminate the bias exactly. Computing U^* directly is exponentially slow; recent work instead approximates it via a variational characterizations, training a non-linear (Guo et al., 2026) or linear (Ren et al., 2026) neural network for each specific reward.

STSL as a special case. Originally motivated by expanding the h-transform $h_t^*(x)$ (7) to second order in X_t to approximate the optimal (reaction free) drift, STSL (Rout et al., 2024) chooses a potential U that drives the trajectory toward low-uncertainty regions of the initial condition X_0. Up to constants, the choice is

$$U(t, x) = \mathrm{tr}(\Sigma_t(x)) = \sum_{i=1}^{d} \lambda_i(t, x) \;\geq\; 0.$$

where $\lambda_i(t, x) \geq 0$ are the eigenvalues of $\Sigma_t(x)$. Since the λ_i are non-negative, smaller values of U correspond to smaller spread in the dominant eigendirections of $\Sigma_t(x)$, which in turn flattens the spatially varying part of c_{DPS} in (8). In practice, this yields a better algorithm with reduced output uncertainty (Rout et al., 2024).

5 Numerical Instabilities of the DPS Algorithm

As discussed in earlier sections, the algorithm evolution introduces a bias when compared to the surrogate evolution. In this section, we study a different issue with the actual implementation of the DPS algorithm, namely instability of the evolution close to the constraint manifold. In the context of the DPS algorithm proposed in Chung et al. (2023), we show that the instability unavoidably occurs in the space parallel to the data manifold due to the systematic violation of the forward Euler stability condition. This phenomenon has indeed been observed in practice (Fig. 3). To mitigate this, a common practice is to "turn off" reward guidance close to the data manifold, which we refer to as **early guidance stopping**, see Algorithm 2. In words, the stochastic evolution starts with both the score (corresponding to the untilted prior) and reward guidance drift terms until some intermediate time $t_{\text{stop}} \in (0, T)$, after which the diffusion proceeds with only the untilted score. We show that early guidance stopping can be explicitly characterized as an appropriately weighted tilt of the prior.

Instability of DPS. We first examine the algorithmic implementation of the DPS algorithm, which is conceptualized as an approximation to the solution of the SDE (DPS SDE). The exact algorithm proposed by the authors of (Chung et al., 2023) is given in Section D. The DPS algorithm progressively denoises over discrete time-steps, with the reward guidance weighted at each time-step through a guidance schedule $\{\zeta_i\}_{i=1}^N$ that is taken to be trajectory-dependent,

$$\zeta_i = \frac{\alpha}{\|y - \mathcal{A}(\hat{x}_{t_i}(x_i))\|_2}, \tag{14}$$

where x_i is the iterate at step i, corresponding to time t_i, and $\hat{x}_{t_i}(x_i)$ is its Tweedie mean, $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^L$ is a general observation operator, $y \in \mathbb{R}^L$ is the observation, and $\alpha \in [0.2, 1]$ is a problem-dependent hyperparameter. In practice, the guidance schedule has a dramatic performance impact.

A first observation is that (14) does not account for the time discretization Δt_i of the SDE (DPS SDE); effectively, this corresponds to multiplying the guidance vector field by a time-dependent factor. In terms of the surrogate (Surrogate Path), this corresponds to the curve

$$t \mapsto \overrightarrow{\mu}_t^\eta = \frac{e^{-\alpha \eta_t \|\mathcal{A}(\hat{x}_t(x)) - y\|_2} \rho_t}{Z_t}, \tag{15}$$

that is, $\overrightarrow{\mu}_t^\eta = e^{\alpha \eta_t R_y(\hat{x}_t)} \rho_t / Z_t$ with reward $R_y(x) := -\|\mathcal{A}(x) - y\|_2$, and annealing schedule (for the linear noising schedule $\{\beta_i\}_{i=1}^{1000}$ used in the classical DDPM, see Section G for details) given by

$$\eta_t \approx \frac{1.8 \cdot 10^4}{\sqrt{1 + 6600\, t}}. \tag{16}$$

The key takeaway is that the schedule weight is large for a reasonable choice of hyperparameters, with the qualitative implication of *strong enforcement of measurement constraints* as we approach the data manifold (small t / low-temperature regime). Indeed notice that this path-dependent guidance schedule yields the target density of:

$$\mu_y^{\text{Target}} = \frac{e^{-\alpha\, 1.8 \cdot 10^4 \|\mathcal{A}(x) - y\|_2} \rho_*}{Z^{\text{Target}}}, \tag{17}$$

in which the reward is unsquared and the constraint $\{\mathcal{A}(x) = y\}$ has a large weight.

Inevitable Oscillations. The unsquared residual in target (17) distorts the dynamics in a way that no choice of step size can repair. To see this, consider the one-dimensional example where gradient flow on $|x|$ under forward Euler is: $x_{n+1} = x_n - \Delta t \operatorname{sign}(x_n)$. The gradient $\operatorname{sign}(x)$ has unbounded Lipschitz constant at the origin, so any $\Delta t > 0$ produces a limit cycle of amplitude $\sim \Delta t$ around the minimum; this is bounded but non-convergent. The DPS guidance integration is the multidimensional analogue.

As Y_t approaches the constraint $\{\mathcal{A}(x) = y\}$, the guidance gradient $\nabla_Y \|\mathcal{A}(\hat{x}_{T-t}(Y)) - y\|_2 = \nabla \hat{x}_{T-t}(Y)^\top \nabla \mathcal{A}(\hat{x}_{T-t}(Y))^\top (\mathcal{A}(\hat{x}_{T-t}(Y)) - y) / \|\mathcal{A}(\hat{x}_{T-t}(Y)) - y\|_2$ does not vanish, while the annealing schedule (16) multiplying the drift cancels the Euler step size exactly (see Section H). The forward Euler stability criterion is therefore inevitably violated near the constraint, and the iteration enters a limit cycle tangent to the data manifold $\mathcal{M}$, of amplitude $\sim \alpha\, \sigma_{\max}(\nabla \mathcal{A}\, P_{T_p \mathcal{M}})$. Here

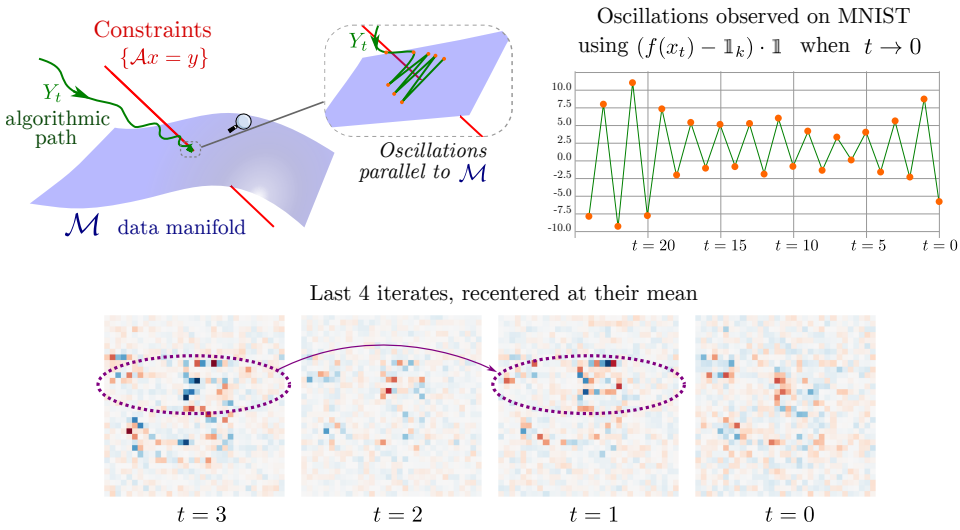


Figure 3: **DPS oscillates near the data manifold. (Top left)** Schematic: near the manifold, the large effective guidance triggers oscillations. **(Top right)** Oscillations in an MNIST posterior sampling task, with guidance toward the digit 3. **(Bottom)** The last four DPS iterates, centered at their mean: alternate iterates deviate alike, a period-two oscillation. Details in Fig. 6 and Section I.

$P_{T_p\mathcal{M}}$ is the orthogonal projection onto the tangent space $T_p\mathcal{M}$ of $\mathcal{M}$ at the nearest point $p \in \mathcal{M}$, the small-noise limit of $\nabla \hat{x}_t$ (Section C.2.1), and $\sigma_{\max}$ denotes the largest singular value. The advantage of the unsquared residual $\|\cdot\|_2$ over $\|\cdot\|_2^2$ in the reward is that the iterates remain *semi-stable* in the sense of Lyapunov: they settle into a limit cycle at distance at most $\alpha \|\nabla \mathcal{A}\|_{\mathrm{op}}$ from the constraint manifold. By contrast, when the forward Euler stability criterion is violated for $\|\cdot\|_2^2$, the oscillations diverge.

Early Guidance Stopping. To avoid these numerical instabilities, practitioners apply early guidance stopping (Algorithm 2), terminating the guidance at some intermediate time $t_{\mathrm{stop}} \in (0, T)$. Combined with Theorem 1, this recovers the output of DPS with early guidance stopping.

Theorem 2 (Early Guidance Stopping)**.** *If guidance is stopped at time $t_{\mathrm{stop}} = T - t_*$, the output of the DPS algorithm, up to time-discretization error, is given by*

$$\nu_y^{\mathrm{DPS},t_*}(x) = \frac{\mathbb{E}_{X\sim \mathrm{OU}}\left[\frac{1}{\omega_{t_*}}(X_{t_*})\, e^{\alpha \eta_{t_*} R_y(\hat{x}_{t_*}(X_{t_*}))} \,\middle|\, X_0 = x\right]}{Z_{t_*}} \rho_*(x),$$

where

$$\frac{1}{\omega_{t_*}}(x) := \mathbb{E}_{X\sim \mathrm{OU}}\left[\frac{\gamma(X_{T-t_*})}{\overrightarrow{\mu}_T^{\eta}(X_{T-t_*})} \exp\left(\int_0^{T-t_*} c^*_{\mathrm{DPS}}(t_* + s, X_s)\, ds\right) \,\middle|\, X_0 = x\right], \tag{18}$$

with

$$c^*_{\mathrm{DPS}}(t,x) = -\frac{\alpha\eta_t \,\mathrm{tr}\big(\Sigma_t(x)(D^2 R_y)(\hat{x}_t(x))\Sigma_t(x)\big) + \alpha^2\eta_t^2 \left\|\Sigma_t(x)\nabla R_y(\hat{x}_t(x))\right\|^2}{(e^t - e^{-t})^2}$$
$$+ \alpha \frac{d\eta_t}{dt} R_y(\hat{x}_t(x)) - \frac{d}{dt}\log Z_t,$$

that is, (8) with R_y replaced by $\alpha\eta_t R_y$, plus a term from the time dependence of η_t. Here α, R_y, the ***annealing schedule*** *η_t and Z_t are as in (14)–(16).*

See for instance (Huang et al., 2026, Proposition 2.6) for the effect of early guidance stopping in the simpler linear-quadratic case.

Reproducibility Statement

All theoretical results are stated with their assumptions, and complete proofs are given in the appendix. Our 2-dimensional Gaussian mixture experiments use standard methods run in their standard form, including DDPM ancestral sampling with exact scores and jacobians, and the calculation of $\frac{1}{\omega}$ for Figure 2 is a standard finite-difference scheme with full configuration given in Section A. The MNIST experiments run the standard DPS algorithm of Chung et al. (2023) as specified in Section D, and their configurations are summarized in Section I.

AI Use Statement

In this work, we used generative AI tools to assist in the writing and checking of proofs (verifying derivations, identifying and correcting errors in intermediate formulas, and drafting revised versions of appendix proofs), to provide feedback on the research and its presentation, and to implement methods used in our numerical experiments. We have not used generative AI tools to formulate the main mathematical claims or conceptual framework, to propose hypotheses or interpret results. Generating synthetic data sets, translation and qualitative or thematic data analysis are not applicable to this work. Additionally, we used generative AI tools to edit the paper for readability, draft parts of the text, to create and edit software code, to search for and identify relevant literature, and to check and format references. We have reviewed all AI-assisted work: every line has been examined and revised by the authors and we take responsibility for the final content of this work, including text, claims and artifacts produced with the aid of generative AI.

References

Tara Akhound-Sadegh, Jungyoon Lee, Joey Bose, Valentin De Bortoli, Arnaud Doucet, Michael M. Bronstein, Dominique Beaini, Siamak Ravanbakhsh, Kirill Neklyudov, and Alexander Tong. Progressive inference-time annealing of diffusion models for sampling from Boltzmann densities. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Brian D. O. Anderson. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982. doi: 10.1016/0304-4149(82)90051-5.

Gautham Govind Anil, Shaan Ul Haque, Nithish Kannen, Dheeraj Mysore Nagaraj, Sanjay Shakkottai, and Karthikeyan Shanmugam. Fine-tuning diffusion models via intermediate distribution shaping. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://arxiv.org/abs/2510.02692.

Benjamin Boys, Mark Girolami, Jakiw Pidstrigach, Sebastian Reich, Alan Mosca, and O. Deniz Akyildiz. Tweedie moment projected diffusions for inverse problems. *Transactions on Machine Learning Research*, 2024. URL https://openreview.net/forum?id=4unJi0qrTE.

Joan Bruna and Jiequn Han. Provable posterior sampling with denoising oracles via tilted transport. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://arxiv.org/abs/2407.00745.

Sitan Chen, Sinho Chewi, Jerry Li, Yuanzhi Li, Adil Salim, and Anru R. Zhang. Sampling is as easy as learning the score: theory for diffusion models with minimal data assumptions. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2209.11215.

Hyungjin Chung, Byeongsu Sim, Dohoon Ryu, and Jong Chul Ye. Improving diffusion models for inverse problems using manifold constraints. In Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho (eds.), *Advances in Neural Information Processing Systems*, 2022. URL https://openreview.net/forum?id=nJJjv0JDJju.

