CLASSIFIER-FREE GUIDANCE FROM DDPM TO DDIM: INVERSE OPTIMAL CONTROL AND TERMINAL LAWS

Anonymous authors
Paper under double-blind review

ABSTRACT

Classifier-free guidance (CFG) applies the same score extrapolation in stochastic and deterministic diffusion samplers, yet generally produces different output distributions. We analyze this sampler dependence through inverse optimal control and density transport in the continuous-time, exact-score setting. We identify explicit running terms that make prescribed CFG policies optimal under specified quadratic effort metrics. The resulting terminal-density representations retain initialization and trajectory-dependent corrections to the endpoint power tilt. They connect stochastic reference-path averaging to deterministic guided transport through a zero-noise limit under stated regularity conditions. The characterization extends to smooth time–evidence potentials and motivates coefficient-budget coordinates for schedule selection and Online-K, which adapts guidance using a numerical evidence estimate. Analytic Gaussian-mixture experiments support the density identities and demonstrate that matching coefficient budgets does not determine the terminal distribution. On CIFAR-10, budget-based selection improves power-cosine schedules (PCS) within the tested search grids. Online-K reduces mean FID relative to PCS by 1.90% in paired DDIM-200 evaluations on CIFAR-10, with improvements across all three sampling seeds, and by 2.98% with pretrained DiT-XL/2 and DDIM-50 on ImageNet 256 × 256.

1 INTRODUCTION

Diffusion models generate data by reversing a progressive noising process (Sohl-Dickstein et al., 2015; Ho et al., 2020; Song et al., 2021b). Classifier-free guidance (CFG) combines conditional and unconditional score estimates to strengthen adherence to a condition without a separate classifier (Ho & Salimans, 2022), and is widely used in latent diffusion models and diffusion transformers (Rombach et al., 2022; Peebles & Xie, 2023). It can be applied to stochastic DDPM and deterministic DDIM sampling (Ho et al., 2020; Song et al., 2021a). However, the same guidance scale can produce different output distributions across these samplers (Bradley & Nakkiran, 2025).

This sampler dependence highlights a distinction between a local score modification and its accumulated distributional effect. With exact scores and a fixed scale, CFG follows the score of a geometric combination of conditional and unconditional densities. This local identity does not ensure that the sampler generates that density, as established in analyses of compositional diffusion sampling and CFG (Du et al., 2023; Bradley & Nakkiran, 2025). The evolving distribution must also satisfy the sampler's dynamics and initial condition; related work develops nonlinear guidance corrections from Fokker–Planck consistency (Zheng & Lan, 2024). We therefore ask:

For a prescribed CFG policy, what objective makes it optimal, and how does that objective relate to its terminal distribution?

Understanding this connection matters when transferring guidance between samplers or changing its allocation over time.

Existing theory characterizes changes in confidence, diversity, and distributional shape under mixture or high-dimensional assumptions (Wu et al., 2024; Chidambaram et al., 2024; Pavasovic et al., 2026). Guidance design also exploits the evolving sampling trajectory. Time-varying schedules and guidance restricted to intermediate noise levels can improve generation (Wang et al., 2024; Kynkäänniemi

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.

et al., 2024), while stage-wise analyses distinguish early mode allocation from later contraction within modes (Jin et al., 2026). Feedback Guidance and Dynamic CFG adapt scales using estimated conditional evidence or latent-space evaluators (Koulischer et al., 2025; Papalampidi et al., 2026). These studies motivate accounting for both guidance timing and the states visited.

Control and density-evolution methods provide closely related perspectives. Schrödinger bridges formulate generative modeling through path-space entropy minimization (De Bortoli et al., 2021), while stochastic control provides samplers for unnormalized densities and a framework for interpreting diffusion-based generation (Zhang & Chen, 2022; Berner et al., 2024). Azangulov et al. (2026) optimize adaptive guidance for a specified terminal-reward objective regularized by path relative entropy. Feynman–Kac correctors use weighted simulation and resampling to realize prescribed density combinations, including corrections involving accumulated score separation (Skreta et al., 2025). Independently developed concurrent work by Jiang & Ma (2026) derives exact deterministic CFG density formulas for constant and time-dependent scales. Our focus is the connection between an inverse-control objective and the terminal distribution of the same prescribed policy, retaining initialization and linking stochastic sampling with deterministic transport.

We study this connection in the continuous-time, exact-score setting. Under specified quadratic effort metrics, we identify the additional running term that makes a prescribed CFG policy optimal. For stochastic reverse diffusion, this control residual also enters the path likelihood and terminal-density correction. Deterministic sampling requires a separate density calculation because changing trajectories can make path relative entropy infinite even when endpoint densities remain smooth. The resulting representation distinguishes averaging over stochastic reference paths conditioned on an endpoint from transport along a deterministic guided trajectory. Under stated regularity conditions, the density correction converges pointwise to its deterministic counterpart as sampler noise vanishes.

The analysis extends to smooth potentials depending on time and conditional evidence, motivating coefficient-budget coordinates for power-cosine schedule (PCS) selection and evidence-responsive guidance with Online-K (Estimating conditonal evidence $log\ p_t(c|X_t)$ step by step during generation process). The budget summarizes schedule coefficients rather than the full control cost, and Online-K uses an estimated evidence signal outside the exact-score guarantees. Gaussian-mixture experiments support the density identities and show that changing guidance timing changes output distributions even when coefficient budgets and scale histograms match. On CIFAR-10, budget-based PCS selection improves FID within the tested grids with equal candidate counts. Online-K lowers mean FID by 1.90% relative to a fixed PCS reference, with improvements in all three paired sampling seeds. A separate ImageNet variant using a batch-shared energy proxy lowers FID by 4.82% in one paired evaluation, with lower feature precision and recall.

Our contributions are threefold. First, we characterize fixed-scale CFG and smooth time–evidence policies through explicit inverse-control objectives and exact optimality-gap identities under specified effort metrics. Second, we derive terminal-density representations retaining initialization and guidance history, and connect stochastic averaging with deterministic transport through a zero-noise limit. Third, we translate this characterization into coefficient-budget coordinates for PCS selection and an evidence-responsive policy, and evaluate their practical performance alongside analytic tests of density corrections and guidance timing.

2 Background on diffusion sampling and CFG

2.1 Sampling dynamics, classifier–free guidance and conditional evidence

DDPM uses stochastic reverse transitions, while deterministic DDIM changes the update mean and removes sampling noise (Ho et al., 2020; Song et al., 2021a). For unconditional and conditional score estimates $\widehat{s}_t$ and $\widehat{s}^c_t$, CFG uses (Ho & Salimans, 2022)

$$\widehat{s}^{(w)}_t = \widehat{s}_t + w(\widehat{s}^c_t - \widehat{s}_t). \tag{1}$$

$w = 0$ is unconditional, $w = 1$ is conditional, and $w > 1$ extrapolates.

We study continuous-time sampling with exact scores. Consider the forward process

$$dZ_s = f_s(Z_s)\, ds + g_s\, dB_s, \qquad s \in [0, T], \tag{2}$$

where Z_0 has the data distribution and $g_s > 0$ is state independent. We use generation time $t = T - s$, from noise at $t = 0$ to data at $t = T$. Let p_t and p_t^c denote the unconditional and conditional densities of Z_{T-t} under the same forward dynamics, with scores $s_t = \nabla \log p_t$ and $s_t^c = \nabla \log p_t^c$.

Set $\sigma_t = g_{T-t} I$, $A_t = \sigma_t \sigma_t^\top$, and $M_t = A_t/2$. The reverse SDE $dX_t = b_t(X_t)\,dt + \sigma_t\,dW_t$ and probability-flow ODE $dX_t = v_t(X_t)\,dt$ have drifts

$$\begin{aligned} b_t(x) &= -f_{T-t}(x) + A_t s_t(x), \\ v_t(x) &= -f_{T-t}(x) + M_t s_t(x). \end{aligned} \tag{3}$$

Both processes have marginals p_t from p_0 (Anderson, 1982; Song et al., 2021b); B, W are standard Brownian motions. State-independent positive-definite A_t are also allowed.

Define the conditional evidence K_t and its terminal value r by

$$K_t(x) = \log \frac{p_t^c(x)}{p_t(x)}, \qquad \nabla K_t = s_t^c - s_t, \qquad r = K_T. \tag{4}$$

For a discrete condition c, Bayes' rule gives $K_t(x) = \log p(c \mid x, t) - \log p(c)$, measuring evidence relative to the prior.

A common representation is

$$\begin{aligned} dX_t = \big[v_t + \eta^2 M_t s_t + (1+\eta^2) a(t,\cdot) M_t \nabla K_t\big](X_t)\,dt \\ + \eta \sigma_t\,dW_t, \qquad X_0 \sim \nu, \end{aligned} \tag{5}$$

where $\eta \in [0,1]$ and $a(t,x)$ is a scalar guidance scale. At $a \equiv w$, $\eta = 1$ adds $wA_t\nabla K_t$ to the reverse-SDE drift, while $\eta = 0$ adds $wM_t\nabla K_t$ to the ODE drift. At $w = 1$, both recover p_t^c when initialized from p_0^c. We denote the resulting marginal density by $\rho_t^{\eta,a}$.

We assume positive $C^{1,2}$ population densities and well-posed dynamics. Appendix A details the correspondence with posterior-variance DDPM and deterministic DDIM on a common sampling grid, and gives sufficient convergence conditions on a truncated variance-preserving noise interval. Additional assumptions are stated with each theoretical result. These exact continuous-time identities do not extend automatically to learned scores, clipping, or arbitrary finite-step samplers.

2.2 Conditional evidence and stochastic control

Conditional evidence provides a common starting point for guidance analysis. Write $\|z\|_B^2 = z^\top B z$ and define the score-separation energy

$$E_t(x) = \|\nabla K_t(x)\|_{M_t}^2. \tag{6}$$

Azangulov et al. (2026) derive the stochastic evolution of this evidence and use a terminal-reward, path-KL control formulation to optimize adaptive guidance. In our notation, the reverse-SDE evidence identity is

$$dK_t(X_t) = (2a_t - 1)E_t(X_t)\,dt + \nabla K_t(X_t)^\top \sigma_t\,dW_t, \tag{7}$$

where $a_t = a(t, X_t)$. Their Lemma 1 gives this identity for Ornstein–Uhlenbeck noising, with $M = I$ and our scale $a = 1 + w_{\mathrm{A}}$, where w_{A} denotes their additional guidance above the conditional score. Their control objective penalizes path KL relative to the conditional reverse diffusion.The link between quadratic drift effort and path KL follows from Girsanov's theorem and is closely related to variational representations of Brownian functionals (Boué & Dupuis, 1998); reward-only control and Doob transforms are classical tools (Fleming & Soner, 2006; Kappen, 2005; Léonard, 2014).(Fleming & Soner, 2006; Kappen, 2005; Léonard, 2014). Appendix A.3 records the covariance and sampler conventions needed here.