Hyungjin Chung, Jeongsol Kim, Michael T. McCann, Marc L. Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=OnD9zGAGT0k.

Giannis Daras, Hyungjin Chung, Chieh-Hsin Lai, Yuki Mitsufuji, Jong Chul Ye, Peyman Milanfar, Alexandros G Dimakis, and Mauricio Delbracio. A survey on diffusion models for inverse problems. *arXiv preprint arXiv:2410.00083*, 2024.

Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer, Berlin, Heidelberg, 2nd edition, 2010. doi: 10.1007/978-3-642-03311-7. Corrected reprint of the second (1998) edition.

Prafulla Dhariwal and Alex Nichol. Diffusion models beat GANs on image synthesis. In *Advances in Neural Information Processing Systems*, volume 34, pp. 8780–8794, 2021. URL https://arxiv.org/abs/2105.05233.

J. L. Doob. *Classical Potential Theory and Its Probabilistic Counterpart*. Springer, 1984.

Zehao Dou and Yang Song. Diffusion posterior sampling for linear inverse problem solving: A filtering perspective. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=tplXNcHZs1.

Wendell H. Fleming and H. Mete Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer, 2nd edition, 2006.

Zhengyi Guo, Wenpin Tang, and Renyuan Xu. Conditional diffusion guidance under hard constraint: A stochastic analysis approach. *arXiv preprint arXiv:2602.05533*, 2026.

Shivam Gupta, Ajil Jalal, Aditya Parulekar, Eric Price, and Zhiyang Xun. Diffusion posterior sampling is computationally intractable. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 17020–17059. PMLR, 21–27 Jul 2024. URL https://proceedings.mlr.press/v235/gupta24a.html.

U. G. Haussmann and E. Pardoux. Time reversal of diffusions. *The Annals of Probability*, 14(4): 1188–1205, 1986. doi: 10.1214/aop/1176992362.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851, 2020. URL https://arxiv.org/abs/2006.11239.

Jerry Y. Huang, Justin Lin, Sheel Shah, Kartik Nair, and Nicholas M. Boffi. How to guide your flow: Few-step alignment via flow map reward guidance. In *International Conference on Machine Learning (ICML)*, 2026. URL https://arxiv.org/abs/2604.27147. arXiv:2604.27147.

Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991. ISBN 0-387-97655-8. doi: 10.1007/978-1-4612-0949-2. URL https://doi.org/10.1007/978-1-4612-0949-2.

Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models. In *Advances in Neural Information Processing Systems*, volume 35, pp. 23593–23606, 2022.

O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva. *Linear and Quasi-linear Equations of Parabolic Type*, volume 23 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1968.

Holden Lee, Jianfeng Lu, and Yixin Tan. Convergence for score-based generative modeling with polynomial complexity. In *Advances in Neural Information Processing Systems*, volume 35, 2022. URL https://arxiv.org/abs/2206.06227.

Ankur Moitra, Andrej Risteski, and Dhruv Rohatgi. Steering diffusion models with quadratic rewards: a fine-grained analysis. In *Proceedings of Thirty Ninth Conference on Learning Theory*, volume 336 of *Proceedings of Machine Learning Research*, pp. 5188–5209. PMLR, 2026. URL https://proceedings.mlr.press/v336/moitra26a.html.

Badr Moufad, Yazid Janati, Lisa Bedin, Alain Oliviero Durmus, Randal Douc, Eric Moulines, and Jimmy Olsson. Variational diffusion posterior sampling with midpoint guidance. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=6EUtjXAvmj.

Advait Parulekar, Litu Rout, Karthikeyan Shanmugam, and Sanjay Shakkottai. Efficient approximate posterior sampling with annealed Langevin Monte Carlo. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://arxiv.org/abs/2508.07631.

Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. Zero-shot text-to-image generation. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 8821–8831. PMLR, 2021. URL https://proceedings.mlr.press/v139/ramesh21a.html.

Yinuo Ren, Wenhao Gao, Lexing Ying, Grant M. Rotskoff, and Jiequn Han. DriftLite: Lightweight drift control for inference-time scaling of diffusion models. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=l01eG3Qikl.

Herbert Robbins. An empirical Bayes approach to statistics. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pp. 157–163, Berkeley, CA, 1956. University of California Press.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10684–10695, 2022.

Litu Rout, Negin Raoof, Giannis Daras, Constantine Caramanis, Alexandros G. Dimakis, and Sanjay Shakkottai. Solving linear inverse problems provably via posterior sampling with latent diffusion models. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Litu Rout, Yujia Chen, Abhishek Kumar, Constantine Caramanis, Sanjay Shakkottai, and Wen-Sheng Chu. Beyond first-order Tweedie: Solving inverse problems using latent diffusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9472–9481, 2024. doi: 10.1109/CVPR52733.2024.00905.

Litu Rout, Yujia Chen, Nataniel Ruiz, Abhishek Kumar, Constantine Caramanis, Sanjay Shakkottai, and Wen-Sheng Chu. RB-Modulation: Training-free stylization using reference-based modulation. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=bnINPG5A32.

Chitwan Saharia, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily L. Denton, Kamyar Ghasemipour, Raphael Gontijo Lopes, Burcu Karagol Ayan, Tim Salimans, Jonathan Ho, David J. Fleet, and Mohammad Norouzi. Photorealistic text-to-image diffusion models with deep language understanding. In *Advances in Neural Information Processing Systems*, volume 35, pp. 36479–36494, 2022. URL https://arxiv.org/abs/2205.11487.

Marta Skreta, Tara Akhound-Sadegh, Viktor Ohanesian, Roberto Bondesan, Alán Aspuru-Guzik, Arnaud Doucet, Rob Brekelmans, Alexander Tong, and Kirill Neklyudov. Feynman-Kac correctors in diffusion: Annealing, guidance, and product of experts. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 55906–55949. PMLR, 2025. URL https://proceedings.mlr.press/v267/skreta25a.html.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 2256–2265. PMLR, 2015. URL https://arxiv.org/abs/1503.03585.

Jiaming Song, Arash Vahdat, Morteza Mardani, and Jan Kautz. Pseudoinverse-guided diffusion models for inverse problems. In *International Conference on Learning Representations (ICLR)*, May 2023.

Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://arxiv.org/abs/1907.05600.

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.

Yang Song, Liyue Shen, Lei Xing, and Stefano Ermon. Solving inverse problems in medical imaging with score-based generative models. In *International Conference on Learning Representations*, 2022. URL https://arxiv.org/abs/2111.08005.

S. R. S. Varadhan. Asymptotic probabilities and differential equations. *Communications on Pure and Applied Mathematics*, 19(3):261–286, 1966. doi: 10.1002/cpa.3160190303.

Santosh S. Vempala and Andre Wibisono. Rapid convergence of the unadjusted Langevin algorithm: Isoperimetry suffices. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://arxiv.org/abs/1903.08568.

Pascal Vincent. A connection between score matching and denoising autoencoders. *Neural Computation*, 23(7):1661–1674, 2011.

Luhuan Wu, Brian L. Trippe, Christian A. Naesseth, David M. Blei, and John P. Cunningham. Practical and asymptotically exact conditional sampling in diffusion models. In *Advances in Neural Information Processing Systems*, volume 36, pp. 31372–31403, 2023. URL https://arxiv.org/abs/2306.17775.

Xingyu Xu and Yuejie Chi. Provably robust score-based diffusion posterior sampling for plug-and-play image reconstruction. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://arxiv.org/abs/2403.17042.

A Experimental Details for Fig. 2

Results. Mode shares are reported as upper/lower. **(a)** $N = 10^5$ exact posterior draws split $50.1\%/49.9\%$. **(b)** The weight $1/\omega = d\nu_y^{\mathrm{DPS}}/d\mu_y$, on a log scale. Warm colors predict over-sampling by DPS, cool colors predict under-sampling, and the white contour is $1/\omega = 1$. Integrating $1/\omega$ against μ_y predicts $99.05\%/0.95\%$, odds of $104:1$. As a check on the solve, $\int (1/\omega)\, d\mu_y = 1$ evaluates to 1.087 on the posterior support. **(c)** $N = 10^5$ DPS samples split $99.07\%/0.93\%$, odds of $106:1$, matching (b) in both direction and magnitude. Panels (a) and (c) show a random subset of 4000 samples, while all shares use the full 10^5.

Setup. The prior is a two-component Gaussian mixture on $\mathbb{R}^2$ with means $(0, \pm 5.5)$ and covariances $\mathrm{diag}(0.3, 1.5)$ (upper) and $\mathrm{diag}(1.5, 0.3)$ (lower). The measurement is $\mathcal{A}(x) = e_2^\top x$ with noise precision $\alpha = 0.4$ and observation $y = 0$, so $R_y(x) = -\frac{\alpha}{2} x_2^2$. The prior weights $0.191/0.809$ are chosen so that the two components of the posterior, which is again a Gaussian mixture in closed form, carry equal mass. Dashed rings are the 1, 2, 3σ ellipses of these posterior components. DPS runs (DPS SDE) by DDPM ancestral sampling on the linear schedule of Section G (1000 steps), using the mixture's exact score, Tweedie mean and Jacobian, so no network is trained. Mode shares are sample averages of the component responsibilities $r_k(x) = p_k \mathcal{N}(x; m_k, \Sigma_k)/\rho_*(x)$, which the tilt e^{R_y} leaves unchanged.

Computing $1/\omega$. By Lemma 6, the ratio $\psi_t := \overleftarrow{\nu}_t / \overleftarrow{\mu}_t$ of the DPS law to the surrogate path solves

$$\partial_t \psi_t = L\psi_t + c_{\mathrm{DPS}}(T - t, \cdot)\, \psi_t, \qquad \psi_0 = \gamma / \overrightarrow{\mu}_T^{\mathrm{DPS}}, \qquad Lf = \Delta f - x \cdot \nabla f, \tag{19}$$

and its terminal value is $1/\omega$: at $t = T$, $\overleftarrow{\nu}_T = \nu_y^{\mathrm{DPS}}$ and $\overleftarrow{\mu}_T = \mu_y$, so $\psi_T = \nu_y^{\mathrm{DPS}}/\mu_y = 1/\omega$ by (9). We solve (19) numerically rather than estimating (11) by Monte Carlo. The term $-\frac{d}{ds} \log Z_s$ in c_{DPS} is constant in space and only rescales the solution. Using $\rho_T \approx \gamma$, we therefore drop it and start instead from $\psi_0 = Z_0 e^{-R_y(\hat{x}_T)}$, which gives the same ψ_T. The remaining reaction term, Z_0, and $\hat{x}_T$ are all available in closed form for a Gaussian mixture. The horizon matches the sampler, $T \approx 5.06$.

Solver. We discretize space with a uniform 801×801 grid covering the plotted window plus a margin of 4 on each side, and impose reflecting (zero-flux) boundary conditions at its edges. The margin makes the boundary irrelevant: OU paths started in the window rarely reach it. Time steps are the 999 OU time steps of the DDPM sampler. They are finest near the data, where the reaction term changes fastest, and they put the solver and the sampler on the same clock. Each step treats the two parts of (19) separately. The reaction part is a pointwise multiplication by an exponential, which is exact. The diffusion part is a Crank–Nicolson step, taken one coordinate at a time, so each step reduces to tridiagonal linear solves. Because (19) is linear, we rescale ψ to unit maximum after every step and track the scale separately, which avoids numerical overflow. The mode shares and the normalization check in (b) come from integrating $1/\omega$ against μ_y on the grid, restricted to the region where the posterior density exceeds 10^{-7} of its maximum.

B Feynman–Kac representations of the Radon–Nikodym derivative

We consider two densities π_t, π'_t evolving under a pair of equations of the form (20) below. The surrogate path $\overleftarrow{\mu}_t$ (Surrogate Path) and the algorithm path $\overleftarrow{\nu}_t$ (Algorithm Path) are such an example, with $\pi = \overleftarrow{\mu}$, $\pi' = \overleftarrow{\nu}$, a common transport field, and $f' \equiv 0$.

Lemma 2 (Feynman–Kac for Density Ratio)**.** *Let $T > 0$ and $\alpha \in (0, 1)$. Consider two initial measures with Lebesgue densities $\pi_0(x)\, dx$ and $\pi'_0(x)\, dx$, where $\pi_0, \pi'_0 \in C^{2+\alpha}_{loc}(\mathbb{R}^d)$ and $\pi_0(x), \pi'_0(x) > 0$ for all $x \in \mathbb{R}^d$. Suppose the Radon–Nikodym derivative $g_0 := d\pi'_0/d\pi_0 = \pi'_0/\pi_0 \in C^2(\mathbb{R}^d)$ satisfies the growth bound $|g_0(x)| + \|\nabla g_0(x)\| + \|D^2 g_0(x)\| \le C\, e^{\lambda \|x\|^2}$ for some C, λ. Suppose also that v, v', f, f' satisfy*

$$v, v' \in C^{\alpha/2,\, 1+\alpha}_{\mathrm{loc}}\big((0, T) \times \mathbb{R}^d\big),$$
$$f, f' \in C^{\alpha/2,\, \alpha}_{\mathrm{loc}}\big((0, T) \times \mathbb{R}^d\big),$$

and the growth conditions, for some fixed K_1, K_2 (independent of t),

$$\|v_t(x)\| + \|v_t'(x)\| \le K_1(1 + \|x\|), \qquad (t,x) \in (0,T) \times \mathbb{R}^d,$$
$$|f_t(x)| + |f_t'(x)| \le K_2(1 + \|x\|^2), \qquad (t,x) \in (0,T) \times \mathbb{R}^d.$$

Then we have the following:

(i) There exist unique classical solutions $\pi, \pi' \in C^{1,2}\big((0,T) \times \mathbb{R}^d\big) \cap C\big([0,T) \times \mathbb{R}^d\big)$ to

$$\begin{cases} \partial_t \pi_t = -\nabla \cdot (v_t\, \pi_t) + \Delta \pi_t + f_t\, \pi_t, & \pi|_{t=0} = \pi_0, \\ \partial_t \pi_t' = -\nabla \cdot (v_t'\, \pi_t') + \Delta \pi_t' + f_t'\, \pi_t', & \pi'|_{t=0} = \pi_0', \end{cases} \tag{20}$$

and $\pi_t, \pi_t' > 0$ for all $t \in (0,T)$.

(ii) The ratio $g_t(x) := \pi_t'(x)/\pi_t(x)$ belongs to $C^{1,2}\big((0,T) \times \mathbb{R}^d\big) \cap C\big([0,T) \times \mathbb{R}^d\big)$ and is a classical solution to

$$\partial_t g_t = \Delta g_t + b_t \cdot \nabla g_t - c_t\, g_t, \qquad g|_{t=0} = g_0, \tag{21}$$

with

$$b_t := 2\nabla \log \pi_t - v_t', \qquad c_t := f_t - f_t' - \nabla \cdot \big(v_t - v_t'\big) - \big(v_t - v_t'\big) \cdot \nabla \log \pi_t. \tag{22}$$

(iii) Fix $t \in (0,T)$ and suppose in addition that

(a) b has linear growth, $\|b_r(x)\| \le K_3(1 + \|x\|)$ for $(r,x) \in (0,t] \times \mathbb{R}^d$, and is locally Lipschitz in x uniformly in $r \in (0,t]$;

(b) c is bounded below, $c_r(x) \ge -K_4$ for $(r,x) \in (0,t] \times \mathbb{R}^d$;

(c) $|g_r(x)| \le C' e^{\lambda' \|x\|^2}$ for $(r,x) \in [0,t] \times \mathbb{R}^d$, where $\lambda' > 0$ is small enough that $\mathbb{E}\left[\sup_{s \le t} e^{\lambda' \|A_s\|^2} \,\middle|\, A_0 = x\right] < \infty$ for the process A below.

Then the SDE

$$dA_s = b_{t-s}(A_s)\, ds + \sqrt{2}\, dB_s, \qquad A_0 = x, \qquad s \in [0,t], \tag{23}$$

where B is a standard d-dimensional Brownian motion, has a unique strong solution, and the ratio admits the Feynman–Kac representation

$$g_t(x) = \mathbb{E}\left[g_0(A_t)\, \exp\left(-\int_0^t c_{t-s}(A_s)\, ds \right) \,\middle|\, A_0 = x \right]. \tag{24}$$

In particular, g is the unique classical solution of (21) satisfying (c).

Condition (c) holds for all sufficiently small λ' (depending on t and K_3) under (a), since then $\sup_{s \le t} \|A_s\| \le e^{K_3 t}\big(\|x\| + K_3 t + \sqrt{2} \sup_{s \le t} \|B_s\|\big)$ by Grönwall's inequality and $\sup_{s \le t} \|B_s\|$ has Gaussian tails.

Remark 2 (The two uses in the main text)**.** In Theorem 1(ii) we take $\pi = \overleftarrow{\mu}^{\mathrm{DPS}}$, $\pi' = \overleftarrow{\nu}^{\mathrm{DPS}}$, with common transport $v_t = v_t' = x + 2\nabla \log \overrightarrow{\mu}^{\mathrm{DPS}}_{T-t} = x + 2\nabla \log \pi_t$, reaction $f_t = -c_{\mathrm{DPS}}(T-t,\cdot)$ and $f_t' = 0$. Then (22) gives

$$b_t(x) = -x, \qquad c_t(x) = -c_{\mathrm{DPS}}(T-t,x),$$

so A in (23) is the OU process (1), and (24) becomes

$$\frac{\overleftarrow{\nu}^{\mathrm{DPS}}_t}{\overleftarrow{\mu}^{\mathrm{DPS}}_t}(x) = \mathbb{E}_{X \sim \mathrm{OU}}\left[\frac{\gamma(X_t)}{\overrightarrow{\mu}^{\mathrm{DPS}}_T(X_t)} \exp\left(\int_0^t c_{\mathrm{DPS}}(T-t+s, X_s)\, ds \right) \,\middle|\, X_0 = x \right],$$

which is Lemma 6. In the proof of Theorem 2, on $[t_{\mathrm{stop}}, T)$ we take $\pi = \overleftarrow{\rho}$ and $\pi' = \overleftarrow{\nu}$ with common transport and no reaction, so again $b_t(x) = -x$ and $c_t \equiv 0$. In both cases b is time-independent, so the time-reversal b_{t-s} in (23) has no effect there; it matters only for general v'.

Proof of Lemma 2. That the assumptions on the coefficients and measures π'_0, π_0 imply (i) is a classical result (Ladyženskaja et al., 1968, Ch. IV, Thm. 5.1).

(ii). By (i), $g_t = \pi'_t/\pi_t$ is well defined, positive, and inherits the regularity of π, π'. Substitute $\pi'_t = g_t\pi_t$ into the second equation of (20), using

$$\partial_t(g_t\pi_t) = \pi_t\,\partial_t g_t + g_t\,\partial_t\pi_t, \qquad \Delta(g_t\pi_t) = \pi_t\Delta g_t + 2\nabla g_t\cdot\nabla\pi_t + g_t\Delta\pi_t,$$
$$\nabla\cdot(v'_t g_t\pi_t) = \pi_t\, v'_t\cdot\nabla g_t + g_t\,\nabla\cdot(v'_t\pi_t).$$

Subtracting g_t times the first equation of (20) gives

$$\pi_t\,\partial_t g_t = \pi_t\Delta g_t + \big(2\nabla\pi_t - v'_t\pi_t\big)\cdot\nabla g_t + \Big(\nabla\cdot\big((v_t - v'_t)\pi_t\big) + (f'_t - f_t)\,\pi_t\Big)\,g_t.$$

Dividing by $\pi_t > 0$ and expanding $\frac{1}{\pi_t}\nabla\cdot\big((v_t - v'_t)\pi_t\big) = \nabla\cdot(v_t - v'_t) + (v_t - v'_t)\cdot\nabla\log\pi_t$ yields (21) with b_t, c_t as in (22).

(iii). By (a), the drift $s \mapsto b_{t-s}$ is locally Lipschitz with linear growth on $[0,t)\times\mathbb{R}^d$, so (23) has a unique, non-exploding strong solution (Karatzas & Shreve, 1991). For $\varepsilon \in (0,t)$ define, for $s \in [0, t-\varepsilon]$, the process

$$M_s := g_{t-s}(A_s)\,\exp\!\left(-\int_0^s c_{t-r}(A_r)\,dr\right),$$

which depends on the whole trajectory. Since $t - s \geq \varepsilon > 0$, g is $C^{1,2}$ along the path, and Itô's formula gives

$$dM_s = e^{-\int_0^s c_{t-r}(A_r)\,dr}\Big[\big(-\partial_t g_{t-s} + \Delta g_{t-s} + b_{t-s}\cdot\nabla g_{t-s} - c_{t-s}\,g_{t-s}\big)(A_s)\,ds + \sqrt{2}\,\nabla g_{t-s}(A_s)\cdot dB_s\Big].$$

The ds term vanishes by (21) evaluated at time $t - s$; this is why the characteristics carry the time-reversed drift b_{t-s}. Hence

$$dM_s = \sqrt{2}\,e^{-\int_0^s c_{t-r}(A_r)\,dr}\,\nabla g_{t-s}(A_s)\cdot dB_s,$$

so M is a local martingale on $[0, t-\varepsilon]$. By (b) and (c),

$$\sup_{s\leq t-\varepsilon}|M_s| \;\leq\; C' e^{K_4 t}\sup_{s\leq t} e^{\lambda'\|A_s\|^2},$$

which is integrable and independent of ε. A local martingale dominated by an integrable random variable is a true martingale, so $\mathbb{E}[M_{t-\varepsilon} \mid A_0 = x] = M_0 = g_t(x)$ (see (Karatzas & Shreve, 1991, Theorem 5.7.6) for the standard presentation). As $\varepsilon \downarrow 0$, $M_{t-\varepsilon} \to g_0(A_t)\exp\big(-\int_0^t c_{t-r}(A_r)\,dr\big)$ almost surely, by continuity of g on $[0,T)\times\mathbb{R}^d$ and of the paths of A. Dominated convergence, with the same bound, gives (24). The same argument applies to any classical solution of (21) satisfying (c), which gives uniqueness. □

C Bounds and Identities on the OU process

C.1 The effective sample backward path

The starting point for sampling from a prior in score-based generative models is the forward path $t \mapsto \rho_t$ that interpolates between the prior $\rho_0 = \rho_*$ and the standard Gaussian $\gamma = \mathcal{N}(0, I)$. This is obtained by solving the OU process $\rho_t = \mathrm{Law}(X_t)$, where $\{X_t\}_{t\geq 0}$ satisfies the SDE

$$\begin{cases} dX_t = -X_t\,dt + \sqrt{2}\,dB_t, \\ X_0 \sim \rho_*, \end{cases}$$

or, equivalently, the Fokker–Planck equation

$$\begin{cases} \partial_t\rho_t = \Delta\rho_t + \nabla\cdot(x\,\rho_t), \\ \rho_0 = \rho_*. \end{cases} \tag{25}$$

By the log-Sobolev inequality for γ and the relative-entropy decay along the OU flow, we have the exponentially decaying bound for $t > 1$,

$$\mathcal{H}(\rho_t\,|\,\gamma) \leq C\,e^{-t}, \tag{26}$$

where C is a universal constant independent of dimension.

To obtain approximate samples from ρ_*, we fix a time horizon T and reverse the path: set $\overleftarrow{\rho}_t = \rho_{T-t}$. The reverse path satisfies

$$\begin{cases} \partial_t \overleftarrow{\rho}_t = \Delta \overleftarrow{\rho}_t - 2\nabla \cdot (\nabla \log \rho_{T-t}\, \overleftarrow{\rho}_t) - \nabla \cdot (x\, \overleftarrow{\rho}_t), \\ \overleftarrow{\rho}_0 = \rho_T, \end{cases} \tag{27}$$

where we have used the diffusion-to-drift identity

$$\Delta \overleftarrow{\rho}_t = -\Delta \overleftarrow{\rho}_t + 2\nabla \cdot (\nabla \log \rho_{T-t}\, \overleftarrow{\rho}_t).$$

In practice, the initial condition is replaced by a standard Gaussian $\overleftarrow{\rho}^{\mathrm{eff}}_0 = \gamma$. The computationally intensive part of this strategy is obtaining a good approximation $s_\theta(x,t) \approx \nabla \log \rho_t(x)$ from samples; see Section 2. The effective samples are then obtained by approximating the solution of

$$\begin{cases} \partial_t \overleftarrow{\rho}^{\mathrm{eff}}_t = \Delta \overleftarrow{\rho}^{\mathrm{eff}}_t - 2\nabla \cdot (s_\theta(\cdot, T-t)\, \overleftarrow{\rho}^{\mathrm{eff}}_t) - \nabla \cdot (x\, \overleftarrow{\rho}^{\mathrm{eff}}_t), \\ \overleftarrow{\rho}^{\mathrm{eff}}_0 = \gamma, \end{cases} \tag{28}$$

which arises as the law of

$$\begin{cases} d\tilde{X}_t = \tilde{X}_t\, dt + 2\, s_\theta(\tilde{X}_t, T-t)\, dt + \sqrt{2}\, dB_t, \\ \tilde{X}_0 \sim \gamma. \end{cases}$$

Differentiating the relative entropy $\frac{d}{dt}\mathcal{H}(\overleftarrow{\rho}^{\mathrm{eff}}_t \mid \overleftarrow{\rho}_t)$ along the flow yields the relative-entropy bound between the effective samples $\tilde{X}_T \sim \overleftarrow{\rho}^{\mathrm{eff}}_T$ and the true distribution $X_0 \sim \rho_*$,

$$\mathcal{H}(\overleftarrow{\rho}^{\mathrm{eff}}_T \mid \rho_*) \;\le\; \underbrace{\mathcal{H}(\rho_T \mid \gamma)}_{\le Ce^{-T}} + \underbrace{\int_0^T \int_{\mathbb{R}^d} \|\nabla \log \rho_t(x) - s_\theta(x,t)\|^2\, \rho_t(x)\, dx\, dt}_{\text{Score-approximation error}}. \tag{29}$$

In what follows, we assume the score-approximation error is negligible, so the samples obtained by denoising are for practical purposes indistinguishable from ρ_*.