Feynman–Kac correctors use weighted diffusion and resampling to realize prescribed density combinations (Skreta et al., 2025). Their CFG correction contains the score-separation term $w(w-1)E_t$; time-dependent density powers also introduce schedule-derivative terms. We further express the terminal laws of prescribed guidance policies through initialization factors and path corrections, connecting stochastic reference-path averaging to deterministic guided transport. The time–evidence formulation links these density corrections to the corresponding control objectives.

2.3 Problem formulation

For prescribed initialization ν and scale w, we seek (i) a running term that makes CFG optimal for terminal reward $wr(X_T)$ under drift-effort metrics A_t^{-1} (SDE) and M_t^{-1} (ODE), and (ii) a terminal-law correction to the power tilt $p_T e^{wr} = p_T^{1-w}(p_T^c)^w$. The metrics match the guidance coefficients in Eq. (5). Sections 3–4 characterize the endpoints $\eta \in \{0, 1\}$; Appendices B–C give the general family. Section 5 extends the potential wK to time- and evidence-dependent policies.

3 An inverse-control characterization of CFG

We identify the running term that makes fixed-scale CFG optimal under the effort metrics specified in Section 2.3. Fix $w \geq 0$ and an initial density ν with $\mathbb{E}_\nu|K_0| < \infty$, and use the energy E_t defined in (6).

Reverse diffusion. Consider the controlled SDE

$$dX_t = \big(b_t(X_t) + u_t\big)\,dt + \sigma_t\,dW_t, \qquad X_0 \sim \nu, \tag{8}$$

and let $\mathbb{E}^u$ denote expectation under this process. Controls are progressively measurable with respect to the state filtration, and all controlled processes start from ν. We impose the verification conditions in Appendix B.1, including nonexplosive dynamics, finite effort, integrable objective terms, a zero-mean verification martingale, and admissibility of the CFG feedback.

A terminal reward wr with quadratic effort in the metric A_t^{-1} generally yields a policy different from CFG. The additional running term below makes the prescribed CFG feedback optimal.

Theorem 1 (CFG objective for reverse diffusion)**.** *For the objective*

$$J_{\mathrm{S}}^w(u) = \mathbb{E}^u\left[wr(X_T) - \int_0^T \left\{w(w-1)E_t(X_t) + \frac{1}{2}\|u_t\|^2_{A_t^{-1}}\right\} dt\right], \tag{9}$$

CFG attains the optimal value:

$$\sup_u J_{\mathrm{S}}^w(u) = w\mathbb{E}_\nu K_0, \qquad u_{\mathrm{S}}^\star(t,x) = wA_t\nabla K_t(x). \tag{10}$$

For every admissible control u, the optimality gap is

$$w\mathbb{E}_\nu K_0 - J_{\mathrm{S}}^w(u) = \frac{1}{2}\mathbb{E}^u\int_0^T \big\|u_t - u_{\mathrm{S}}^\star(t, X_t)\big\|^2_{A_t^{-1}}\,dt. \tag{11}$$

The gap identity quantifies the loss from any admissible deviation from the CFG feedback. For $w > 1$, the objective balances terminal evidence, quadratic control effort, and a running penalty on score separation. The running term vanishes at $w = 1$. Under the additional Girsanov conditions in Appendix B.1, the expected quadratic effort equals the path KL divergence relative to the uncontrolled reverse SDE with the same initialization.

Deterministic sampling. Now consider $dX_t = (v_t(X_t) + u_t)\,dt$, again with $X_0 \sim \nu$. The effort metric is M_t^{-1}. Define

$$\begin{aligned} H_t(x) &= \frac{\operatorname{div}\big(p_t M_t \nabla K_t\big)(x)}{p_t(x)}, \\ R_w^0(t,x) &= -wH_t(x) + \left(\frac{w^2}{2} - w\right) E_t(x). \end{aligned} \tag{12}$$

The quantity H_t describes how the field $M_t\nabla K_t$ redistributes the reference density. Expectations below are taken under the controlled ODE.

Corollary 2 (CFG objective for deterministic sampling)**.** *Under the deterministic assumptions of Theorem 6, define*

$$J_{\mathrm{D}}^w(u) = \mathbb{E}^u\left[wr(X_T) - \int_0^T \left\{R_w^0(t, X_t) + \frac{1}{2}\|u_t\|^2_{M_t^{-1}}\right\} dt\right]. \tag{13}$$

Then $\sup_u J_{\mathrm{D}}^w(u) = w\mathbb{E}_\nu K_0$, attained by $u_{\mathrm{D}}^\star(t,x) = wM_t\nabla K_t(x)$. The gap identity (11) holds with $(J_{\mathrm{S}}^w, u_{\mathrm{S}}^\star, A)$ replaced by $(J_{\mathrm{D}}^w, u_{\mathrm{D}}^\star, M)$.

The running term now contains H_t and need not be nonnegative. The quadratic effort remains finite for admissible controls, but has no general path-KL interpretation: a controlled law assigning positive mass to paths that do not solve the reference ODE has infinite KL relative to the reference law.

Common structure. Both objectives arise from the potential $F_t = wK_t$, with effort metrics matched to the respective CFG drift coefficients. Their running terms differ because conditional evidence changes differently along the stochastic and deterministic reference trajectories. In the inverse-control sense of seeking an objective compatible with a prescribed policy (Kalman, 1964), our characterization is tied to these metrics and does not identify a unique objective independently of them. All proofs, the reward-only benchmark, and the general-η formulation are given in Appendix B.

We next characterize the terminal distributions induced by these policies. The stochastic running term reappears in the density correction, while the deterministic case requires a separate transport calculation.

4 The terminal distribution of CFG

Fix $w \geq 0$ and an initial density $\nu \ll p_0$. Define the boundary factor

$$C_0^w(x) = \frac{\nu(x)}{p_0(x)} e^{-wK_0(x)}, \tag{14}$$

which accounts for both initialization and initial conditional evidence. Even when $\nu = p_0$, this factor need not be constant.

Reverse diffusion. Let P denote the unconditional reverse-SDE law initialized from p_0, so its terminal density is p_T.

Theorem 3 (Terminal density of stochastic CFG)**.** *Under the Girsanov and integrability conditions in Appendix C.1, the CFG process initialized from ν has, for p_T-almost every y, terminal density*

$$\rho_T^{1,w}(y) = p_T(y) e^{wr(y)} \times \mathbb{E}_P\left[C_0^w(X_0) \exp\left\{ -w(w-1) \int_0^T E_t(X_t)\, dt \right\} \middle| X_T = y \right]. \tag{15}$$

The prefactor is the power tilt $p_T e^{wr} = p_T^{1-w} (p_T^c)^w$. Its correction averages the product of the boundary factor and the exponential path weight over unconditional reference paths ending at y. Both the initial state and intermediate path remain random under this conditioning. For $w > 1$, the exponential factor downweights paths with greater accumulated score separation. The running term $w(w-1)E_t$ is exactly the one in the SDE control objective of Section 3.

Deterministic sampling. Assume the guided ODE generates an invertible C^1 flow on $[0, T]$. Density evolution along such flows is governed by the divergence of the velocity field, as in continuous normalizing flows (Chen et al., 2018). Let x_t^y denote its trajectory ending at y:

$$\dot{x}_t^y = v_t(x_t^y) + wM_t \nabla K_t(x_t^y), \qquad x_T^y = y. \tag{16}$$

Corollary 4 (Terminal density of deterministic CFG)**.** *Under the density regularity conditions of Theorem 7, its $\eta = 0$ specialization gives*

$$\rho_T^{0,w}(y) = p_T(y) e^{wr(y)} C_0^w(x_0^y) \times \exp\left\{ -w(w-1) \int_0^T E_t(x_t^y)\, dt \right\}. \tag{17}$$

At a fixed endpoint, the correction is evaluated along a single guided trajectory, which itself depends on w. Flow compression is already accounted for by this density formula. Its running term $w(w-1)E_t$ differs from the ODE control term R_w^0: the two characterize density transport and the optimization objective, respectively.

When the power tilt is exact. Both terminal laws are normalized as written. If $Z_w = \int p_T(y)e^{wr(y)}\,dy < \infty$, the sampler generates the normalized power tilt precisely when its correction equals Z_w^{-1} almost everywhere. Thus the correction identifies the endpoint-dependent distortion left unspecified by the local score field. Even a constant boundary factor does not ensure a constant correction: accumulated score separation can change the output shape, as the Gaussian example in Appendix F shows.

What initialization changes. The factor C_0^w carries the initial density and initial evidence into the terminal law. Initializing from p_0 removes the density ratio ν/p_0, but leaves e^{-wK_0}. At $w = 1$ with $\nu = p_0^c$, both the boundary and path factors equal one, recovering p_T^c. The terminal formula therefore separates an initialization effect from the effect accumulated during guidance. Finite-noise initialization has also been identified as a practical source of mismatch between training and sampling (Lin et al., 2024).

What the sampler changes. The same energy coefficient $w(w-1)$ produces different corrections through different trajectories: reverse diffusion averages over unconditional reference bridges, while deterministic sampling evaluates a single guided characteristic. The deterministic transport coefficient also differs from its control residual R_w^0. Consequently, a control objective alone does not specify the power-tilt correction. For fixed ν and w, the general-family correction converges pointwise to (17) as $\eta \downarrow 0$ under the conditions in Appendix C.3. Appendix C provides the auxiliary-process representation and proofs.

Together, Sections 3 and 4 describe each fixed CFG policy through a control objective and its terminal density. The next section extends the common potential wK to functions of time and evidence, so that guidance design can account for both the objective and the induced distribution.

5 From Characterization to Guidance Design

Extending the potential wK connects guidance timing and evidence response to their control objectives and terminal laws.

5.1 A Common Potential for Guidance Design

For a smooth function $\phi(t, z)$, define

$$F_t(x) = \phi(t, K_t(x)), \qquad a(t, x) = \phi_z(t, K_t(x)). \tag{18}$$

Corollary 5 (Time–evidence potentials)**.** *Under the assumptions of Theorems 6 and 7, the potential F gives optimal drift increments $A_t a \nabla K_t$ for the SDE and $M_t a \nabla K_t$ for the ODE, under the respective effort metrics of Section 3. The pointwise value is $F_t(x)$, the terminal reward is $\phi(T, r)$, and the control running terms are*

$$\begin{aligned} R_\phi^1 = R_\phi^{\mathrm{tr}} &= \phi_t + (\phi_{zz} + \phi_z^2 - \phi_z)E, \\ R_\phi^0 &= \phi_t - \phi_z H + \left(\tfrac{1}{2}\phi_z^2 - \phi_z\right) E. \end{aligned} \tag{19}$$

Derivatives of ϕ are evaluated at $(t, K_t(x))$. Both density representations use R_ϕ^{tr}, with terminal factor $p_T e^{\phi(T,r)}$ and boundary factor $(\nu/p_0)e^{-\phi(0,K_0)}$.

The SDE control and transport terms coincide, whereas the ODE control term retains a different correction. The shared transport residual nevertheless provides a common basis for design: choosing ϕ specifies the guidance response together with its control and distributional consequences. Optimality is relative to this induced objective; it does not rank arbitrary guidance policies by sample quality. We consider time schedules, $\phi(t, z) = a(t)z$, and evidence responses, $\phi(t, z) = \psi(z)$. Proofs and the general-η formulas are in Appendix D.