C.2 Tweedie's identities

A widely used heuristic is to estimate X_0 from a noisy observation X_t via the conditional expectation

$$\hat{x}_t(x) \;:=\; \mathbb{E}[X_0 \mid X_t = x], \tag{30}$$

which, under the negligible-score-approximation assumption of Section C.1, also equals $\mathbb{E}[\tilde{X}_T \mid \tilde{X}_{T-t} = x]$ for the effective backward process. Tweedie's formula expresses (30) in closed form via the score:

$$\hat{x}_t(x) \;=\; e^t\, x \;+\; (e^t - e^{-t})\, \nabla \log \rho_t(x). \tag{31}$$

Derivation. The OU semigroup admits the Gaussian kernel

$$\rho_{t|0}(x \mid x_0) \;=\; \frac{1}{(2\pi(1-e^{-2t}))^{d/2}} \exp\!\left(-\frac{\|x - e^{-t}x_0\|^2}{2(1-e^{-2t})}\right), \tag{32}$$

so $\nabla_x \log \rho_{t|0}(x \mid x_0) = -(x - e^{-t}x_0)/(1-e^{-2t})$. Differentiating $\rho_t(x) = \int \rho_{t|0}(x \mid x_0)\, \rho_*(x_0)\, dx_0$ in x and dividing by $\rho_t(x)$,

$$\nabla \log \rho_t(x) \;=\; -\frac{x - e^{-t}\, \hat{x}_t(x)}{1 - e^{-2t}}.$$

Solving for $\hat{x}_t(x)$ and using $e^t(1-e^{-2t}) = e^t - e^{-t}$ yields (31). □

Second-order identity. Differentiating (31) in x,

$$\nabla \hat{x}_t(x) \;=\; e^t\, I \;+\; (e^t - e^{-t})\, D^2 \log \rho_t(x). \tag{33}$$

The right-hand side admits a probabilistic interpretation as a rescaled conditional covariance:

$$\Sigma_t(x) \;:=\; \mathrm{Cov}(X_0 \mid X_t = x) \;=\; (e^t - e^{-t})\, \nabla \hat{x}_t(x). \tag{34}$$

Proof of (34). Differentiating ρ_t twice via (32) and subtracting $(\nabla \log \rho_t)(\nabla \log \rho_t)^\top$ to convert from $D^2\rho_t/\rho_t$ to $D^2 \log \rho_t$,

$$D^2 \log \rho_t(x) \;=\; -\frac{I}{1-e^{-2t}} \;+\; \frac{e^{-2t}}{(1-e^{-2t})^2}\,\Sigma_t(x). \tag{35}$$

Substituting into (33), the identity $(e^t - e^{-t})/(1-e^{-2t}) = e^t$ cancels the $e^t I$ contribution, and the identity $(e^t - e^{-t})\,e^{-2t}/(1-e^{-2t})^2 = e^{-t}/(1-e^{-2t}) = 1/(e^t - e^{-t})$ collapses the covariance term, giving

$$\nabla \hat{x}_t(x) \;=\; \frac{\Sigma_t(x)}{e^t - e^{-t}},$$

which rearranges to (34). □

C.2.1 The zero-noise limit

Throughout this subsection we assume ρ_* is supported on a smooth, compact, k-dimensional submanifold $\mathcal{M} \subset \mathbb{R}^d$ of positive reach $\tau_{\mathcal{M}} > 0$, with a smooth positive density with respect to the volume measure on $\mathcal{M}$. The orthogonal projection $P_{\mathcal{M}} : x \mapsto \arg\min_{x_0 \in \mathcal{M}} \|x - x_0\|$ is then well-defined and smooth on the tubular neighborhood $\mathcal{N}_{\tau_{\mathcal{M}}} := \{x \in \mathbb{R}^d : \mathrm{dist}(x, \mathcal{M}) < \tau_{\mathcal{M}}\}$, and the case $x \in \mathcal{M}$ corresponds to $\mathrm{dist}(x, \mathcal{M}) = 0$. We compute the small-$t$ behavior of $\hat{x}_t$ and $\nabla \hat{x}_t$ via Laplace's method on $\mathcal{M}$; standard references include Varadhan (1966); Dembo & Zeitouni (2010).

Setup. As $t \to 0^+$, $1 - e^{-2t} = 2t + O(t^2)$ and $e^{-t} = 1 + O(t)$, so the OU kernel (32) concentrates as

$$\rho_{t|0}(x \mid x_0) \;=\; \frac{1 + O(t)}{(4\pi t)^{d/2}} \exp\!\left(-\frac{\|x - x_0\|^2}{4t} + O(1)\right) \qquad \text{as } t \to 0^+.$$

The integrals defining $\hat{x}_t(x)$ are of Laplace form on $\mathcal{M}$ with phase $\Phi(x_0) = \|x - x_0\|^2$ at temperature $4t$.

Local geometry. Fix $x \in \mathcal{N}_{\tau_{\mathcal{M}}}$, set $p := P_{\mathcal{M}}(x)$, and let $P_{T_p\mathcal{M}}$ be the orthogonal projection onto the tangent space $T_p\mathcal{M}$. Parametrize $\mathcal{M}$ near p by tangent vectors,

$$\varphi(v) \;=\; p + v + O(\|v\|^2), \qquad v \in T_p\mathcal{M},$$

where the $O(\|v\|^2)$ correction lies in the normal space $N_p\mathcal{M} := (T_p\mathcal{M})^\perp$ and is bounded by the second fundamental form. Since $x - p \in N_p\mathcal{M}$ is orthogonal to $v \in T_p\mathcal{M}$,

$$\|x - \varphi(v)\|^2 \;=\; \|x - p\|^2 + \|v\|^2 + O(\|v\|^3), \tag{36}$$

so the phase $\Phi(x_0) = \|x - x_0\|^2$ is locally quadratic in tangent coordinates with Hessian $2\,I_{T_p\mathcal{M}}$ at the minimizer p.

Limit of $\hat{x}_t$. On $\mathcal{N}_{\tau_{\mathcal{M}}}$, the phase $\Phi|_{\mathcal{M}}$ has unique global minimizer p with $\Phi(p) = \|x - p\|^2$. Laplace's method on $\mathcal{M}$, applied with the local expansion (36), gives the asymptotic

$$\int_{\mathcal{M}} f(x_0)\, e^{-\Phi(x_0)/(4t)}\, \rho_*(x_0)\, d\mathrm{vol}(x_0) \;=\; (4\pi t)^{k/2}\, e^{-\Phi(p)/(4t)}\, \rho_*(p)\, \big(f(p) + O(t)\big)$$

for any C^1 function f on $\mathcal{M}$. Recall that the Tweedie estimate is the ratio

$$\hat{x}_t(x) \;=\; \frac{\int_{\mathcal{M}} x_0\, e^{-\Phi(x_0)/(4t)}\, \rho_*(x_0)\, d\mathrm{vol}(x_0)}{\int_{\mathcal{M}} e^{-\Phi(x_0)/(4t)}\, \rho_*(x_0)\, d\mathrm{vol}(x_0)}.$$

Applying the asymptotic to numerator and denominator, the common prefactor $(4\pi t)^{k/2}\, e^{-\Phi(p)/(4t)}\, \rho_*(p)$ cancels, leaving $\hat{x}_t(x) = p + O(t)$, where the $O(t)$ remainder collects the next-order Laplace corrections. The $O(\sqrt{t})$ contributions from tangential fluctuations vanish by Gaussian symmetry on $T_p\mathcal{M}$, since the linear function $v \mapsto v$ has zero mean under a centered Gaussian. Hence

$$\hat{x}_t(x) \;=\; P_{\mathcal{M}}(x) + O(t), \qquad \lim_{t \to 0^+} \hat{x}_t(x) \;=\; P_{\mathcal{M}}(x). \tag{37}$$

Limit of $\nabla\hat{x}_t$. By (34) and $e^t - e^{-t} = 2t + O(t^3)$,

$$\nabla\hat{x}_t(x) \;=\; \frac{\Sigma_t(x)}{e^t - e^{-t}} \;=\; \frac{\Sigma_t(x)}{2t}\big(1 + O(t^2)\big), \tag{38}$$

so it suffices to compute $\Sigma_t(x)$ to leading order. By (36), the conditional law $\rho_{0|t}$ is asymptotically Gaussian on $T_p\mathcal{M}$ with covariance $2t\, I_{T_p\mathcal{M}}$, and the normal component of $X_0 - p$ is of order $\|v\|^2 = O(t)$ and contributes only at order t^2. Hence

$$\Sigma_t(x) \;=\; 2t\, P_{T_p\mathcal{M}} + O(t^2), \tag{39}$$

and substituting into (38) yields

$$\lim_{t\to 0^+} \nabla\hat{x}_t(x) \;=\; P_{T_{P_{\mathcal{M}}(x)}\mathcal{M}}. \tag{40}$$

Consequence. In the small-noise regime relevant near the constraint set, $\hat{x}_t$ acts as the orthogonal projection onto the data manifold and $\nabla\hat{x}_t$ acts as the projection onto the corresponding tangent space. This is the geometric structure that drives the manifold-tangent oscillations of the DPS guidance analyzed in Section 5.

D THE DPS ALGORITHM (CHUNG ET AL., 2023)

Algorithm 1 DPS

Require: N, $\mathbf{y}$, $\{\zeta_i\}_{i=1}^N$, $\{\tilde{\sigma}_i\}_{i=1}^N$
1: $\mathbf{x}_N \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
2: **for** $i = N-1$ **to** 0 **do**
3: $\quad\hat{\mathbf{s}} \leftarrow \mathbf{s}_\theta(\mathbf{x}_i, i)$
4: $\quad\hat{\mathbf{x}}_0 \leftarrow \frac{1}{\sqrt{\bar{\alpha}_i}}(\mathbf{x}_i + (1-\bar{\alpha}_i)\hat{\mathbf{s}})$
5: $\quad\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
6: $\quad\mathbf{x}'_{i-1} \leftarrow \frac{\sqrt{\alpha_i}(1-\bar{\alpha}_{i-1})}{1-\bar{\alpha}_i}\mathbf{x}_i + \frac{\sqrt{\bar{\alpha}_{i-1}}\beta_i}{1-\bar{\alpha}_i}\hat{\mathbf{x}}_0 + \tilde{\sigma}_i\mathbf{z}$
7: $\quad\mathbf{x}_{i-1} \leftarrow \mathbf{x}'_{i-1} - \zeta_i\nabla_{\mathbf{x}_i}\|\mathbf{y} - \mathcal{A}(\hat{\mathbf{x}}_0)\|_2^2$
8: **end for**
9: **return** $\hat{\mathbf{x}}_0$

Here $\hat{\mathbf{x}}_0$ is the Tweedie mean $\hat{x}_{t_i}(\mathbf{x}_i)$ of the iterate at step i, $\mathcal{A} : \mathbb{R}^d \to \mathbb{R}^L$ is a general linear or non-linear observation operator and $y \in \mathbb{R}^L$ is the observation. The guidance schedule $\{\zeta_i\}_{i=1}^N$ is taken to be trajectory-dependent,

$$\zeta_i = \frac{\alpha}{\|y - \mathcal{A}(\hat{x}_{t_i}(x_i))\|_2}, \tag{41}$$

where $\alpha \in [0.2, 1]$ is a hyperparameter chosen depending on the inverse problem to be solved. In practice, the choice of guidance schedule significantly affects the performance of the algorithm.

E PROOF OF THEOREM 1

In this section, we will prove the following theorem

Theorem 1. *The terminal law $\nu_y^{\mathrm{DPS}} := \overleftarrow{\nu}_T^{\mathrm{DPS}}$ of the DPS-SDE (DPS SDE) differs from the true posterior μ_y by a pointwise multiplicative weight:*

$$\mu_y(x) = \omega(x)\, \nu_y^{\mathrm{DPS}}(x). \tag{9}$$

The weight ω admits two equivalent Feynman–Kac representations in terms of c_{DPS} defined in (8):

(i) ***Backward path*** *(condition on the DPS denoising process arriving at $Y_T = x$):*

$$\omega(x) = \mathbb{E}_{Y\sim(\text{DPS SDE})}\left[\frac{\overrightarrow{\mu}_T^{\mathrm{DPS}}(Y_0)}{\gamma(Y_0)}\,\exp\Big(-\int_0^T c_{\mathrm{DPS}}(T-s, Y_s)\,ds\Big)\;\Big|\; Y_T = x\right]. \tag{10}$$

(ii) ***Forward path*** *(condition on the OU process (1) starting at $X_0 = x$)*:

$$\frac{1}{\omega(x)} = \mathbb{E}_{X\sim \mathrm{OU}}\left[\frac{\gamma(X_T)}{\overrightarrow{\mu}_T^{\mathrm{DPS}}(X_T)}\,\exp\Big(\int_0^T c_{\mathrm{DPS}}(s, X_s)\,ds\Big)\;\Big|\; X_0 = x\right]. \tag{11}$$

Both path functionals are expressible in terms of quantities obtainable from the score oracle and its Jacobian via Tweedie's formula. Importance-weighting DPS samples by ω recovers μ_y exactly.

Sketch. The proof has three steps (elaborated below).

Step 1. We record an evolution equation satisfied by any tilted prior path $t \mapsto \overrightarrow{\mu}_t := h_t\rho_t/Z_t$.

$$\partial_t \overrightarrow{\mu}_t = \Delta \overrightarrow{\mu}_t + \nabla\cdot(x\overrightarrow{\mu}_t) + c[h_t, \rho_t](x)\,\overrightarrow{\mu}_t,$$

for an appropriate choice of $c[h_t, \rho_t]$ (Lemma 3).

Step 2. Specializing to the DPS tilt $h_t = e^{R_y\circ\hat{x}_t}$ and using the Kolmogorov backward equation for the conditional mean $\hat{x}_t(x) = \mathbb{E}[X_0 \mid X_t = x]$, we identify the reaction term as exactly c_{DPS} from (8) (Lemma 4).

Step 3. In Lemma 5, we use Lemmas 3 and 4 and apply the Feynman–Kac formula Lemma 2 to the resulting evolution equation to obtain an expression for the weights $\frac{\mu_y}{\nu_y^{\mathrm{DPS}}}$. Reversing the path integral, in Lemma 6 we recast the path expectation along the DPS reverse SDE (DPS SDE) with the forward OU path measure. □

E.1 Two Tweedie identities and the Kolmogorov backward equation

We collect three facts used repeatedly. Throughout, ρ_t denotes the marginal of the OU forward process (1) starting at $X_0 \sim \rho_*$, and $\hat{x}_t, \Sigma_t$ are the conditional mean and covariance of X_0 given X_t.

(F1) First-order Tweedie. A direct integration of the OU semigroup yields

$$\hat{x}_t(x) = e^t x + (e^t - e^{-t})\nabla \log \rho_t(x). \tag{42}$$

(F2) Second-order Tweedie. Differentiating (42) in x and using the standard identity $\Sigma_t(x) = (e^t - e^{-t})^2 D^2 \log \rho_t(x) + (e^t - e^{-t})e^t I$ (which follows from a second-order expansion of the OU posterior, or equivalently from differentiating Tweedie under Bayes' rule):

$$\nabla \hat{x}_t(x) = \frac{1}{e^t - e^{-t}}\,\Sigma_t(x). \tag{43}$$

In particular, $\nabla\hat{x}_t$ is symmetric and positive semidefinite.

(F3) Kolmogorov backward equation for $\hat{x}_t$. By Anderson's reversal (Anderson, 1982), the time-reversed OU process $\overleftarrow{X}_s := X_{T-s}$ satisfies the SDE

$$d\overleftarrow{X}_s = (\overleftarrow{X}_s + 2\nabla \log \rho_{T-s}(\overleftarrow{X}_s))\,ds + \sqrt{2}\,d\overleftarrow{B}_s, \tag{44}$$

with generator acting on $f \in C_C^\infty$,

$$\overleftarrow{L}_t^{OU} f := \Delta f + (x + 2\nabla \log \rho_{T-t}(x))\cdot\nabla f \tag{45}$$

Since

$$\hat{x}_t(x) = \mathbb{E}[X_0 \mid X_t = x] = \mathbb{E}[\overleftarrow{X}_T \mid \overleftarrow{X}_{T-t} = x],$$

$\hat{x}_t$ is a Kolmogorov backward solution along $\overleftarrow{X}$, and hence

$$\partial_t \hat{x}_t(x) = \Delta\hat{x}_t(x) + \nabla\hat{x}_t(x)\cdot\big(x + 2\nabla \log \rho_t(x)\big). \tag{46}$$

E.2 Step 1: Evolution of tilted prior paths

Lemma 3 (Tilted-prior Fokker–Planck)**.** *Let $h \in C^{1,2}([0,T]\times\mathbb{R}^d)$ be positive with $h_t \in L^1(\rho_t\,dx)$, and define the tilted prior path*

$$\overrightarrow{\mu}_t(x) \coloneqq \frac{h_t(x)\rho_t(x)}{Z_t}, \qquad Z_t \coloneqq \int h_t(x)\rho_t(x)\,dx.$$

Then $\overrightarrow{\mu}_t$ obeys

$$\partial_t \overrightarrow{\mu}_t = L^\dagger \overrightarrow{\mu}_t + c[h_t,\rho_t](x)\,\overrightarrow{\mu}_t, \tag{47}$$

where $L^\dagger f \coloneqq \Delta f + \nabla\cdot(xf)$ is the OU Fokker–Planck operator and the reaction term is

$$c[h_t,\rho_t](x) \coloneqq \partial_t \log h_t - \big(2\nabla\log\rho_t + x\big)\cdot\nabla\log h_t - \|\nabla\log h_t\|^2 - \Delta\log h_t - \tfrac{d}{dt}\log Z_t. \tag{48}$$

Proof. Using the product rule and writing Z_t terms as log-derivatives:

$$\partial_t \overrightarrow{\mu}_t = \frac{(\partial_t h_t)\rho_t}{Z_t} + \frac{h_t(\partial_t\rho_t)}{Z_t} - \Big(\tfrac{d}{dt}\log Z_t\Big)\overrightarrow{\mu}_t.$$

Writing $\partial_t h_t = h_t\,\partial_t\log h_t$ and substituting the OU Fokker–Planck equation $\partial_t\rho_t = L^\dagger\rho_t$:

$$\partial_t \overrightarrow{\mu}_t = (\partial_t\log h_t)\,\overrightarrow{\mu}_t + \frac{h_t}{Z_t}L^\dagger\rho_t - (\tfrac{d}{dt}\log Z_t)\,\overrightarrow{\mu}_t.$$

The OU Fokker–Planck operator is $L^\dagger\rho = \Delta\rho + \nabla\cdot(x\rho)$. Using $\Delta\rho_t = \rho_t(\|\nabla\log\rho_t\|^2 + \Delta\log\rho_t)$ and $\nabla\cdot(x\rho_t) = \rho_t(d + x\cdot\nabla\log\rho_t)$:

$$h_t L^\dagger\rho_t = h_t\rho_t\Big[\|\nabla\log\rho_t\|^2 + \Delta\log\rho_t + d + x\cdot\nabla\log\rho_t\Big].$$

Set $\varphi = h_t/Z_t$ so that $\overrightarrow{\mu}_t = \varphi\rho_t$. We compute $L^\dagger\overrightarrow{\mu}_t = L^\dagger(\varphi\rho_t)$ via the product rule applied to each term of $L^\dagger = \Delta + \nabla\cdot(x\,\cdot)$:

$$\begin{aligned}\Delta(\varphi\rho_t) &= \varphi\Delta\rho_t + 2\nabla\varphi\cdot\nabla\rho_t + \rho_t\Delta\varphi,\\ \nabla\cdot(x\varphi\rho_t) &= \varphi\,\nabla\cdot(x\rho_t) + \rho_t\,x\cdot\nabla\varphi.\end{aligned}$$

Summing, we have

$$L^\dagger(\varphi\rho_t) = \varphi\,L^\dagger\rho_t + 2\nabla\varphi\cdot\nabla\rho_t + \rho_t\Delta\varphi + \rho_t\,x\cdot\nabla\varphi.$$

Rearranging to isolate $\varphi L^\dagger\rho_t$:

$$\varphi\,L^\dagger\rho_t = L^\dagger(\varphi\rho_t) - 2\nabla\varphi\cdot\nabla\rho_t - \rho_t(\Delta\varphi + x\cdot\nabla\varphi)\,.$$

Since Z_t does not depend on x, we have $\nabla\varphi = \nabla h_t/Z_t$ and $\Delta\varphi = \Delta h_t/Z_t$. Substituting $\varphi\rho_t = \overrightarrow{\mu}_t$, $\nabla\rho_t = \rho_t\nabla\log\rho_t$, and $\rho_t/Z_t = \overrightarrow{\mu}_t/h_t$, and using the identities $\nabla h_t/h_t = \nabla\log h_t$ and $\Delta h_t/h_t = \|\nabla\log h_t\|^2 + \Delta\log h_t$:

$$\frac{h_t}{Z_t}L^\dagger\rho_t = L^\dagger\overrightarrow{\mu}_t - \overrightarrow{\mu}_t\Big[2\nabla\log\rho_t\cdot\nabla\log h_t + \|\nabla\log h_t\|^2 + \Delta\log h_t + x\cdot\nabla\log h_t\Big].$$

Collecting all $\overrightarrow{\mu}_t$ terms:

$$\partial_t\overrightarrow{\mu}_t = L^\dagger\overrightarrow{\mu}_t + c[h_t,\rho_t](x)\,\overrightarrow{\mu}_t,$$

where the reaction coefficient is

$$c[h_t,\rho_t](x) = \partial_t\log h_t - (2\nabla\log\rho_t + x)\cdot\nabla\log h_t - \|\nabla\log h_t\|^2 - \Delta\log h_t - \tfrac{d}{dt}\log Z_t$$

□

E.3 Step 2: Specialization to the DPS tilt

Lemma 4 (Reaction term for the DPS tilt). *With $h_t(x) = \exp(R_y(\hat{x}_t(x)))$, so that $\overrightarrow{\mu}_t = \overrightarrow{\mu}_t^{\mathrm{DPS}}$ in (DPS Surrogate Path), the reaction $c[h_t, \rho_t]$ from (48) reduces to*

$$c[h_t, \rho_t](x) = -\frac{1}{(e^t - e^{-t})^2}\Big[\operatorname{tr}\left(\Sigma_t(x)\, D^2 R_y(\hat{x}_t(x))\, \Sigma_t(x)\right) + \|\Sigma_t(x) \nabla R_y(\hat{x}_t(x))\|^2\Big] - \tfrac{d}{dt}\log Z_t, \tag{49}$$

that is, $c[h_t, \rho_t](x) = c_{\mathrm{DPS}}(t, x)$ as defined in (8).

Proof. Set $\log h_t = R_y \circ \hat{x}_t$ and apply the chain rule component-wise:

$$\begin{aligned}
\partial_t \log h_t &= \nabla R_y(\hat{x}_t) \cdot \partial_t \hat{x}_t, \\
\nabla \log h_t &= (\nabla \hat{x}_t)\, \nabla R_y(\hat{x}_t), \\
\Delta \log h_t &= \nabla R_y(\hat{x}_t) \cdot \Delta \hat{x}_t + \operatorname{tr}\big(\nabla \hat{x}_t\, D^2 R_y(\hat{x}_t)\, \nabla \hat{x}_t\big),
\end{aligned}$$

using symmetry of $\nabla \hat{x}_t$ from (43). Substituting into (48) and grouping terms by their dependence on ∇R_y and $D^2 R_y$, we get the following expression for $c[h_t, \rho_t]$:

$$\begin{aligned}
c[h_t, \rho_t](x) + \tfrac{d}{dt}\log Z_t = \nabla R_y(\hat{x}_t) \cdot \partial_t \hat{x}_t - (2\nabla \log \rho_t + x) \cdot (\nabla \hat{x}_t)\, \nabla R_y(\hat{x}_t) \\
- \|(\nabla \hat{x}_t)\, \nabla R_y(\hat{x}_t)\|^2 - \nabla R_y(\hat{x}_t) \cdot \Delta \hat{x}_t - \operatorname{tr}\big(\nabla \hat{x}_t\, D^2 R_y(\hat{x}_t)\, \nabla \hat{x}_t\big)
\end{aligned}$$

Grouping the terms of ∇R together, we have

$$\begin{aligned}
c[h_t, \rho_t](x) + \tfrac{d}{dt}\log Z_t = \nabla R_y(\hat{x}_t) \cdot [\partial_t \hat{x}_t - \Delta \hat{x}_t - (2\nabla \log \rho_t + x) \cdot (\nabla \hat{x}_t)] \\
- \|(\nabla \hat{x}_t)\, \nabla R_y(\hat{x}_t)\|^2 - \operatorname{tr}\big(\nabla \hat{x}_t\, D^2 R_y(\hat{x}_t)\, \nabla \hat{x}_t\big)
\end{aligned}$$

By the Kolmogorov backward equation (46), we have $\partial_t \hat{x}_t - \nabla \hat{x}_t\,(x + 2\nabla \log \rho_t) - \Delta \hat{x}_t = 0$. This is the key cancellation underlying the bias formula. The remaining contributions are $-\|\nabla \log h_t\|^2$ and the trace piece of $-\Delta \log h_t$:

$$-\|(\nabla \hat{x}_t) \nabla R_y(\hat{x}_t)\|^2 - \operatorname{tr}\big(\nabla \hat{x}_t\, D^2 R_y(\hat{x}_t)\, \nabla \hat{x}_t\big).$$

Using (43) to substitute $\nabla \hat{x}_t = \Sigma_t/(e^t - e^{-t})$ and pulling out the common scalar factor yields (49). □

E.4 Step 3: Feynman–Kac two ways

We now combine Lemmas 3 and 4 to prove Theorem 1: part (i) in Lemma 5 and part (ii) in Lemma 6.

Lemma 5. *The terminal law $\nu_y^{\mathrm{DPS}} := \overleftarrow{\nu}_T^{\mathrm{DPS}}$ of the DPS-SDE (DPS SDE) differs from the true posterior μ_y by a pointwise multiplicative weight:*

$$\mu_y(x) = \omega(x)\, \nu_y^{\mathrm{DPS}}(x).$$

The weight ω admits a Feynman–Kac representation in terms of the reaction term c_{DPS} defined in (8):

$$\omega(x) = \mathbb{E}_{Y \sim (\text{DPS SDE})}\left[\frac{\overrightarrow{\mu}_T^{\mathrm{DPS}}(Y_0)}{\gamma(Y_0)}\, \exp\Big(-\int_0^T c_{\mathrm{DPS}}(T - s, Y_s)\, ds\Big) \;\Big|\; Y_T = x\right].$$

Proof. Writing $\overrightarrow{\mu}_t$ for $\overrightarrow{\mu}_t^{\mathrm{DPS}}$ in this proof, setting $\overleftarrow{\mu}_t := \overrightarrow{\mu}_{T-t}$, and applying (47) together with the diffusion-to-drift identity

$$-L^\dagger \overleftarrow{\mu}_t = \Delta \overleftarrow{\mu}_t - \nabla \cdot \big((x + 2\nabla \log \overleftarrow{\mu}_t)\, \overleftarrow{\mu}_t\big)$$

gives exactly

$$\begin{cases}
\partial_t \overleftarrow{\mu}_t = \Delta \overleftarrow{\mu}_t - 2\,\nabla \cdot \big(\nabla \log \overrightarrow{\mu}_{T-t}\, \overleftarrow{\mu}_t\big) - \nabla \cdot (x\, \overleftarrow{\mu}_t) - c_{\mathrm{DPS}}(T - t, x)\, \overleftarrow{\mu}_t, \\
\overleftarrow{\mu}_0(x) = \overrightarrow{\mu}_T(x),
\end{cases} \tag{50}$$

with reaction $-c_{\mathrm{DPS}}$ by Lemma 4, which is (DPS Surrogate PDE). By construction, $\overleftarrow{\mu}_T = \overrightarrow{\mu}_0 = \mu_y$. Equation (50) is the Fokker–Planck equation associated with the DPS reverse SDE

$$\begin{cases} dY_t = \left(Y_t + 2\nabla \log \rho_{T-t}(Y_t) + \frac{2}{e^{T-t} - e^{-(T-t)}} \Sigma_{T-t}(Y_t)\, \nabla R_y(\hat{x}_{T-t}(Y_t))\right) dt + \sqrt{2}\, dB_t, \\ Y_0 \sim \gamma, \end{cases} \tag{51}$$

augmented by a multiplicative reaction $-c_{\mathrm{DPS}}$. The DPS algorithm itself directly simulates (51) from $Y_0 \sim \gamma$, that is, *without* the reaction term. This produces a marginal $\overleftarrow{\nu}_t$ with $\overleftarrow{\nu}_T = \nu_y^{\mathrm{DPS}}$. Effectively, the DPS algorithm solves

$$\begin{cases} \partial_t \overleftarrow{\nu}_t = \Delta \overleftarrow{\nu}_t - 2\nabla \cdot (\nabla \log \overrightarrow{\mu}_{T-t}\, \overleftarrow{\nu}_t) - \nabla \cdot \left(x\, \overleftarrow{\nu}_t\right), \\ \overleftarrow{\nu}_0(x) = \gamma(x), \end{cases} \tag{52}$$