5.2 Time Allocation through Budget Coordinates

For $\phi(t, z) = a(t)z$, the transport residual becomes

$$R_a^{\mathrm{tr}} = \dot{a}K + (a^2 - a)E. \tag{20}$$

The density correction therefore depends jointly on schedule changes, guidance strength, and the states visited. We summarize the coefficient contribution by

$$B_{\mathrm{coef}}[a] = \frac{1}{2}\int_0^1 \left[a(T\tau)^2 - a(T\tau)\right] d\tau, \tag{21}$$

and use this coefficient statistic as a coordinate for schedule search. For schedules with $a \geq 1$, it summarizes guidance strength and duration in a form that is easy to compute from schedule parameters. It omits $\dot{a}K$, local energy, and trajectory dependence, so it neither determines the terminal distribution on its own nor represents the complete control cost. The matched-budget timing experiment tests precisely this remaining dependence on when guidance is applied.

We apply these coordinates to power-cosine schedules (PCS) (Wang et al., 2024). On the implemented $N = 200$ grid, let $\tau_k = k/(N-1)$ and define

$$\begin{aligned} r_s(\tau) &= \frac{1-\cos(\pi\tau^s)}{2}, \qquad a_k = 1 + (W-1)r_s(\tau_k), \\ B_N(W,s) &= \frac{1}{2N}\sum_{k=0}^{N-1}(a_k^2 - a_k), \quad W > 1, \quad s > 0. \end{aligned} \tag{22}$$

For fixed peak W, $B_N(W,s)$ strictly decreases with s. Each feasible peak–budget pair therefore determines a unique shape, found by bisection; the discrete mean is $m = 1+(W-1)N^{-1}\sum_k r_s(\tau_k)$. Appendix I.1 gives the feasible interval and inversion details.

We prescribe a finite set of peak–budget pairs, convert them to PCS parameters, and select the candidate with the lowest development FID, subject to the accuracy requirement when applicable. The selected configuration is locked before independent final evaluation. Theory supplies the search coordinates and their inverse; development performance selects the operating point.

5.3 Evidence-responsive guidance with Online-K

Time schedules assign the same scale to every sample at a given step. To make guidance respond to conditional evidence, we choose

$$\begin{aligned} \psi(z) &= z + \frac{\lambda}{\gamma}\log\left(1 + e^{\gamma(z-\kappa)}\right), \\ a_\psi(t,x) &= \psi'(K_t(x)) = 1 + \lambda\,\mathrm{sigmoid}[\gamma(K_t(x)-\kappa)], \end{aligned} \tag{23}$$

where $\lambda \geq 0$ and $\gamma > 0$. The scale lies in $[1, 1+\lambda]$, with threshold κ and transition sharpness γ. For $\lambda > 0$, higher evidence receives stronger guidance. The potential's slope sets the scale, while its curvature contributes to the transport residual in (19).

The score difference gives ∇K, while the value of K also depends on its evolution in time. Along a guided ODE,

$$\frac{d}{dt}K_t(X_t) = -H_t(X_t) + (a_t - 1)E_t(X_t). \tag{24}$$

Online-K integrates the full density-ratio differential along the actual DDIM endpoints using trapezoidal quadrature, one input vector–Jacobian product (VJP) per endpoint to compute a Hutchinson estimate of the divergence (Hutchinson, 1989), and the time-coordinate correction. Stochastic Jacobian-trace estimation is also used for scalable likelihood evaluation in continuous normalizing flows (Grathwohl et al., 2019). Starting from $\widehat{K}_0 = 0$, Online-K completes the preceding interval's evidence update before setting $a_j = 1 + \lambda\,\mathrm{sigmoid}[\gamma(\widehat{K}_j - \kappa)]$. The same predictions are reused for sampling. Appendix J.4 gives the update equations and execution order; Appendix J.7 examines initialization, learned-score, quadrature, and trace-estimation errors. Corollary 5 applies to exact evidence, while the image experiments evaluate the estimated policy.

6 Experiments

We evaluate three aspects of our analysis: terminal-law corrections, budget-based schedule selection, and evidence-responsive guidance. Experiments use Gaussian mixtures, CIFAR-10 (Krizhevsky, 2009), and ImageNet (Deng et al., 2009; Russakovsky et al., 2015). The seed groups and scope of each comparison are specified below and in Appendices I and J.

6.1 GAUSSIAN MIXTURES: TERMINAL CORRECTIONS AND GUIDANCE TIMING

Gaussian mixtures with analytic scores allow us to test the terminal-law corrections without score-estimation error. For an eight-component, two-dimensional mixture at $w = 3$, the complete SDE law (15) yields a mode discrepancy of 0.002586 relative to direct reverse sampling, compared with 0.119934 when the initial boundary factors are omitted and 0.086362 for the endpoint-only tilt. The unnormalized estimate has total mass 1.0003 ± 0.0024 (SE). These results support agreement at $w = 1, 3$; uncertainty in the importance weights leaves the check at $w = 8$ inconclusive (Appendix G).

For a one-dimensional, three-component mixture, the ODE law (17) agrees with the flow-Jacobian density to within 3×10^{-11} in half-L^1 discrepancy. Because both density calculations use the same numerical characteristics, we also use an independent forward-quantile solver, which gives CDF errors below 4×10^{-6}. By comparison, the endpoint-only tilt has TV error 0.28035 at $w = 3$. Omitting $\dot{a}K$ for a power ramp yields raw mass 1.6738 and normalized TV error 0.08687 (Appendix H).

We next test whether coefficient budgets alone determine the terminal law. Early and late pulses have identical initial, terminal, peak, and mean scales, the same scale histogram, and a shared budget $B_{\mathrm{coef}} = 5/8$. Nevertheless, their terminal laws differ by TV 0.174187; spatial refinement changes this value by only 2.8×10^{-6} (Figure H.1). Thus, budgets provide interpretable search coordinates, while guidance timing remains a separate design choice.

6.2 BUDGET COORDINATES IMPROVE PCS SELECTION

To evaluate budget coordinates for PCS selection, we compare a conventional mean–shape grid (m, s) with our peak–budget grid (W, B_{200}). Each grid contains 12 candidates, with a shared anchor yielding 23 distinct development configurations. Minimizing development FID, either with or without a 98.2% development-accuracy requirement, selects the same winner within each grid. Both selected configurations and the original PCS exceed 98% accuracy in every final seed.

Table 1: PCS selection on CIFAR-10/DDIM-200. FID is mean ± sample SD over three paired seeds, with 50,000 images per configuration and seed. Each search evaluates 12 candidates.

PCS configuration	FID ↓
Original PCS	2.6334 ± 0.0096
Conventional-grid winner	2.6232 ± 0.0116
Budget-grid winner (ours)	$\mathbf{2.5993} \pm 0.0110$

Appendix I gives complete metrics and search definitions.

Budget-based selection reduces FID by 0.913% relative to conventional selection and by 1.298% relative to the original PCS, with improvements in all three paired seeds (Table 1). These coordinates therefore identify a lower-FID configuration with the same number of candidates. The selected schedule increases both peak and budget by 31.17% over the anchor and changes the guidance timing (Figure I.2). Because the two grids vary timing and strength jointly and cover different candidate domains, this gain is specific to selection within these matched-size grids.

6.3 ONLINE-K AND BASELINE GUIDANCE METHODS

We compare Online-K with fixed CFG, Wang's scheduler family (Wang et al., 2024), and VP/DDIM adaptations of β-CFG and Feedback Guidance (FBG) (Malarz et al., 2025; Koulischer et al., 2025). Table 2 reports one configuration per family, selected by minimum development FID. This criterion selects PCS from Wang's family and Online-K with $(\lambda, \gamma, \kappa) = (2, 4, 4)$ and one trace probe. Each row aggregates all three final seeds for the selected configuration. Appendix J.3 provides the baseline parameters and selection details.

Online-K achieves the lowest reported mean FID in Table 2, 1.65% below PCS. Because these two entries differ in numerical precision and evaluation seeds, we also conduct the paired evaluation below.

In this paired evaluation, we rerun the same PCS configuration, $(m, s) = (1.3, 4)$, using the same FP32 arithmetic, initial noise, labels, and clipping rule as Online-K. PCS obtains FID 2.6114 ± 0.0141.

Table 2: Best-FID configurations for five guidance families on CIFAR-10/DDIM-200.

Method	FID ↓	Precision ↑	Recall ↑
Online-K (ours)	$\mathbf{2.5617} \pm 0.0140$	0.6861 ± 0.0016	0.5780 ± 0.0002
Wang PCS	2.6048 ± 0.0275	0.6873 ± 0.0033	$\mathbf{0.5784} \pm 0.0022$
Fixed CFG	2.9474 ± 0.0311	0.6844 ± 0.0022	0.5778 ± 0.0016
β-CFG†	3.0866 ± 0.0348	0.6950 ± 0.0028	0.5648 ± 0.0048
FBG†	4.5264 ± 0.0209	$\mathbf{0.7301} \pm 0.0010$	0.5154 ± 0.0019

†VP/DDIM adaptation.

Online-K reduces this mean by 1.90%, with a paired difference of -0.0497 ± 0.0033 and lower FID in all three seeds. PCS remains fixed during this development round, so the search budgets are not matched. This PCS reference also differs from the budget-selected schedule in Section 6.2.

Classifier agreement and feature precision/recall provide secondary diagnostics. Appendix J.3 reports these metrics for the four baseline configurations; Appendices J.5 and J.9 report the paired evaluation and the additional Online-K parameter setting, respectively. All methods in the CIFAR experiment use 200 joint conditional/unconditional forward calls per trajectory; Online-K additionally requires 200 input VJPs and 199 evidence updates. These comparisons evaluate the complete policies and do not isolate the effect of feedback from guidance strength and timing.

On ImageNet 256×256, we evaluate pretrained DiT-XL/2 (Peebles & Xie, 2023) with DDIM-50. Across three seeds, the DiT Online-K variant, which uses a batch-shared expected-energy proxy, achieves a mean FID of 3.508, a 2.98% reduction relative to PCS (Table 3). This comparison evaluates complete guidance profiles, in which feedback and time allocation remain coupled. Appendix K provides the parameters, fixed-scale screening results, and generated samples.

Table 3: ImageNet DiT-XL/2, DDIM-50 steps, 50K images per method. All three methods use seeds 0, 1, and 2.