Two operators (50) and (52) differ only by a multiplicative reaction and an initial condition. They can be related by a Feynman–Kac formula of Lemma 2. The ground truth satisfies $\mu_y(x) = \overleftarrow{\mu}_T(x)$, for any test function $\varphi(x)$ we have

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x) \mu_y(x)\, dx &= \int_{\mathbb{R}^d} \varphi(x) \overleftarrow{\mu}_T(x)\, dx \\ &= \mathbb{E}_{Y \sim (51)} \left[\varphi(Y_T) \frac{\overrightarrow{\mu}_T(Y_0)}{\gamma(Y_0)}\, e^{-\int_0^T c_{\mathrm{DPS}}(T-s, Y_s)\, ds} \right] \\ &= \mathbb{E}_{Y \sim (51)} \left[\varphi(Y_T)\, \mathbb{E} \underbrace{\left[\frac{\overrightarrow{\mu}_T(Y_0)}{\gamma(Y_0)}\, e^{-\int_0^T c_{\mathrm{DPS}}(T-s, Y_s)\, ds} | Y_T \right]}_{\omega(Y_T)} \right] \\ &= \int_{\mathbb{R}^d} \varphi(x)\, \omega(x)\, \nu_y^{\mathrm{DPS}}(x)\, dx, \end{aligned}$$

where we conditioned on the value of Y_T, used the law of total expectation, and the observation $Y_T \sim \overleftarrow{\nu}_T$. Concretely,

$$\frac{\mu_y(x)}{\nu_y^{\mathrm{DPS}}(x)} = \frac{\overleftarrow{\mu}_T(x)}{\overleftarrow{\nu}_T(x)} = \mathbb{E}_{Y \sim (51)} \left[\frac{\overrightarrow{\mu}_T(Y_0)}{\gamma(Y_0)}\, e^{-\int_0^T c_{\mathrm{DPS}}(T-s, Y_s)\, ds} \;\middle|\; Y_T = x \right]. \tag{53}$$

The boundary factor $\overrightarrow{\mu}_T / \gamma$ accounts for the mismatch between $\overleftarrow{\mu}_0 = \overrightarrow{\mu}_T$ and $\overleftarrow{\nu}_0 = \gamma$. □

The formula for the weight using OU, can now be derived using Anderson the time reversal of SDEs. For clarity and brevity, we instead provide a derivation using PDE satisfied by the ratio of the algorithmic path to the PDE for the ratio of the algorithmic path to the surrogate path. As in (Surrogate Path) and (Algorithm Path) we here define

$$\begin{aligned} \overleftarrow{\mu}_t &:= \overleftarrow{\mu}_t^{\mathrm{DPS}} = \frac{1}{Z_{T-t}} \exp(R_y \circ \hat{x}_{T-t}(x)) \rho_{T-t}(x) \\ \overleftarrow{\nu}_t &:= \overleftarrow{\nu}_t^{\mathrm{DPS}} = \mathrm{Law}(Y_t). \end{aligned}$$

and for $0 \le t < T$, as in Section B define the density ratio $\psi_t(x)$,

$$\psi_t \overleftarrow{\mu}_t = \overleftarrow{\nu}_t \tag{54}$$

noting $\lim_{t \to T} \psi_t = \psi_T = \frac{1}{\omega(x)}$ as in (11).

Lemma 6. *The ratio $\psi_t(x)$ solves the parabolic initial value problem,*

$$\begin{cases} \partial_t \psi_t = \Delta \psi_t - x \cdot \nabla \psi_t + c_{\mathrm{DPS}}(T-t, x)\, \psi_t \\ \psi_0(x) = \frac{\gamma(x)}{\overrightarrow{\mu}_T^{\mathrm{DPS}}(x)}, \end{cases} \tag{55}$$

and thus, by Feynman–Kac formula, with $t < T$,

$$\psi_t(x) = \mathbb{E}_{X \sim \mathrm{OU}} \left[\frac{\gamma(X_t)}{\overrightarrow{\mu}_T^{\mathrm{DPS}}(X_t)} \exp\left(\int_0^t c_{\mathrm{DPS}}(T-t+s, X_s)\, ds \right) \middle| X_0 = x \right] \tag{56}$$

Taking the limit $t \to T$ yields the statement of Theorem 1 with $\frac{1}{\omega(x)} = \psi_T(x)$ as the ratio.

Proof. The following calculations hold classically, since for $t < T$, both $\overleftarrow{\mu}_t(x)$ and $\overleftarrow{\nu}_t^{\mathrm{DPS}}(x)$ are smooth, positive densities. Note $\log \psi_t = \log \overleftarrow{\nu}_t - \log \overleftarrow{\mu}_t$, using (50) and (52) we have

$$\begin{aligned}\partial_t \log \overleftarrow{\nu}_t &= \Delta \log \overleftarrow{\nu}_t + \|\nabla \log \overleftarrow{\nu}_t\|^2 - d - 2\Delta \log \overleftarrow{\mu}_t \\ &\quad - x \cdot \nabla \log \overleftarrow{\nu}_t - 2\nabla \log \overleftarrow{\mu}_t \cdot \nabla \log \overleftarrow{\nu}_t \\ \partial_t \log \overleftarrow{\mu}_t &= \Delta \log \overleftarrow{\mu}_t + \|\nabla \log \overleftarrow{\mu}_t\|^2 - d - 2\Delta \log \overleftarrow{\mu}_t \\ &\quad - x \cdot \nabla \log \overleftarrow{\mu}_t - 2\|\nabla \log \overleftarrow{\mu}_t\|^2 - c_{\mathrm{DPS}}(T-t,\cdot)\end{aligned}$$

Taking the difference and completing the square shows

$$\begin{aligned}\partial_t \log \psi_t &= \partial_t \log \overleftarrow{\nu}_t - \partial_t \log \overleftarrow{\mu}_t \\ &= \Delta \log \psi_t + \|\nabla \log \psi_t\|^2 - x \cdot \nabla \log \psi_t + c_{\mathrm{DPS}}(T-t,\cdot)\end{aligned} \tag{57}$$

Applying the Cole–Hopf transformation ($\log \psi_t \overset{\exp(\cdot)}{\mapsto} \psi_t$) obtains (55). Finally, applying Feynman–Kac to an initial-value problem (rather than terminal-value) induces time-reversal of the reaction term $c_{\mathrm{DPS}}(T-t+s, X_s)$ in the path-functional. □

F Early Guidance Stopping, Proof of Theorem 2

Algorithm 2 DPS with Early Guidance Stopping

Require: i_{stop}, N, $\mathbf{y}$, $\{\zeta_i\}_{i=1}^N$, $\{\tilde{\sigma}_i\}_{i=1}^N$
$\mathbf{x}_N \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
for $i = N-1$ **to** 0 **do**
  $\hat{\mathbf{s}} \leftarrow \mathbf{s}_\theta(\mathbf{x}_i, i)$
  $\hat{\mathbf{x}}_0 \leftarrow \frac{1}{\sqrt{\bar{\alpha}_i}}\big(\mathbf{x}_i + (1-\bar{\alpha}_i)\hat{\mathbf{s}}\big)$
  $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
  $\mathbf{x}'_{i-1} \leftarrow \frac{\sqrt{\alpha_i}(1-\bar{\alpha}_{i-1})}{1-\bar{\alpha}_i}\mathbf{x}_i + \frac{\sqrt{\bar{\alpha}_{i-1}}\beta_i}{1-\bar{\alpha}_i}\hat{\mathbf{x}}_0 + \tilde{\sigma}_i \mathbf{z}$
  if $i > i_{\mathrm{stop}}$ **then**
    $\mathbf{x}_{i-1} \leftarrow \mathbf{x}'_{i-1} - \zeta_i \nabla_{\mathbf{x}_i} \|\mathbf{y} - \mathcal{A}(\hat{\mathbf{x}}_0)\|_2^2$
  else
    $\mathbf{x}_{i-1} \leftarrow \mathbf{x}'_{i-1}$
  end if
end for
return $\hat{\mathbf{x}}_0$

The output Algorithm 2, up to discretization error, is characterized in the following result.

Theorem 2 (Early Guidance Stopping)**.** *If guidance is stopped at time $t_{\mathrm{stop}} = T - t_*$, the output of the DPS algorithm, up to time-discretization error, is given by*

$$\nu_y^{\mathrm{DPS},t_*}(x) = \frac{\mathbb{E}_{X\sim \mathrm{OU}}\left[\frac{1}{\omega_{t_*}}(X_{t_*})\, e^{\alpha \eta_{t_*} R_y(\hat{x}_{t_*}(X_{t_*}))} \,\middle|\, X_0 = x\right]}{Z_{t_*}} \rho_*(x),$$

where

$$\frac{1}{\omega_{t_*}}(x) := \mathbb{E}_{X\sim \mathrm{OU}}\left[\frac{\gamma(X_{T-t_*})}{\overrightarrow{\mu}_T^{\eta}(X_{T-t_*})} \exp\left(\int_0^{T-t_*} c^*_{\mathrm{DPS}}(t_*+s, X_s)\, ds\right) \,\middle|\, X_0 = x\right], \tag{18}$$

with

$$\begin{aligned}c^*_{\mathrm{DPS}}(t,x) = &-\frac{\alpha\eta_t \,\mathrm{tr}\big(\Sigma_t(x)(D^2 R_y)(\hat{x}_t(x))\Sigma_t(x)\big) + \alpha^2\eta_t^2 \left\|\Sigma_t(x)\nabla R_y(\hat{x}_t(x))\right\|^2}{(e^t - e^{-t})^2} \\ &+ \alpha \frac{d\eta_t}{dt} R_y(\hat{x}_t(x)) - \frac{d}{dt}\log Z_t,\end{aligned}$$

that is, (8) with R_y replaced by $\alpha\eta_t R_y$, plus a term from the time dependence of η_t. Here α, R_y, the ***annealing schedule*** *η_t and Z_t are as in (14)–(16).*

Proof. We apply (Surrogate Path) to the annealed path (15):

$$t \mapsto \overrightarrow{\mu}_t := \overrightarrow{\mu}^{\eta}_t = \frac{e^{\alpha \eta_t R_y(\hat{x}_t(x))}\, \rho_t}{Z_t}.$$

By Lemma 3, the reverse equation reads

$$\begin{aligned}\partial_t \overleftarrow{\mu}_t = \Delta \overleftarrow{\mu}_t - \nabla \cdot (x\, \overleftarrow{\mu}_t) - 2\nabla \cdot (\nabla \log \rho_{T-t}\, \overleftarrow{\mu}_t) - 2\alpha \eta_{T-t}\, \nabla \cdot (\nabla (R_y \circ \hat{x}_{T-t})\, \overleftarrow{\mu}_t) \\ - c^*_{\mathrm{DPS}}(T-t, x)\, \overleftarrow{\mu}_t,\end{aligned}$$

where $R_y(x) = -\|\mathcal{A}(x) - y\|_2$ for the (possibly non-linear) observation operator $\mathcal{A}$, and c^*_{DPS} is the reaction of Theorem 2: Lemma 4 applied to the tilt $h_t = e^{\alpha \eta_t R_y \circ \hat{x}_t}$, i.e. with R_y replaced by $\alpha \eta_t R_y$, plus the term $\alpha \frac{d\eta_t}{dt}\, R_y \circ \hat{x}_t$ coming from $\partial_t \log h_t$.

The corresponding algorithmic SDE (Algorithm Path) with early stopping at time $t_{\mathrm{stop}} := T - t_*$ is

$$\partial_t \overleftarrow{\nu}_t = \begin{cases} \Delta \overleftarrow{\nu}_t - \nabla \cdot (x\, \overleftarrow{\nu}_t) - 2\nabla \cdot (\nabla \log \rho_{T-t}\, \overleftarrow{\nu}_t) - 2\alpha \eta_{T-t}\, \nabla \cdot (\nabla (R_y \circ \hat{x}_{T-t})\, \overleftarrow{\nu}_t), & t \in (0, t_{\mathrm{stop}}), \\ \Delta \overleftarrow{\nu}_t - \nabla \cdot (x\, \overleftarrow{\nu}_t) - 2\nabla \cdot (\nabla \log \rho_{T-t}\, \overleftarrow{\nu}_t), & t \in [t_{\mathrm{stop}}, T). \end{cases} \tag{58}$$

On $(0, t_{\mathrm{stop}})$, both $\overleftarrow{\mu}_t$ and $\overleftarrow{\nu}_t$ satisfy the same equation, so Theorem 1 applies directly and yields

$$\frac{\overleftarrow{\nu}_{t_{\mathrm{stop}}}(x)}{\overleftarrow{\mu}_{t_{\mathrm{stop}}}(x)} = \frac{1}{\omega_{t_*}}(x) = \mathbb{E}_{X \sim \mathrm{OU}}\left[\frac{\gamma(X_{T-t_*})}{\overrightarrow{\mu}_T(X_{T-t_*})} \exp\left(\int_0^{T-t_*} c^*_{\mathrm{DPS}}(t_* + s,\, X_s)\, ds \right) \,\middle|\, X_0 = x \right].$$

Substituting the explicit form of $\overleftarrow{\mu}_{t_{\mathrm{stop}}}$,

$$\overleftarrow{\nu}_{T-t_*}(x) = \frac{1}{\omega_{t_*}}(x)\, \overleftarrow{\mu}_{t_{\mathrm{stop}}}(x) = \frac{1}{\omega_{t_*}}(x)\, \frac{e^{\alpha \eta_{t_*} R_y(\hat{x}_{t_*}(x))}}{Z_{t_*}}\, \overleftarrow{\rho}_{T-t_*}(x). \tag{59}$$

On $[t_{\mathrm{stop}}, T)$, the SDE for $\overleftarrow{\nu}_t$ is the unbiased reverse OU equation, started from (59). Setting $s = T - t$ for the corresponding forward time, the Radon–Nikodym derivative of the initial condition with respect to the OU forward marginal ρ_{t_*} is

$$g_{t_*}(x) := \frac{d\overleftarrow{\nu}_{T-t_*}}{d\overleftarrow{\rho}_{T-t_*}}(x) = \frac{1}{\omega_{t_*}(x)}\, \frac{e^{\alpha \eta_{t_*} R_y(\hat{x}_{t_*}(x))}}{Z_{t_*}}.$$

Applying the Feynman–Kac identity for ratios (Lemma 2),

$$\frac{d\overleftarrow{\nu}_T}{d\rho_*}(x_0) = \frac{d\overleftarrow{\nu}_T}{d\overleftarrow{\rho}_T}(x_0) = \mathbb{E}_{X \sim \mathrm{OU}}\left[g_{t_*}(X_{t_*}) \,\middle|\, X_0 = x_0 \right], \tag{60}$$

where $\{X_s\}_{s \geq 0}$ is the OU process started from $X_0 \sim \rho_*$. Substituting the expression for g_{t_*} yields the claim. □

G TIME DISCRETIZATION OF DDPM

To establish the correspondence between the discrete variance schedule β_i used in Denoising Diffusion Probabilistic Models (DDPM) Ho et al. (2020) and the continuous time steps Δt_i of the underlying Ornstein–Uhlenbeck (OU) process, we compare their respective transition kernels.

The forward Markov jump process in DDPM defines the transition from step i to $i+1$ as:

$$q(x_{i+1}|x_i) = \mathcal{N}(x_{i+1}; \sqrt{1-\beta_i} x_i, \beta_i I) \tag{61}$$

The continuous-time forward (OU) SDE under consideration is given by:

$$dX_t = -X_t dt + \sqrt{2} dB_t \tag{62}$$

For a finite time increment Δt_i, the exact solution to this SDE yields the transition:

$$p(x_{t+\Delta t_i}|x_t) = \mathcal{N}(x_{t+\Delta t_i}; e^{-\Delta t_i} x_t, (1 - e^{-2\Delta t_i}) I) \tag{63}$$

For the discrete Markov chain to exactly discretize the continuous SDE, the coefficients of the mean and variance must be consistent across regimes:

$$e^{-\Delta t_i} = \sqrt{1-\beta_i} \qquad 1 - e^{-2\Delta t_i} = \beta_i. \tag{64}$$

Solving for Δt_i we obtain

$$\Delta t_i = -\frac{1}{2}\log(1-\beta_i) \tag{65}$$

The linear noise schedule of DDPM is given by

$$\beta_i = \beta_{min} + i\frac{\beta_{max}-\beta_{min}}{N} \approx 10^{-4} + 2i\,10^{-5} \qquad \text{for } i = 1,\dots,1000,$$

with the choices $\beta_{min} = 10^{-4}$, $\beta_{max} = 0.02$, and $N = 1000$ steps. Applying the first-order Taylor expansion $\log(1-\epsilon) \approx -\epsilon$, we obtain the approximately linear relationship between Δt_i and β_i:

$$\Delta t_i \approx \frac{1}{2}\beta_i \tag{66}$$

Next, we derive a relationship in time between the discrete steps t_i and the continuous time t by summing over the increments:

$$t_i = \sum_{j=1}^{i}\Delta t_j \approx \frac{1}{2}\sum_{j=1}^{i}\beta_j = \frac{1}{2}\sum_{j=1}^{i}\left(10^{-4} + 2j\,10^{-5}\right) = \frac{1}{2}\left(10^{-4}i + 2\cdot 10^{-5}\frac{i(i+1)}{2}\right). \tag{67}$$

Next, we solve for a function $i(t)$ that maps continuous time to discrete steps by inverting the quadratic relationship $\frac{1}{2}\left(10^{-4}i + 2\cdot 10^{-5}\frac{i(i+1)}{2}\right) = t$:

$$i(t) = \frac{\sqrt{121 + 8t\,10^5} - 11}{2} \tag{68}$$

This function $i(t)$ provides a mapping from continuous time t to the corresponding discrete step index i in the DDPM framework, allowing us to understand the time step behavior in the continuum limit. Substituting $i(t)$ into (66), we obtain

$$\begin{aligned}\Delta t(t) \approx \frac{1}{2}\beta_{i(t)} &= \frac{1}{2}\left(10^{-4} + 2\cdot 10^{-5}\,i(t)\right)\\ &= \frac{1}{2}\left(10^{-4} + 10^{-5}\left(\sqrt{121 + 8\cdot 10^5\,t} - 11\right)\right)\\ &= \frac{10^{-5}}{2}\left(\sqrt{121 + 8\cdot 10^5\,t} - 1\right).\end{aligned} \tag{69}$$

A naive large-t approximation $\Delta t(t) \sim \sqrt{2\cdot 10^{-5}\,t}$ would incorrectly vanish at $t = 0$. To preserve the nonzero constant floor at the origin, we drop the small -1 term (negligible compared to $\sqrt{121} = 11$) but keep the constant 121 inside the square root. Pulling the prefactor $\frac{10^{-5}}{2}$ inside the radical yields the compact form

$$\Delta t(t) \approx \sqrt{2\cdot 10^{-5}\,t + \Delta t_0^2}, \qquad \Delta t_0 = 5.5\cdot 10^{-5}. \tag{70}$$

H FORWARD EULER INSTABILITY

In terms of implementation, DPS Algorithm 1 integrates the different terms of (DPS SDE) in different ways. The denoising step, corresponding to the terms $\left(Y_t + 2\nabla \log \rho_{T-t}(Y_t)\right)dt + \sqrt{2}\,dB_t$, is integrated implicitly via DDPM in Step 6, avoiding numerical instabilities. The guidance term $\alpha\eta_{T-t}\,\frac{2}{e^{T-t}-e^{-(T-t)}}\,\Sigma_{T-t}(Y_t)\,\nabla R_y(\hat{x}_{T-t}(Y_t))$, however, is integrated explicitly via forward Euler in Step 7. The *annealing schedule* $\eta_t = 1/\Delta t(t)$ in (16) is an auxiliary quantity we introduce to compensate for the missing time step Δt in Step 7. In place of Δt, the algorithm multiplies by the path-dependent factor

$$\zeta_i = \frac{\alpha}{\|y - \mathcal{A}(\hat{x}_{t_i}(x_i))\|_2},$$

applied to the gradient of the squared residual $\|y - \mathcal{A}(\hat{x}_{t_i}(x_i))\|_2^2$ in Step 7. Since $\nabla\|\cdot\|_2^2 = 2\|\cdot\|_2\nabla\|\cdot\|_2$, one guidance step equals $2\alpha\,\nabla(R_y \circ \hat{x}_{t_i})(x_i)$ with the reward $R_y(x) = -\|\mathcal{A}(x) - y\|_2$ of (15). In our analysis, this is the guidance drift $2\alpha\eta_t\nabla(R_y \circ \hat{x}_t)$ of the annealed path (15) integrated over one step, since $\eta_t\,\Delta t = 1$ for the annealing schedule (16).

To derive the oscillations, we assume that the prior ρ_* is a smooth distribution supported on a smooth lower-dimensional manifold $\mathcal{M} \subset \mathbb{R}^d$ embedded in the ambient space. In the limit $t \to 0^+$,

$$\hat{x}_t(x) \to P_{\mathcal{M}}(x), \qquad \nabla\hat{x}_t(x) \to P_{T_{P_{\mathcal{M}}(x)}\mathcal{M}},$$

where $P_{\mathcal{M}}$ denotes the orthogonal projection onto $\mathcal{M}$ and $P_{T_{P_{\mathcal{M}}(x)}\mathcal{M}}$ denotes the orthogonal projection onto the tangent space at $P_{\mathcal{M}}(x)$, see Section C.2.1. Consequently, as $t \to 0^+$, the guidance is well-approximated by a flow on $\mathcal{M}$,

$$dY_t = -\frac{1}{\Delta t_0}\,P_{T_{Y_t}\mathcal{M}}\,\nabla\|\mathcal{A}(Y_t) - y\|_2\,dt,$$

where $P_{T_{Y_t}\mathcal{M}}$ denotes the orthogonal projection onto the tangent space at Y_t. Taking a forward Euler step of size Δt_0, the prefactor $1/\Delta t_0$ cancels the step size exactly, so one Euler step corresponds to one full projected-gradient step on $\mathcal{M}$, independently of Δt_0.

Local Lipschitz constant. The local Lipschitz constant of the projected drift on $\mathcal{M}$ scales as

$$L_{\text{lip}} \;\sim\; \frac{1}{\Delta t_0}\,\frac{\sigma_{\max}(\nabla\mathcal{A}\,P_{T_Y\mathcal{M}})^2}{\|\mathcal{A}(Y) - y\|_2}, \tag{71}$$

where the residual in the denominator originates from the gradient of the unsquared norm, which renormalizes to unit magnitude as Y approaches the constraint.

Stability criterion and inevitable oscillations. Forward Euler stability requires $\Delta t_0 \cdot L_{\text{lip}} \le 2$, i.e.,

$$\sigma_{\max}(\nabla\mathcal{A}\,P_{T_Y\mathcal{M}})^2 \;\le\; 2\,\|\mathcal{A}(Y) - y\|_2. \tag{72}$$

As Y approaches the constraint set $\{Y : \mathcal{A}(Y) = y\}$, the right-hand side tends to zero while the left-hand side depends only on $\mathcal{A}$ and the local geometry of $\mathcal{M}$. The criterion is therefore inevitably violated near the constraint. This is the standard pathology of forward Euler applied to the unsquared norm: the gradient does not vanish as the residual shrinks, but merely renormalizes to unit magnitude along $\nabla\mathcal{A}^\top(\mathcal{A}(Y) - y)/\|\mathcal{A}(Y) - y\|_2$. The iteration overshoots and oscillates around the constraint, and no choice of step size can restore stability/convergence. We note that these oscillations occur only parallel to the data manifold. Implicit integration avoids these instabilities, see Fig. 4.

I Empirical Evidence of Instabilities

LQG instability. *Setup.* Take the prior $\rho_* = \mathcal{N}(\mu, \sigma^2 I)$ on $\mathbb{R}^2$ with $\mu = (1, 1)$ and $\sigma^2 = 0.05$, and observe the second coordinate, $\mathcal{A}(x) = a^\top x$ with $a = (0, 1)$ and $y = 1.5$. The constraint line $\{x_2 = y\}$ sits 2.2σ into the tail of the prior. We run $N = 1000$ steps of the linear schedule of Section G with guidance strength $\alpha = 0.35$. This experiment steps along the gradient of the unsquared residual, half the step of Algorithm 1, so under the schedule (14) DPS targets (17) with $\alpha/2$ in place of α, which here is

$$\mu_y^{\text{Target}}(x) \;\propto\; e^{-\frac{\alpha}{2}\eta_0\,|x_2 - y|}\,\rho_*(x), \qquad \tfrac{\alpha}{2}\eta_0 = \alpha/(2\Delta t_0) \approx 3.2 \times 10^3 :$$

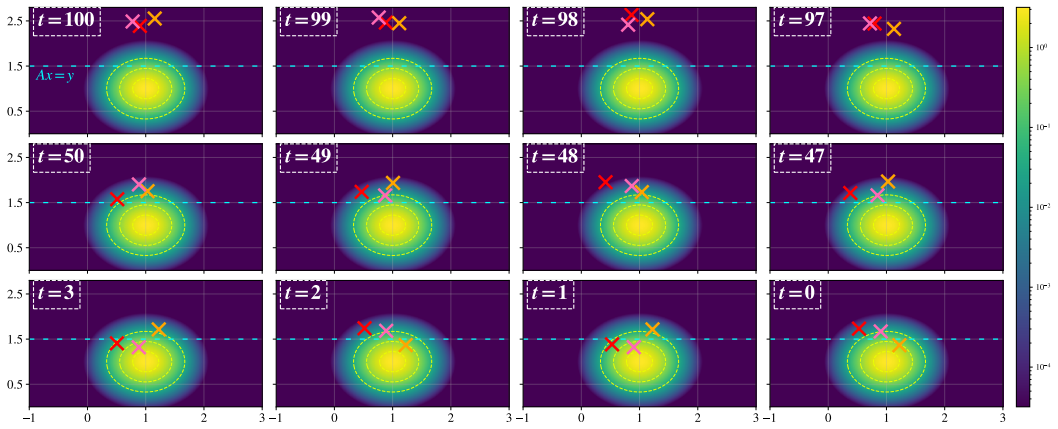
a Laplace tilt of the prior which, to within 10^{-3}, is the prior restricted to the line, i.e. $x_2 = y$ and $x_1 \sim \mathcal{N}(1, \sigma^2)$.