Method	FID ↓	Precision ↑	Recall ↑
Fixed CFG ($w = 1.5$)	3.7376 ± 0.0442	$\mathbf{0.7101 \pm 0.0018}$	0.5887 ± 0.0009
PCS	3.6157 ± 0.0675	0.7036 ± 0.0016	$\mathbf{0.6172 \pm 0.0008}$
Online-K (energy proxy)	$\mathbf{3.5079 \pm 0.0149}$	0.6802 ± 0.0008	0.6117 ± 0.0003

7 Conclusion and Limitations

We characterize prescribed CFG policies through inverse-control objectives under specified quadratic effort metrics and terminal laws that depend on the sampler. These laws expose initialization and trajectory corrections to the local power tilt, connecting stochastic path averaging with deterministic transport through a zero-noise limit under stated regularity conditions. Time–evidence potentials motivate coefficient-budget coordinates for PCS selection and Online-K. Gaussian-mixture experiments support the density identities and show that guidance timing matters even when coefficient budgets match. On CIFAR-10, budget-based selection improves PCS within the tested grids, while Online-K reduces mean FID relative to a fixed PCS reference across three paired seeds. ImageNet extends the evaluation to a batch-shared energy variant.

The guarantees require exact scores, regularity, and integrability; learned-score and finite-step errors remain outside their scope. High-guidance stochastic validation remains inconclusive. The image comparisons couple feedback, strength, and timing: PCS searches use different candidate domains, and the paired CIFAR-10 comparison holds PCS fixed during development. Online-K incurs VJP cost and evidence-estimation error. ImageNet has one paired seed, with lower precision and recall despite improved FID; fixed guidance uses a separate VAE protocol. Extending the theory to learned scores and improving the accuracy and computational cost of evidence estimation remain future directions.

AI USE STATEMENT

Generative AI assisted with manuscript organization, English drafting and editing, reference checking, and LaTeX preparation. It also supported sampling-code inspection, consistency checks between tables and experiment records, and compilation and layout checks. The authors remain responsible for the paper's claims, proofs, references, and reported results.

REPRODUCIBILITY STATEMENT

Appendices A–F state the assumptions and derivations, including finite-step identities and the zero-noise limit. Appendices G and H give the analytic-mixture protocols. Appendix I specifies the PCS search grids, inverse budget map, selected configurations, and nine final runs. Appendix J gives the baseline guidance rules and twelve baseline runs, together with the Online-K estimator, numerical diagnostics, selection rules, and nine runs in the paired FP32 comparison. Appendix K records the ImageNet protocol and available metrics. The accompanying source includes the CIFAR-10 Online-K/PCS sampling core, baseline policy implementations, CPU estimator checks, scalar metric records, editable result tables, and reporting scripts, plus the supplied ImageNet result summary and sample grid. Model weights, full generated-image collections, and feature caches are not included; the qualitative subsets contain 60 CIFAR-10 images and 16 ImageNet images. Reporting checks reconstruct summaries from saved metrics and do not recompute image-level FID.

ETHICS STATEMENT

The experiments concern synthetic Gaussian mixtures and conditional image generation on CIFAR-10 and ImageNet. The evaluation does not assess deployment risks such as harmful content or representational bias in conditional generation.

REFERENCES

Brian D. O. Anderson. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982.

Iskander Azangulov, Peter Potaptchik, Qinyu Li, Eddie Aamari, George Deligiannidis, and Judith Rousseau. Adaptive diffusion guidance via stochastic optimal control. In *Proceedings of The 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 4087–4095. PMLR, 2026. URL https://proceedings.mlr.press/v300/azangulov26a.html.

Julius Berner, Lorenz Richter, and Karen Ullrich. An optimal control perspective on diffusion-based generative modeling. *Transactions on Machine Learning Research*, 2024. URL https://openreview.net/forum?id=oYIjw37pTP.

Michelle Boué and Paul Dupuis. A variational representation for certain functionals of Brownian motion. *The Annals of Probability*, 26(4):1641–1659, 1998. doi: 10.1214/aop/1022855876. URL https://doi.org/10.1214/aop/1022855876.

Arwen Bradley and Preetum Nakkiran. Classifier-free guidance is a predictor-corrector. *Transactions on Machine Learning Research*, 2025. URL https://arxiv.org/abs/2408.09000.

Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David K. Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, volume 31, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/69386f6bb1dfed68692a24c8686939b9-Abstract.html.

Muthu Chidambaram, Khashayar Gatmiry, Sitan Chen, Holden Lee, and Jianfeng Lu. What does guidance do? a fine-grained analysis in a simple setting. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/9a3942c235daa9c5f62a8598ae81a946-Abstract-Conference.html.

Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion Schrödinger bridge with applications to score-based generative modeling. In *Advances in Neural Information Processing Systems*, volume 34, pp. 17695–17709, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/940392f5f32a7ade1cc201767cf83e31-Abstract.html.

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255. IEEE, 2009. URL https://image-net.org/static_files/papers/imagenet_cvpr09.pdf.

Yilun Du, Conor Durkan, Robin Strudel, Joshua B. Tenenbaum, Sander Dieleman, Rob Fergus, Jascha Sohl-Dickstein, Arnaud Doucet, and Will Sussman Grathwohl. Reduce, reuse, recycle: Compositional generation with energy-based diffusion models and MCMC. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 8489–8510. PMLR, 2023. URL https://proceedings.mlr.press/v202/du23a.html.

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

Will Grathwohl, Ricky T. Q. Chen, Jesse Bettencourt, Ilya Sutskever, and David Duvenaud. FFJORD: Free-form continuous dynamics for scalable reversible generative models. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=rJxgknCcK7.

Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In *Advances in Neural Information Processing Systems*, 2017. URL https://arxiv.org/abs/1706.08500.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022. URL https://arxiv.org/abs/2207.12598.

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

M. F. Hutchinson. A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. *Communications in Statistics—Simulation and Computation*, 18(3):1059–1076, 1989. doi: 10.1080/03610918908812806. URL https://doi.org/10.1080/03610918908812806.

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

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

R. E. Kalman. When is a linear control system optimal? *Journal of Basic Engineering*, 86(1):51–60, 1964. doi: 10.1115/1.3653115. URL https://doi.org/10.1115/1.3653115.

Hilbert J. Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of Statistical Mechanics: Theory and Experiment*, pp. P11011, 2005.

Felix Koulischer, Florian Handke, Johannes Deleu, Thomas Demeester, and Luca Ambrogioni. Feedback guidance of diffusion models. In *Advances in Neural Information Processing Systems*, volume 38, 2025. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/73c90c0c16fe835bc58a815d4a8f28be-Abstract-Conference.html.

Alex Krizhevsky. Learning multiple layers of features from tiny images. Technical report, University of Toronto, 2009. URL https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf.

Tuomas Kynkäänniemi, Tero Karras, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Improved precision and recall metric for assessing generative models. In *Advances in Neural Information Processing Systems*, 2019. URL https://arxiv.org/abs/1904.06991.

Tuomas Kynkäänniemi, Miika Aittala, Tero Karras, Samuli Laine, Timo Aila, and Jaakko Lehtinen. Applying guidance in a limited interval improves sample and distribution quality in diffusion models. In *Advances in Neural Information Processing Systems*, 2024. URL https://arxiv.org/abs/2404.07724.

Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete and Continuous Dynamical Systems A*, 34(4):1533–1574, 2014.

Shanchuan Lin, Bingchen Liu, Jiashi Li, and Xiao Yang. Common diffusion noise schedules and sample steps are flawed. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 5404–5411, 2024. URL https://openaccess.thecvf.com/content/WACV2024/html/Lin_Common_Diffusion_Noise_Schedules_and_Sample_Steps_Are_Flawed_WACV_2024_paper.html.

Dawid Malarz, Artur Kasymov, Maciej Zięba, Jacek Tabor, and Przemysław Spurek. Classifier-free guidance with adaptive scaling. *arXiv preprint arXiv:2502.10574*, 2025. URL https://arxiv.org/abs/2502.10574.

Pinelopi Papalampidi, Olivia Wiles, Ira Ktena, Aleksandar Shtedritski, Emanuele Bugliarello, Ivana Kajić, Isabela Albuquerque, and Aida Nematzadeh. Dynamic classifier-free diffusion guidance via online feedback. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/cef8b22f26953d8bdaf93bb64b7cc72f-Abstract-Conference.html.

Krunoslav Lehman Pavasovic, Jakob Verbeek, Giulio Biroli, and Marc Mezard. Overshoot and shrinkage in classifier-free guidance: From theory to practice. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/d7168ad82b4780797ffa8202c531ef6a-Abstract-Conference.html.

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4195–4205, 2023. URL https://openaccess.thecvf.com/content/ICCV2023/html/Peebles_Scalable_Diffusion_Models_with_Transformers_ICCV_2023_paper.html.

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*, pp. 10684–10695, 2022. URL https://openaccess.thecvf.com/content/CVPR2022/html/Rombach_High-Resolution_Image_Synthesis_With_Latent_Diffusion_Models_CVPR_2022_paper.html.

Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 115 (3):211–252, 2015. doi: 10.1007/s11263-015-0816-y. URL https://doi.org/10.1007/s11263-015-0816-y.

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 *International Conference on Machine Learning*, 2025. URL https://arxiv.org/abs/2503.02819.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International Conference on Machine Learning*, 2015. URL https://arxiv.org/abs/1503.03585.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021a. URL https://arxiv.org/abs/2010.02502.

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*, 2021b. URL https://arxiv.org/abs/2011.13456.

Xi Wang, Nicolas Dufour, Nefeli Andreou, Marie-Paule Cani, Victoria Fernandez Abrevaya, David Picard, and Vicky Kalogeiton. Analysis of classifier-free guidance weight schedulers. *arXiv preprint arXiv:2404.13040*, 2024. URL https://arxiv.org/abs/2404.13040.

Yuchen Wu, Minshuo Chen, Zihao Li, Mengdi Wang, and Yuting Wei. Theoretical insights for diffusion guidance: A case study for Gaussian mixture models. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 53291–53327. PMLR, 2024. URL https://proceedings.mlr.press/v235/wu24b.html.

Qinsheng Zhang and Yongxin Chen. Path integral sampler: A stochastic control approach for sampling. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=_uCb2ynRu7Y.

Candi Zheng and Yuan Lan. Characteristic guidance: Non-linear correction for diffusion model at large guidance scale. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 61386–61412. PMLR, 2024. URL https://proceedings.mlr.press/v235/zheng24f.html.

A SAMPLER CORRESPONDENCE AND POPULATION EVIDENCE

The fixed-guidance endpoints of Eq. (5) are

$$\begin{aligned} \text{SDE:} \quad & dX_t = [b_t + wA_t\nabla K_t](X_t)\,dt + \sigma_t\,dW_t, \\ \text{ODE:} \quad & dX_t = [v_t + wM_t\nabla K_t](X_t)\,dt. \end{aligned} \tag{A.1}$$

Both start from the prescribed density ν.

This appendix fixes the finite-step convention and derives the population identities used in the main text. Sampling time increases from noise to data. Throughout, $M_t = A_t/2$, $A_t = \sigma_t\sigma_t^\top$ is state independent, $s_t = \nabla \log p_t$, and

$$K_t = \log(p_t^c/p_t), \qquad r = K_T, \qquad E_t = \nabla K_t^\top M_t \nabla K_t, \qquad H_t = \frac{\operatorname{div}(p_t M_t \nabla K_t)}{p_t}.$$

The initial law of the guided process is denoted by ν. It is not assumed to equal either population marginal unless stated explicitly.