The DPS SDE. Write $f_s(x) \coloneqq \alpha\,|y - a^\top\hat{x}_s(x)|$. With the schedule (14), the guidance term of (DPS SDE) becomes a gradient flow on f scaled by the annealing schedule $\eta_t = 1/\Delta t(t)$ (Section H):

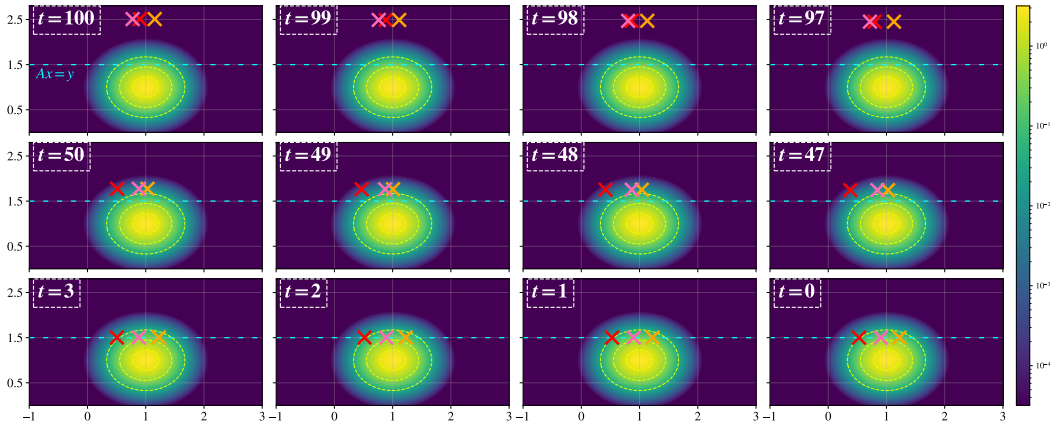
$$dY_t = \big(Y_t + 2\nabla\log\rho_{T-t}(Y_t)\big)\,dt \;-\; \eta_{T-t}\,\nabla f_{T-t}(Y_t)\,dt \;+\; \sqrt{2}\,dB_t.$$

DPS integrates the first term with the DDPM step, $\tilde{Y}_{i-1} = \text{DDPM}(Y_i)$, and the guidance term with forward Euler over the step's OU time Δt_i:

$$Y_{i-1} = \tilde{Y}_{i-1} - \Delta t_i\,\eta_i\nabla f_i(Y_i) = \tilde{Y}_{i-1} - \nabla f_i(Y_i) = \tilde{Y}_{i-1} + \alpha\,\text{sign}(r_i)\,J_i^\top a,$$



(a) DPS, final iterations



(b) Implicit-Integration DPS, final iterations

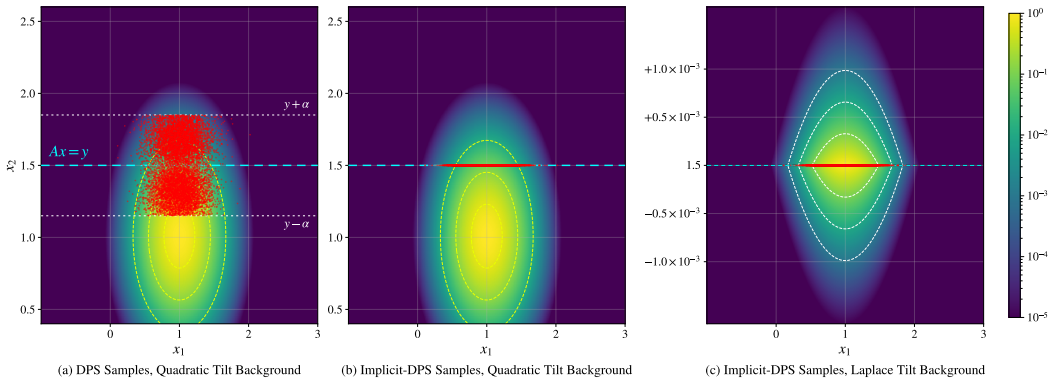
Figure 4: Three particles under **(a)** DPS and **(b)** implicitly integrated DPS (same prior, measurement, $\alpha = 0.35$, noise schedule and seed). Rows show steps $i = 100$–97 and 3–0 ($i = 0$ is the output), over the prior density (log scale) with its $1, 2, 3\sigma$ contours dashed; the cyan line is $\{x_2 = y\}$. Color identifies a particle, fixed across panels. DPS particles jump across the line at every step, while the implicit particles settle onto it.

where $r_i = y - a^\top \hat{x}_{t_i}(Y_i)$ and $J_i = \nabla \hat{x}_{t_i}(Y_i)$. The step size cancels the schedule, so every step has fixed length $\alpha \|J_i^\top a\| \to \alpha$ however small the residual. Once $|r_i| < \alpha$ the step overshoots, the sign flips, and the next step returns: a period-2 orbit of amplitude α about the constraint. Here $\alpha = 0.35$ exceeds the prior width $\sigma \approx 0.22$. Figure 4(a) shows this orbit on three particles.

Implicit-Guidance-DPS. Backward Euler on the same drift over the same step evaluates ∇f at the new point and the new noise level, with the schedule held at its step-i value $\eta_i = 1/\Delta t_i$ (it stands in for that step's missing Δt_i):

$$Y_{i-1} = \tilde{Y}_{i-1} - \Delta t_i \, \eta_i \, \nabla f_{i-1}(Y_{i-1}) = \tilde{Y}_{i-1} - \nabla f_{i-1}(Y_{i-1}) \quad \Longleftrightarrow \quad Y_{i-1} = \mathrm{prox}_{f_{i-1}}(\tilde{Y}_{i-1}).$$

Since $f_{i-1} = \alpha \, |y - a^\top \hat{x}_{t_{i-1}}|$, this is the prox of the residual magnitude with parameter α, the same at every step: as in DPS, the step size and the schedule cancel. The prior score stays in the explicit DDPM step; only the guidance is implicit. Because the prior is Gaussian, $\hat{x}_{t_{i-1}}$ is affine with a constant Jacobian J_{i-1}, so f_{i-1} is α times the absolute value of an affine function, and its prox (the unique solution of the implicit equation) is a soft-threshold of the residual at level $\alpha \|v\|^2$ along



(a) DPS Samples, Quadratic Tilt Background (b) Implicit-DPS Samples, Quadratic Tilt Background (c) Implicit-DPS Samples, Laplace Tilt Background

Figure 5: Final states of 10^4 particles, same setup as Fig. 4. **(a)** DPS and **(b)** implicitly integrated DPS, over the density $\propto e^{-\alpha(x_2-y)^2/2}\rho_*$ (log scale) with its $1, 2, 3\sigma$ contours dashed; the dotted lines in (a) mark $y \pm \alpha$. **(c)** A different background: the samples of (b) over the Laplace tilt μ_y^{Target} (shown as a fraction of its maximum, log scale), on the same x_1 range but magnified $669\times$ in x_2, with contours at whole decades. DPS spreads its samples over the band $y \pm \alpha$; the implicit rule places every sample on the line, where μ_y^{Target} concentrates.

$v = J_{i-1}^\top a$. With $\tilde{r} = y - a^\top \hat{x}_{t_{i-1}}(\tilde{Y}_{i-1})$,

$$Y_{i-1} = \tilde{Y}_{i-1} + \lambda\, v, \qquad \lambda = \operatorname{sign}(\tilde{r})\, \min\bigl(\alpha,\ |\tilde{r}|/\|v\|^2\bigr).$$

Far from the constraint this is the same α-sized step as DPS; within reach it sets the residual to exactly zero rather than overshooting. At the last step $\hat{x} = x$ and $J = I$, so the output satisfies $x_2 = y$ exactly; Fig. 4(b) shows the same three particles under this rule.

Both rules are run with the same seed and noise, so they differ only in how the guidance is integrated. Under DPS (Fig. 4(a)) the particles alternate sides of the line at every step with x_1 fixed, and end at $x_2 \in \{1.73,\ 1.36,\ 1.67\}$: the period-2 orbit predicted above. Under implicit integration (Fig. 4(b)) they descend onto the line and stay, with residual exactly zero in float32 at the output. Figure 5 repeats the comparison on the final states of 10^4 particles.

Across 10^4 particles (Fig. 5), the DPS orbit, stopped at a random phase, fills a band of half-width α about the line, while the implicit rule places every sample on it. The guidance never moves x_1, which keeps its prior law $\mathcal{N}(1, \sigma^2)$, as under μ_y^{Target}. The entire discrepancy between the two rules is therefore the $\pm\alpha$ oscillation in x_2.

Conditional Guidance for MNIST digits We consider the setting of posterior sampling with the MNIST prior. This dataset consists of paired images of handwritten digits and their corresponding labels, denoted (x, y). We train a simple MLP classifier $\text{softmax}(f(x)) \approx \mathbb{1}_y$ over the MNIST dataset, and set the reward to be $R(x) = -\|f(x) - \mathbb{1}_k\|$ for some fixed target k. We run DPS with guidance schedule constant at 0.1. The evolution of $\|f(x_t) - \mathbb{1}_k\|$ as well as $(f(x_t) - \mathbb{1}_k) \cdot \mathbb{1}$ is plotted below.

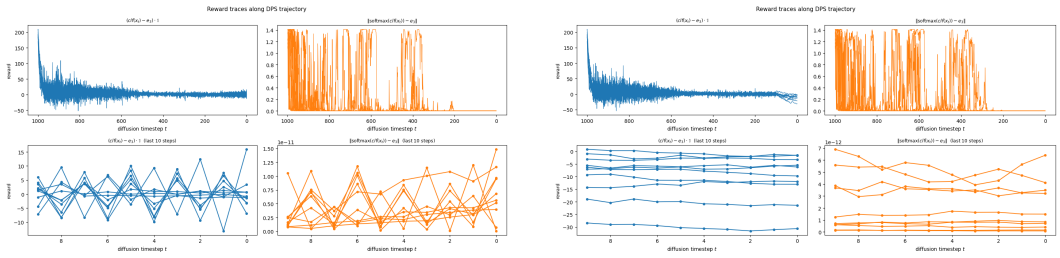


Figure 6: We plot a projected discrepancy $(f(x_t) - \mathbb{1}_k) \cdot \mathbb{1}$ (**(Columns 1 and 3)**) and the distance $\|f(x_t) - \mathbb{1}_k\| = -R(x_t)$ (**(Columns 2 and 4)**) across t where denoising proceeds from left (most noise) to right (least noise). The second row depicts a close-up plot of just the last 10 steps to highlight the oscillations. **(Columns 1 and 2)** are run with a constant guidance schedule Algorithm 1, while **(Column 3 and 4)** are run with early guidance stopping Algorithm 2 with parameter $i_{\text{stop}} = 100$.

We see a distinct oscillatory pattern is sustained throughout the trajectory when the guidance schedule is constant Algorithm 1. Turning off the guidance schedule Algorithm 2 at time-step $i_{\text{stop}} = 100$ eliminates the oscillations in that period, though now the distance $\|f(x_t) - \mathbb{1}_k\|$ is not pulled toward 0. This indicates that the instability is associated with the reward guidance. We see in either case that a softmax applied to the logits of the classifier results in a very high confidence prediction of the correct class despite these oscillations.

We also plot the alignment between consecutive steps of the algorithm $\delta_t = x_t - x_{t-1}$. Because these vectors lie in 784 dimensions, to emphasize the step over step alignment we maintain a subspace described by the most recent $\ell = 50$ such steps. In particular, let $\mathcal{P}_t = [\delta_{t-\ell+1}, \delta_{t-\ell+2}, \cdots \delta_t] \in \mathbb{R}^{784\times\ell}$, and let $\mathcal{P}_t^{:k}$ denote just the projection onto the top k principal axes. We plot $\alpha_t = (\mathcal{P}_t^{:k}\delta_{t-1}) \cdot (\mathcal{P}_t^{:k}\delta_t)$ over the trajectory in Fig. 7. For a purely oscillating trajectory, we expect $\delta_t \approx -\delta_{t-1}$, resulting in $\alpha_t \approx -1$. When δ_t is "unrelated" to δ_{t-1}, we expect $\alpha_t \approx 0$.

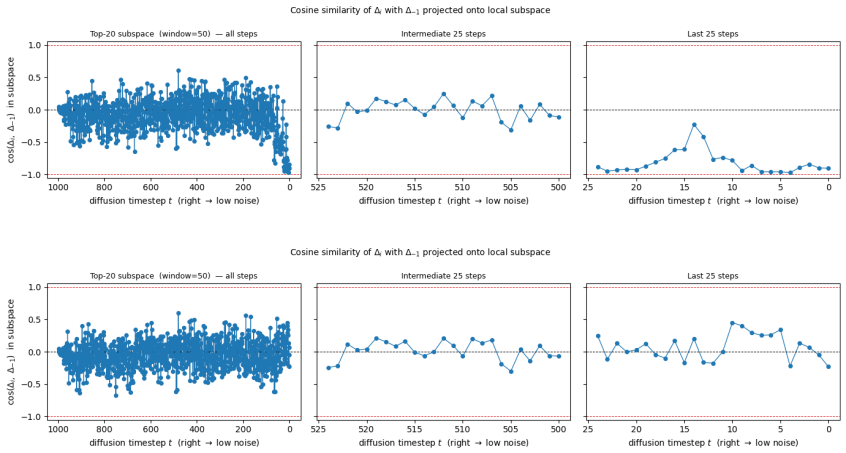


Figure 7: **Top** Are plots associated to the standard DPS algorithm Algorithm 1, **Top Left** We plot $\alpha_t = (\mathcal{P}_t^{:k}\delta_t) \cdot (\mathcal{P}_t^{:k}\delta_{t-1})$ along the DPS trajectories for a constant guidance schedule $\zeta = 0.1$. **Top Middle** A close-up of steps $525 \to 500$. **Top Right** A close-up of steps $25 \to 0$. Note that α_t is close to 0 at the intermediate noise levels, but drops to ≈ -1 towards the low noise levels. **Bottom** Are the same plots associated to Algorithm 2 with $i_{\text{stop}} = 100$. We observe that α_t remains close to 0 both at intermediate noise levels and low noise levels.

All experiments were run in a few minutes on a single NVIDIA H100 GPU.