A.1 THE FINITE-GRID DDPM–DDIM RELATION

Let $k = 0, \dots, N$ index a sampling grid with $0 < \bar\alpha_0 < \cdots < \bar\alpha_N \le 1$. For $\epsilon_k^a = \epsilon_k^u + a_k(\epsilon_k^c - \epsilon_k^u)$, the unclipped DDIM-family update is

$$X_{k+1} = m_k^{\eta,a}(X_k) + \sigma_{k,\eta} Z_k, \qquad Z_k \sim \mathcal{N}(0, I),$$

$$\sigma_{k,\eta}^2 = \eta^2 \frac{1-\bar\alpha_{k+1}}{1-\bar\alpha_k}\left(1 - \frac{\bar\alpha_k}{\bar\alpha_{k+1}}\right), \tag{A.2}$$

$$m_k^{\eta,a}(x) = \sqrt{\frac{\bar\alpha_{k+1}}{\bar\alpha_k}}\,x + \left[\sqrt{1-\bar\alpha_{k+1}-\sigma_{k,\eta}^2} - \sqrt{\frac{\bar\alpha_{k+1}(1-\bar\alpha_k)}{\bar\alpha_k}}\right]\epsilon_k^a(x). \tag{A.3}$$

The parameter η determines both the mean and variance: $\eta = 0$ gives deterministic DDIM, while $\eta = 1$ gives the posterior-variance DDPM update derived below.

For conventional noising indices $j < i$, write $\alpha_- = \bar\alpha_i$, $\alpha_+ = \bar\alpha_j$, and $\chi = \alpha_-/\alpha_+$. The forward transition is $X_i \mid X_j \sim \mathcal{N}(\sqrt{\chi}X_j, (1-\chi)I)$. Conditioning additionally on the clean state $X_0 = x_0$ gives

$$q(x_j \mid x_i, x_0) = \mathcal{N}\left(\frac{\sqrt{\alpha_+}(1-\chi)}{1-\alpha_-}x_0 + \frac{\sqrt{\chi}(1-\alpha_+)}{1-\alpha_-}x_i,\ \frac{(1-\alpha_+)(1-\chi)}{1-\alpha_-}I\right). \tag{A.4}$$

Substitute $\widehat{x}_0 = (x_i - \sqrt{1-\alpha_-}\,\epsilon_i^a)/\sqrt{\alpha_-}$ into the DDIM mean. With $\widetilde\beta = (1-\alpha_+)(1-\chi)/(1-\alpha_-)$, the identity $1 - \alpha_+ - \widetilde\beta = \chi(1-\alpha_+)^2/(1-\alpha_-)$ yields, at $\eta = 1$,

$$m_i^{1,a}(x) = \chi^{-1/2}\left[x - \frac{1-\chi}{\sqrt{1-\alpha_-}}\epsilon_i^a(x)\right], \qquad \sigma_{i,1}^2 = \widetilde\beta. \tag{A.5}$$

Thus $\eta = 1$ recovers the posterior-variance DDPM kernel on the chosen grid. This correspondence uses the direct transition between the selected noise levels. Appendix E gives the finite-step terminal and control identities for deterministic maps and stochastic kernels.

A.2 A SUFFICIENT CONTINUOUS-LIMIT SETTING

Consider a truncated VP interval with $\delta \le \bar\alpha(t) \le 1 - \delta$ for some $\delta > 0$, a smooth positive rate $\beta(t)$, and $\bar\alpha(t+h) = \bar\alpha(t)(1 + \beta(t)h) + O(h^2)$. Exact population predictors satisfy

$$\epsilon^u(t,x) = -\sqrt{1-\bar\alpha(t)}\,s_t(x), \qquad \epsilon^c(t,x) - \epsilon^u(t,x) = -\sqrt{1-\bar\alpha(t)}\,\nabla K_t(x).$$

Expanding (A.2)–(A.3) gives

$$\mathbb{E}[\Delta X \mid X_t = x] = h\left[\frac{\beta(t)}{2}x + \frac{1+\eta^2}{2}\beta(t)\{s_t(x) + a(t,x)\nabla K_t(x)\}\right] + O(h^2), \tag{A.6}$$

$$\operatorname{Cov}(\Delta X \mid X_t = x) = \eta^2\beta(t)hI + O(h^2). \tag{A.7}$$

The remainders are locally uniform in x. Thus the limiting drift and covariance are precisely those in (5).

For completeness, a sufficient convergence argument assumes that the guided drift is globally Lipschitz in x, has linear growth uniformly in time, and has enough time regularity for the local expansions above; the initial laws coincide and have finite second moment. Assume also uniform linear-growth bounds for the finite-step drift increments. These bounds give uniform second moments and compact containment. After stopping on compact sets, the drift/martingale decomposition gives the stopping-time increment estimate $\mathbb{E}\|X_{\tau+\theta} - X_\tau\|^2 \le C(\theta + h)$, up to the stopping boundary, which supplies the required time-tightness criterion. The localized Gaussian increments also imply that the maximum jump tends to zero. Generator convergence for smooth compactly supported test functions then identifies every subsequential limit with the limiting martingale problem; uniqueness gives convergence in law. At $\eta = 0$, consistency and Lipschitz stability yield the ODE limit, with first-order trajectory accuracy when the local deterministic error is uniformly $O(h^2)$. These are sufficient assumptions for the analytic limit, not verified properties of an arbitrary neural predictor. Reaching zero signal or a singular clean endpoint requires an additional justified limit.

A.3 Evidence identities for the sampler family

Section 2.2 collects the stochastic evidence and control foundations. We record their form under our covariance convention, then extend the evidence equation to the full sampler family.

Let $q_{t|T}(x \mid y)$ be the forward corruption density from clean y to the noise level indexed by sampling time t. For the unconditional reverse reference P started from p_0, Bayes' rule gives

$$\mathbb{E}_P[e^{r(X_T)} \mid X_t = x] = \int \frac{p_T^c(y)}{p_T(y)} \frac{q_{t|T}(x \mid y) p_T(y)}{p_t(x)}\, dy = \frac{p_t^c(x)}{p_t(x)} = e^{K_t(x)}. \tag{A.8}$$

Let $L_t = b_t \cdot \nabla + M_t : \nabla^2$, where $b_t = v_t + M_t s_t$. The backward equation for e^K implies

$$(\partial_t + L_t)K_t = -E_t, \qquad (\partial_t + v_t \cdot \nabla)K_t = -H_t - E_t. \tag{A.9}$$

Indeed, $H = M : \nabla^2 K + (Ms) \cdot \nabla K$. Hence, with $L_t^\eta = (v_t + \eta^2 M_t s_t) \cdot \nabla + \eta^2 M_t : \nabla^2$,

$$(\partial_t + L_t^\eta)K_t = -(1 - \eta^2)H_t - E_t. \tag{A.10}$$

Itô's formula along the guided family gives the exact evidence dynamics

$$\begin{aligned} dK_t(X_t) = {} & \left[-(1-\eta^2)H_t + \{(1+\eta^2)a(t,\cdot) - 1\}E_t\right](X_t)\, dt \\ & + \eta \nabla K_t(X_t)^\top \sigma_t\, dW_t. \end{aligned} \tag{A.11}$$

The Doob identity (A.8) is specific to the unconditional reverse-SDE reference. Sharing population marginals does not give the same backward conditional moment for every η.

For a discrete class condition, $K_t(x) = \log p(c \mid x, t) - \log p(c) \le -\log p(c)$. Thus evidence thresholds have a fixed calibration convention and a feasible upper range. Replacing K by $\log p(c \mid x, t)$ shifts thresholds by the same class-dependent constant. A finite-noise K_0 need not be constant.

B Control verification and general stochasticity

This appendix proves the inverse-control results in Section 3 and extends them to the full sampler family (5).

B.1 Assumptions and admissible controls

Let $b_t^\eta = v_t + \eta^2 M_t s_t$. Denote by $Q^{\eta,u}$ the path law of $dX_t = (b_t^\eta(X_t) + u_t)\, dt + \eta \sigma_t\, dW_t$ with $X_0 \sim \nu$, and by $P^{\eta,\nu}$ the law with $u \equiv 0$. At $\eta = 1$, write $Q^u = Q^{1,u}$ and $P = P^{1,\nu}$.

Work on a finite interval with continuous uniformly positive definite M_t, positive $C^{1,2}$ population densities, and unique nonexplosive reference dynamics. Candidate potentials F are $C^{1,2}$. Candidate feedbacks are locally Lipschitz and induce nonexplosive dynamics. An admissible progressively

measurable control induces a well-posed controlled process, has finite expected quadratic effort, and has integrable terminal and running terms. The stochastic integral used below must be a true zero-mean martingale; a sufficient condition is

$$\mathbb{E}_{Q^{\eta,u}} \int_t^T \|\nabla F_s(X_s)\|^2_{A_s}\, ds < \infty. \tag{B.1}$$

For the fixed initial density ν, require $\mathbb{E}_\nu |F_0| < \infty$. The maximizing feedback must itself be admissible. These conditions can be imposed for each starting pair (t, x) to obtain the pointwise value, or integrated against a prescribed initial density ν.

Path-likelihood identities additionally require absolute continuity and a true Girsanov exponential conditional on the initial state. Comparing with the reference started from p_0 requires $\nu \ll p_0$. For $\eta > 0$, write $A^\eta_t = 2\eta^2 M_t$; a sufficient Girsanov condition is

$$\mathbb{E}_{P^{\eta,\nu}} \exp\left\{ \frac{1}{2} \int_0^T \|u_t\|^2_{(A^\eta_t)^{-1}}\, dt \right\} < \infty. \tag{B.2}$$

Density representations require classical solutions, uniqueness in the chosen solution class, and the Feynman–Kac integrability conditions specified below. Singular clean distributions are handled only through truncation followed by a justified limit.

B.2 A GENERAL POTENTIAL AND ITS VERIFICATION

Set $G^\eta_t = (1+\eta^2)M_t$ and $L^\eta_t = b^\eta_t \cdot \nabla + \eta^2 M_t : \nabla^2$.

Theorem 6 (General inverse-control identity)**.** *Under the preceding assumptions, let F be a $C^{1,2}$ potential and define*

$$\begin{aligned} R^\eta_F &= (\partial_t + L^\eta_t)F + \tfrac{1}{2}\|\nabla F\|^2_{G^\eta_t}, \\ J^{\eta,F}_{t,x}(u) &= \mathbb{E}^{\eta,u}_{t,x}\left[F_T(X_T) - \int_t^T \left\{ R^\eta_F(s, X_s) + \tfrac{1}{2}\|u_s\|^2_{(G^\eta_s)^{-1}} \right\} ds \right]. \end{aligned} \tag{B.3}$$

For every admissible control,

$$F_t(x) - J^{\eta,F}_{t,x}(u) = \tfrac{1}{2}\mathbb{E}^{\eta,u}_{t,x} \int_t^T \|u_s - G^\eta_s \nabla F_s(X_s)\|^2_{(G^\eta_s)^{-1}}\, ds. \tag{B.4}$$

Hence the value is $F_t(x)$, attained by $u^*_F = G^\eta \nabla F$. The statement applies to each starting pair (t, x) for which the assumptions hold, or after integration against the prescribed initial density ν.

Proof. Under $dX = (b^\eta + u)dt + \eta\sigma dW$, Itô's formula gives

$$dF_t(X_t) = \{(\partial_t + L^\eta_t)F_t + u_t \cdot \nabla F_t\}(X_t)dt + \eta \nabla F_t(X_t)^\top \sigma_t\, dW_t.$$

Integrate, subtract the running term and effort, and take expectations. Completion of the square yields

$$\begin{aligned} &\mathbb{E}^{\eta,u}_{t,x}\left[F_T(X_T) - \int_t^T \left\{ R^\eta_F(s, X_s) + \tfrac{1}{2}\|u_s\|^2_{(G^\eta_s)^{-1}} \right\} ds \right] \\ &\qquad = F_t(x) - \frac{1}{2}\mathbb{E}^{\eta,u}_{t,x} \int_t^T \|u_s - G^\eta_s \nabla F_s(X_s)\|^2_{(G^\eta_s)^{-1}}\, ds. \end{aligned} \tag{B.5}$$

This proves the value and the feedback $u^*_F = G^\eta \nabla F$. The ordinary chain rule proves the same identity at $\eta = 0$. An admissible control attains the value precisely when it equals u^*_F $dt \otimes Q^{\eta,u}$-almost everywhere; uniqueness of the closed-loop dynamics then gives uniqueness of the optimal law. No uniqueness is asserted for control values at states the process never visits.

Using (A.10), the fixed-CFG choice $F = wK$ has

$$R^\eta_w = -(1-\eta^2)wH + \left[\frac{1+\eta^2}{2} w^2 - w \right] E, \qquad u^*_w = w(1+\eta^2)M\nabla K. \tag{B.6}$$

This is inverse optimality for the specified running term and effort metric. It does not uniquely determine an objective independently of the policy.

□

Specialization to the main-text results. For $F = wK$ and $\eta = 1$, $G^1 = A$ and $R^1_w = w(w-1)E$. For $\eta = 0$, $G^0 = M$ and $R^0_w = -wH + (w^2/2 - w)E$. Taking $t = 0$ and integrating the general identity against ν proves Theorem 1 and Corollary 2, including their gap identities.

B.3 The reward-only benchmark and path KL

We use the entropy-regularized control foundation summarized in Section 2.2. The identities below fix the unconditional reference and preserve the prescribed initial law ν.

For the reverse SDE, let $h^w_t(x) = \mathbb{E}_P[e^{wr(X_T)} \mid X_t = x]$ and $K^w_t = \log h^w_t$. Whenever these quantities and the associated laws satisfy the preceding conditions,

$$\frac{dQ^{*,w}}{dP^{1,\nu}} = \frac{e^{wr(X_T)}}{h^w_0(X_0)}$$

preserves the fixed initial density ν. For any admissible Q with that initial law,

$$w\mathbb{E}_Q r(X_T) - D_{\mathrm{KL}}(Q\|P^{1,\nu}) = \mathbb{E}_\nu K^w_0(X_0) - D_{\mathrm{KL}}(Q\|Q^{*,w}). \tag{B.7}$$

The corresponding HJB equation is $(\partial_t + L_t)K^w + \frac{1}{2}\|\nabla K^w\|^2_{A_t} = 0$, with terminal value wr, and optimal control $A\nabla K^w$. Equation (A.8) gives $K^1 = K$; in general, $K^w \neq wK$.

For $\eta > 0$, with matching initial laws, canonical state-filtration controls, true Girsanov likelihoods, and finite expected effort,

$$\frac{1}{2}\mathbb{E}_{Q^{\eta,u}}\int_0^T \|u_t\|^2_{(G^\eta_t)^{-1}}dt = \frac{1}{c_\eta}D_{\mathrm{KL}}(Q^{\eta,u}\|P^{\eta,\nu}), \qquad c_\eta = \frac{1+\eta^2}{2\eta^2}. \tag{B.8}$$

Thus the chosen effort equals path KL at $\eta = 1$. At intermediate stochasticity, multiplying the entire control objective by c_η gives unit-weight KL, terminal utility $c_\eta F_T$, and running term $c_\eta R^\eta_F$. For $F = \phi(t, K)$ and $g = \phi_z$, the optimal instantaneous KL effort is $\frac{(1+\eta^2)^2}{4\eta^2}g^2E$. Its coefficient diverges as $\eta \downarrow 0$; divergence of the integrated expected effort additionally requires that its path-averaged g^2E not vanish in the limit. If initial laws differ, full path KL includes $D_{\mathrm{KL}}(\nu\|p_0)$.

At $\eta = 0$, let $\mathcal{G}_0$ be the measurable set of paths satisfying the reference ODE's integral equation. The reference law assigns it probability one. A controlled law assigning positive mass to $\mathcal{G}^c_0$ is not absolutely continuous with respect to the reference and has infinite path KL. If trajectories differ for almost every controlled initial state and the reference flow is unique, the path laws are mutually singular. Endpoint densities may still overlap: the complete path includes its initial coordinate and its dynamics. In contrast, (B.5) remains a valid finite-effort identity with metric M^{-1}. For reward-only deterministic control, the equation is $(\partial_t + v\cdot\nabla)U + \frac{1}{2}\|\nabla U\|^2_M = 0$. The terminal reward transported along the reference flow solves linear transport and need not solve this Hamilton–Jacobi equation.

Combining Theorem 1 with (B.8) gives

$$\sup_u \left\{ w\mathbb{E}_{Q^u} r(X_T) - w(w-1)\mathbb{E}_{Q^u}\int_0^T E_t(X_t)\,dt - D_{\mathrm{KL}}(Q^u\|P^{1,\nu}) \right\} = w\mathbb{E}_\nu K_0. \tag{B.9}$$

Here the path laws share the same initial density ν, and the additional conditions for the effort–KL identity are required.

C Terminal-density proofs and the zero-noise limit

We first derive the reverse-SDE terminal law from its path likelihood, then obtain the general-density representation and its deterministic limit.

C.1 Reverse-SDE likelihood and terminal conditioning

Throughout the terminal-law derivations below, let $P = P^{1,p_0}$ be the unconditional reverse-SDE reference started from p_0, so its marginals are p_t. This differs from the reference $P^{1,\nu}$ used for the same-initialization control comparison. Let Q^F start from $\nu \ll p_0$ and have additional drift

$u_F = A\nabla F$. Assume the absolute-continuity and conditional Girsanov conditions in Appendix B.1. Then

$$\frac{dQ^F}{dP} = \frac{\nu(X_0)}{p_0(X_0)} \exp\left\{ \int_0^T \nabla F_t(X_t)^\top \sigma_t \, dW_t - \frac{1}{2} \int_0^T \|\nabla F_t(X_t)\|^2_{A_t} dt \right\}.$$

Applying Itô's formula under P to the stochastic integral gives

$$\frac{dQ^F}{dP} = \frac{\nu(X_0)}{p_0(X_0)} \exp\left\{ F_T(X_T) - F_0(X_0) - \int_0^T R^1_F(t, X_t) dt \right\}, \tag{C.1}$$

where R^1_F is the control residual in (B.3). Conditioning on $X_T = y$ yields, for p_T-almost every y,

$$\rho^F_T(y) = p_T(y) e^{F_T(y)} \, \mathbb{E}_P \left[\frac{\nu(X_0)}{p_0(X_0)} e^{-F_0(X_0) - \int_0^T R^1_F(t, X_t) dt} \,\middle|\, X_T = y \right]. \tag{C.2}$$

This identity follows by integrating against arbitrary bounded terminal test functions. Its total integral is one because it is the density of Q^F_T, so no additional normalizer is required. With $F = wK$, $R^1_w = w(w-1)E$ and the initial factor is C^w_0 from (14). This proves Theorem 3 and (15). Both the initial-density ratio and the initial-potential factor remain inside the bridge expectation.

C.2 A TERMINAL-DENSITY REPRESENTATION FOR GENERAL POTENTIALS

Let $F_t(x) = \phi(t, K_t(x))$, with ϕ sufficiently smooth that F is $C^{1,2}$, and set $g_t(x) = \phi_z(t, K_t(x))$. For this subsection, write $\rho^{\eta,\phi}_t$ for the density in (5) with $a = g$ and initial density ν. Define

$$R^{\mathrm{tr}}_\phi(t, x) = \phi_t + (\phi_{zz} + \phi_z^2 - \phi_z) E_t(x), \tag{C.3}$$

where the derivatives of ϕ are evaluated at $(t, K_t(x))$. This transport residual is independent of η, unlike the control residual R^η_F.

Theorem 7 (General terminal correction and zero-noise limit)**.** *Let $\eta \in [0,1]$ and let $C_0 = \nu p_0^{-1} e^{-F_0}$. Assume the population regularity stated in Appendix B.1, classical density solutions, uniqueness in the chosen solution class, nonexplosive auxiliary dynamics, and the Feynman–Kac integrability conditions detailed in the proof below. From $Y_0 = y$, run*

$$\begin{aligned} dY_s &= \left[-v_t + \eta^2 M_t s_t - (1-\eta^2) g_t M_t \nabla K_t \right](Y_s) \, ds + \eta \sigma_t \, dB_s, \\ t &= T - s. \end{aligned} \tag{C.4}$$

Then

$$\rho^{\eta,\phi}_T(y) = p_T(y) e^{\phi(T, r(y))} \mathbb{E}_y \left[C_0(Y_T) \exp\left\{ - \int_0^T R^{\mathrm{tr}}_\phi(T-s, Y_s) \, ds \right\} \right]. \tag{C.5}$$

At $\eta = 1$, the auxiliary process is the unconditional reference P conditioned on $X_T = y$, read backward. At $\eta = 0$, it is the inverse guided characteristic. Writing $x^y_T = y$ and $\dot{x}^y_t = (v_t + g_t M_t \nabla K_t)(x^y_t)$, its nonexplosive inverse flow on $[0,T]$ gives

$$\rho^{0,\phi}_T(y) = p_T(y) e^{\phi(T, r(y))} C_0(x^y_0) \exp\left\{ - \int_0^T R^{\mathrm{tr}}_\phi(t, x^y_t) \, dt \right\}. \tag{C.6}$$

For fixed ν and ϕ, the expectation factor in (C.5) converges pointwise to the characteristic factor in (C.6) as $\eta \downarrow 0$ if the auxiliary drifts are uniformly Lipschitz with linear growth, differ from the limiting drift by $O(\eta^2)(1 + \|x\|)$, σ_t is bounded, and $C_0, R^{\mathrm{tr}}_\phi$ are bounded and continuous. Localization with uniform integrability extends the limit to unbounded data, as detailed in Appendix C.3.

Proof of the density representations. Write $g = \phi_z(t, K)$, $j = \phi_{zz}(t, K)$, and $u = (1+\eta^2) g M \nabla K$. Let $R = \rho / p$ temporarily denote the density ratio. Subtracting the reference Fokker–Planck equation from the guided one gives

$$\partial_t R = \eta^2 M : \nabla^2 R + [-v + \eta^2 M s - u] \cdot \nabla R - (1+\eta^2)(gH + jE) R. \tag{C.7}$$

Set $R = e^F C$. The coefficient of ∇C is

$$D^{\eta,\phi} = -v + \eta^2 M s - (1-\eta^2) g M \nabla K. \tag{C.8}$$

Using $\nabla^2 F = g\nabla^2 K + j\nabla K \nabla K^\top$ and (A.9), the zeroth-order coefficient reduces to $-\phi_t - (j + g^2 - g)E = -R_\phi^{\mathrm{tr}}$; the H terms cancel. Thus

$$\partial_t C = \eta^2 M : \nabla^2 C + D^{\eta,\phi} \cdot \nabla C - R_\phi^{\mathrm{tr}} C, \qquad C_0 = \frac{\nu}{p_0} e^{-F_0}. \tag{C.9}$$

The standard Feynman–Kac representation of (C.9), with coefficients read at $T - s$ and $Y_0 = y$, gives (C.5). Section 2.2 summarizes its use for diffusion density corrections. Sufficient conditions are bounded continuous boundary and residual data, a bounded classical solution, nonexplosive Lipschitz auxiliary coefficients, and uniqueness in this solution class. Unbounded data require localization and uniform integrability of the stopped Feynman–Kac weights.

At $\eta = 1$, $D^1 = -b + A\nabla \log p$. Indeed, the reversed generator of the unconditional reference is

$$\widehat{L}_t f = \frac{L_t^*(p_t f) - f L_t^* p_t}{p_t} = (-b_t + A_t \nabla \log p_t) \cdot \nabla f + M_t : \nabla^2 f. \tag{C.10}$$

The auxiliary law is $P = P^{1,p_0}$ conditioned on its terminal point and read backward. At $\eta = 0$, $D^{0,\phi} = -(v + gM\nabla K)$, so $Y_s = x^y_{T-s}$ is the inverse guided characteristic. This proves (C.6). □

Specialization to fixed-scale CFG. For $\phi(t, z) = wz$, one has $g = w$, $R_\phi^{\mathrm{tr}} = w(w-1)E$, and $C_0 = C_0^w$ as defined in (14). At $\eta = 1$, (C.5) agrees with the reference-path average in (15). At $\eta = 0$, (C.6) proves Corollary 4 and (17). At $w = 1$ with $\nu = p_0^c$, $C_0^w = 1$ and the residual vanishes, recovering p_T^c at both endpoints. Starting instead from p_0 generally leaves a nonconstant boundary correction.

C.3 A JUSTIFIED ZERO-NOISE LIMIT

Keep ν and ϕ fixed. Assume $D^{\eta,\phi}$ is uniformly globally Lipschitz with linear growth, $\|D_t^{\eta,\phi}(x) - D_t^{0,\phi}(x)\| \le C\eta^2(1 + \|x\|)$, and σ_t is bounded. Couple Y^η and Y^0 from the same terminal point y. Their integral equations, the Burkholder–Davis–Gundy inequality, and Gronwall's inequality give

$$\mathbb{E} \sup_{s \le T} \|Y_s^\eta - Y_s^0\|^2 \le C_T \eta^2 (1 + \|y\|^2). \tag{C.11}$$

For bounded continuous C_0 and R_ϕ^{tr}, the weights converge in probability and are uniformly bounded, so their expectations converge. Localization and uniform integrability extend this conclusion to unbounded data. Hence the stochastic correction converges, for each such y, to the deterministic characteristic correction. This completes the zero-noise assertion of Theorem 7. It is a density-representation limit; it does not assert convergence of path KL, nor a total-variation or uniform density rate without further assumptions.

D TIME AND EVIDENCE POTENTIALS

D.1 PROOF OF COROLLARY 5

Work under the assumptions of Theorem 6 and, for the density statements, Theorem 7. For $F_t(x) = \phi(t, K_t(x))$, write $g = \phi_z$ and $j = \phi_{zz}$, with all derivatives evaluated at $(t, K_t(x))$. The chain rule and (A.10) give

$$(\partial_t + L_t^\eta) F = \phi_t - (1 - \eta^2) g H + (\eta^2 j - g) E. \tag{D.1}$$

Adding $\frac{1}{2}\|\nabla F\|^2_{G^\eta}$ gives the control residual

$$R_\phi^\eta = \phi_t - (1 - \eta^2)\phi_z H + \left[\eta^2 \phi_{zz} + \frac{1 + \eta^2}{2}\phi_z^2 - \phi_z\right] E. \tag{D.2}$$

Theorem 6 supplies the pointwise value $\phi(t, K_t(x))$, terminal utility $\phi(T, r)$, and feedback $u^* = G^\eta \phi_z \nabla K$. For the prescribed initial density ν, the initial optimal value is $\mathbb{E}_\nu \phi(0, K_0)$ when the integrability conditions hold. Theorem 7 supplies the density residual

$$R_\phi^{\mathrm{tr}} = \phi_t + (\phi_{zz} + \phi_z^2 - \phi_z) E.$$

Setting $\eta = 1$ and $\eta = 0$ in (D.2) gives the two control residuals in (19). The control and transport residuals coincide at $\eta = 1$, but not generally elsewhere. This proves Corollary 5.

Which policies and objectives this parameterization covers. Every sufficiently smooth scalar response $\bar{a}(t, z)$ has a potential

$$\phi(t, z) = b(t) + \int_{z_*}^{z} \bar{a}(t, \zeta)\, d\zeta.$$

The arbitrary time-only term $b(t)$ changes the value and running term by additive quantities and leaves the feedback unchanged. Bounds on ϕ_z impose the same bounds on guidance, while $\phi_{zz} \geq 0$ makes the scale nondecreasing in evidence. Conversely, specifying a terminal utility U and running term q requires solving $R^{\eta}_{\phi}(t, x) = q(t, x)$ with $\phi(T, z) = U(z)$. Such a solution need not exist within this family: states with identical K_t can have different E_t and H_t.

Time-dependent scales and an alternative factorization. For $F = a(t)K$, the residuals are

$$R^{\eta}_{a} = \dot{a}K - (1 - \eta^2)aH + \left[\frac{1 + \eta^2}{2}a^2 - a\right] E,$$
$$R^{\mathrm{tr}}_{a} = \dot{a}K + (a^2 - a)E. \tag{D.3}$$

This gives (20). For $F = \psi(K)$, replace ϕ_t by zero, ϕ_z by ψ', and ϕ_{zz} by ψ'' in the general formulas. Absolutely continuous schedules are covered when their almost-everywhere chain-rule terms are integrable.

Along a deterministic guided characteristic, $dK_t/dt = -H_t + (a(t) - 1)E_t$. Integration by parts therefore also gives

$$\log \frac{\rho^{0,a}_{T}(y)}{p^{c}_{T}(y)} = \log \frac{\nu(x^{y}_{0})}{p^{c}_{0}(x^{y}_{0})} - \int_{0}^{T} (a(t) - 1)(H_t + E_t)(x^{y}_{t})\, dt. \tag{D.4}$$

There is no explicit $\dot{a}$ here because the boundary factor uses p^c instead of pe^{aK}. Omitting $\dot{a}K$ while retaining the latter factorization changes the density formula; changing the factorization consistently does not.

D.2 Schedule formulas and budget conversion

Let $W > 1$, $\Delta = W - 1$, and $\tau = t/T$. For $a = 1 + \Delta f$, the coefficient budget is

$$B_{\mathrm{coef}}[a] = \frac{1}{2}\int_{0}^{1} (a^2 - a)\, d\tau = \frac{\Delta}{2}\int_{0}^{1} f\, d\tau + \frac{\Delta^2}{2}\int_{0}^{1} f^2\, d\tau.$$

Table D.1 collects the resulting formulas. The linear schedule is the power family at $p = 1$. The continuous power budget decreases strictly with p, so a target $B_{\mathrm{coef}} = \lambda_{\mathrm{budget}} W$ has a unique $p > 0$ exactly when $0 < \lambda_{\mathrm{budget}} < (W - 1)/2$.

For the implemented PCS grid ($N = 200$), the feasible interval is

$$\frac{W(W - 1)}{2N} < B_N(W, s) < \frac{(N - 1)W(W - 1)}{2N}.$$

The endpoint weights make its lower limit positive. Appendix I.1 specifies the inverse and conversion to the PCS mean–shape parameters. In physical time, $\frac{1}{2}\int_{0}^{T}(a^2 - a)\, dt = TB_{\mathrm{coef}}$. These budgets summarize coefficients; the full terminal correction also retains $\dot{a}K$, local energy, and trajectory dependence.

D.3 Exact evidence and Online-K feedback

For the potential $\psi(z) = z + \lambda\gamma^{-1}\log(1 + e^{\gamma(z-\kappa)})$, with $\lambda \geq 0$ and $\gamma > 0$, let $S(z) = \mathrm{sigmoid}(\gamma(z - \kappa))$. Then

$$\psi' = 1 + \lambda S, \qquad \psi'' = \lambda\gamma S(1 - S), \qquad R^{\mathrm{tr}}_{\psi} = [\lambda\gamma S(1 - S) + \lambda S + \lambda^2 S^2]E \geq 0.$$

This recovers the ideal guidance response in (23). The deterministic control residual need not be nonnegative: it is $R^{0}_{\psi} = -\psi' H + (\frac{1}{2}\psi'^2 - \psi')E$. For $\lambda > 0$, requiring $S(k_{\mathrm{low}}) \leq \varepsilon$ and

Table D.1: Coefficient budgets and conversion rules. Continuous budgets integrate over $\tau \in [0,1]$; PCS uses an equally weighted N-point grid including both endpoints. Write $\overline{r} = N^{-1}\sum_k r(\tau_k)$ and $\overline{r^2} = N^{-1}\sum_k r(\tau_k)^2$.

Schedule	Guidance	Budget	Conversion / range
Constant	$a = W$	$W(W-1)/2$	Fixed by W
Power	$a = 1 + \Delta\tau^p, p > 0$	$\frac{\Delta}{2(p+1)} + \frac{\Delta^2}{2(2p+1)}$	Strictly decreasing; $0 < B < W(W-1)/2$
Linear	$a = 1 + \Delta\tau$	$\Delta/4 + \Delta^2/6$	Power schedule, $p = 1$
Discrete PCS	$a_k = 1 + \Delta r_s(\tau_k)$; $r_s(\tau) = [1 - \cos(\pi\tau^s)]/2$	$[\Delta\overline{r}_s + \Delta^2\overline{r_s^2}]/2$	Unique $s > 0$ by bisection on the feasible interval
Early / late pulses	$1 + 2\sin^2(2\pi\tau)$ on the selected half; 1 elsewhere	$5/8$ for either pulse	Same coefficient budget; different timing

$S(k_{\text{high}}) \geq 1 - \varepsilon$, with $k_{\text{low}} < k_{\text{high}}$ and $0 < \varepsilon < 1/2$, is equivalent to

$$\gamma \geq \frac{2L_\varepsilon}{k_{\text{high}} - k_{\text{low}}}, \qquad k_{\text{low}} + \frac{L_\varepsilon}{\gamma} \leq \kappa \leq k_{\text{high}} - \frac{L_\varepsilon}{\gamma}, \qquad L_\varepsilon = \log\frac{1-\varepsilon}{\varepsilon}. \tag{D.5}$$

The chosen evidence levels must be attainable; for a discrete class, $k_{\text{high}} \leq -\log p(c)$. The response is bounded above by $1 + \lambda$, but this algebraic cap need not be reached on the population evidence range. At $\kappa = 0$ and $K = 0$, the extra guidance is $\lambda/2$.

Specializing the shared evidence dynamics (A.11) to deterministic sampling gives

$$\frac{d}{dt}K_t(X_t) = -H_t(X_t) + (a_t - 1)E_t(X_t). \tag{D.6}$$

The divergence contribution remains. Thus accumulating energy alone does not recover $K_t(X_t)$. A known constant shift can be absorbed into κ; a nonconstant initial K_0 cannot.

The Online-K update in (J.8) integrates the spatial and time-coordinate terms of the VP density-ratio differential on the actual DDIM endpoints. Its divergence term is estimated by input VJPs, and the gate uses the updated $\widehat{K}_j$ at the current endpoint. Appendix J.4 specifies the algorithm; Appendix J.7 derives the underlying differential and separates initialization, quadrature, learned-score, and random-trace errors. Corollary 5 applies to the exact state function $K_t(x)$. Applying the same scalar response to the accumulated estimate $\widehat{K}_j$ yields a numerical policy whose performance and calibration require separate evaluation.

The gate remains in $[1, 1 + \lambda]$ for every accumulated estimate. Appendix J.8 reports the calibration diagnostic and interprets the empirical threshold on the learned-estimate scale.

E Exact finite-step identities and compatibility

This appendix gives finite-step counterparts of the terminal-law and inverse-control characterizations: a transport identity for deterministic DDIM maps, path-density and KL identities for Gaussian kernels, and a Bellman identity with an explicit evidence-compatibility term.

E.1 Deterministic maps

Without nonlinear clipping, write the deterministic DDIM update as $T_k = m_k^{\text{u}} + a_k d_k$, where m_k^{u} is the unconditional update and $d_k = m_k^{\text{c}} - m_k^{\text{u}}$ is the conditional update difference. Let $G = T_{N-1} \circ \cdots \circ T_0$. The terminal measure is exactly $G_{\#}\nu$. If the maps are C^1 bijections with

nonzero Jacobian determinants, then along the inverse path ending at y,

$$\log \rho_N(y) = \log \nu(x_0) - \sum_{k=0}^{N-1} \log |\det DT_k(x_k)|, \qquad DT_k = Dm_k^{\mathrm{u}} + a_k Dd_k + d_k(\nabla a_k)^\top. \tag{E.1}$$

Finite-to-one nonsingular maps require summation over preimages. Collapsing maps can give singular pushforwards.

Let p_k^D be the actual unconditional discrete-sampler densities and define

$$\mathcal{T}_k^F(x) = \log |\det DT_k(x)| + \log \frac{p_{k+1}^D(T_k(x))}{p_k^D(x)} + F_{k+1}(T_k(x)) - F_k(x). \tag{E.2}$$

Telescoping yields the exact identity

$$\rho_N(y) = p_N^D(y) e^{F_N(y)} \frac{\nu(x_0)}{p_0^D(x_0)} \exp\left[-F_0(x_0) - \sum_{k=0}^{N-1} \mathcal{T}_k^F(x_k)\right]. \tag{E.3}$$

For $F = \phi(t, K)$ and $g = \phi_z(t, K)$, suppose the update approximates the smooth guided flow with local C^1 error $O(h^2)$. The Jacobian and population-density terms then contribute $h(gH + \phi_{zz}E) + O(h^2)$; the potential increment contributes $h[\phi_t - gH + g(g-1)E] + O(h^2)$. Their sum is $hR_\phi^{\mathrm{tr}} + O(h^2)$. Replacing p_k^D by forward-noising population densities requires density and derivative convergence in addition to trajectory consistency.

E.2 Nondegenerate Gaussian kernels

Suppose reference, conditional, and guided transitions have the same actual positive definite covariance Σ_k:

$$P_k^0 = \mathcal{N}(m_k^{\mathrm{u}}, \Sigma_k), \quad P_k^1 = \mathcal{N}(m_k^{\mathrm{u}} + d_k, \Sigma_k), \quad Q_k^a = \mathcal{N}(m_k^{\mathrm{u}} + a_k d_k, \Sigma_k).$$

Means, covariances, and scales are evaluated at the current state; the scale may also depend on the preceding history. Assume the initial guided density ν is absolutely continuous with respect to the reference density π_0, and that the expectations below are finite. Define the finite-step separation $\mathcal{E}_k = d_k^\top \Sigma_k^{-1} d_k$ and

$$\ell_k(x, y) = \log \frac{P_k^1(y \mid x)}{P_k^0(y \mid x)} = d_k^\top \Sigma_k^{-1}(y - m_k^{\mathrm{u}}) - \tfrac{1}{2}\mathcal{E}_k. \tag{E.4}$$

Gaussian expansion gives

$$\log \frac{dQ^a}{dP^0} = \log \frac{\nu(X_0)}{\pi_0(X_0)} + \sum_k \left[a_k \ell_k - \tfrac{1}{2}(a_k^2 - a_k)\mathcal{E}_k\right], \tag{E.5}$$

$$D_{\mathrm{KL}}(Q^a \| P^0) = D_{\mathrm{KL}}(\nu \| \pi_0) + \tfrac{1}{2}\mathbb{E}_{Q^a} \sum_k a_k^2 \mathcal{E}_k, \tag{E.6}$$

$$\mathbb{E}_{Q^a}[\ell_k \mid X_{0:k}] = (a_k - \tfrac{1}{2})\mathcal{E}_k, \qquad \mathrm{Var}_{Q^a}(\ell_k \mid X_{0:k}) = \mathcal{E}_k. \tag{E.7}$$

Positive definite Σ_k ensures the path-law absolute continuity used above. A step reaching $\bar{\alpha} = 1$ has zero posterior variance and must be treated separately; deterministic DDIM follows the map formulation above. The Online-K implementation integrates the VP density-ratio differential at DDIM endpoints (Appendix J.4).

E.3 Bellman inverse optimality and the compatibility defect

For an unrestricted reference Markov kernel and finite exponential moments, define

$$\mathcal{B}_k F(x) = \log \int e^{F(y)} P_k^0(dy \mid x), \quad R_k^F = \mathcal{B}_k F_{k+1} - F_k, \quad Q_k^F = e^{F_{k+1} - \mathcal{B}_k F_{k+1}} P_k^0.$$

At each current state,

$$\mathbb{E}_Q F_{k+1} - D_{\mathrm{KL}}(Q\|P_k^0) = \mathcal{B}_k F_{k+1} - D_{\mathrm{KL}}(Q\|Q_k^F). \tag{E.8}$$

Subtracting R_k^F and iterating gives the value F_k and the optimal kernel Q_k^F. This exponential tilt is generally non-Gaussian for nonlinear F_{k+1}. Likewise, a finite reverse conditional $q(x_j \mid x_i)$ is generally a mixture after marginalizing the clean state, even with exact conditional-mean prediction. The Gaussian CFG family connects to terminal evidence through the compatibility condition below.

Let $K_k^D(x) = \log \mathbb{E}_{P^0}[e^{r(X_N)} \mid X_k = x]$. The compatibility defect

$$\varepsilon_k(x,y) = \ell_k(x,y) - [K_{k+1}^D(y) - K_k^D(x)] \tag{E.9}$$

vanishes exactly when P_k^1 is the unit-reward Doob kernel, up to null sets. For fixed guidance,

$$\begin{aligned}\log \frac{dQ^w}{dP^0} = \log \frac{\nu}{\pi_0}(X_0) + wr(X_N) - wK_0^D(X_0) \\ - \tfrac{1}{2}(w^2 - w)\sum_k \mathcal{E}_k + w\sum_k \varepsilon_k.\end{aligned} \tag{E.10}$$

Terminal conditioning retains the defect inside the reference-chain bridge expectation. For a time schedule, the telescoping part becomes

$$\sum_{k=0}^{N-1} a_k(K_{k+1}^D - K_k^D) = a_{N-1}r(X_N) - a_0K_0^D(X_0) - \sum_{k=1}^{N-1}(a_k - a_{k-1})K_k^D(X_k). \tag{E.11}$$

For $a_k = \psi'(K_k^D)$, define instead

$$R_k^{\psi,\mathrm{step}} = \psi(K_{k+1}^D) - \psi(K_k^D) - a_k\ell_k + \tfrac{1}{2}(a_k^2 - a_k)\mathcal{E}_k.$$

Then the path log ratio equals the initial log ratio plus $\psi(r(X_N)) - \psi(K_0^D(X_0)) - \sum_k R_k^{\psi,\mathrm{step}}$. This exact adjacent-state residual is the finite-step counterpart of the continuous running correction; the continuous form follows under the corresponding limit and compatibility conditions.

F An exact Gaussian transport example

Consider forward Ornstein–Uhlenbeck noising $dZ_s = -Z_s ds + \sqrt{2}\, dW_s$ and the sampling-time marginals

$$p_t = \mathcal{N}(0,1), \qquad p_t^c = \mathcal{N}(0,q_t), \qquad q_t = 1 - \tfrac{1}{2}e^{2(t-1)}, \qquad 0 \le t \le 1.$$

Here $M = 1$, $v = 0$, and with $\delta_t = 1 - q_t^{-1}$,

$$K_t(x) = -\tfrac{1}{2}\log q_t + \tfrac{1}{2}\delta_t x^2, \quad H_t(x) = \delta_t(1 - x^2), \quad E_t(x) = \delta_t^2 x^2.$$

Deterministic fixed guidance solves $\dot{X}_t = w\delta_t X_t$, hence $X_t = (q_t/q_0)^{w/2}X_0$. At $w = 2$, choose

$$\nu = Z_0^{-1}p_0 e^{2K_0}, \qquad Z_0 = [q_0(2 - q_0)]^{-1/2}, \qquad \nu = \mathcal{N}\left(0, \frac{q_0}{2 - q_0}\right). \tag{F.1}$$

This makes the entire initial boundary factor constant. Transport gives

$$\mathrm{Var}(X_1) = \frac{q_1^2}{q_0(2 - q_0)} = \frac{1}{4 - e^{-4}} \simeq 0.251150. \tag{F.2}$$

The normalized powered endpoint law $p_1 e^{2K_1}$ instead has variance $1/3$.

Along the inverse characteristic $x_t^y = yq_t/q_1$,

$$\int_0^1 E_t(x_t^y)dt = \frac{y^2(1 - e^{-4})}{4}.$$

Since $w^2 - w = 2$, the complete correction is $Z_0^{-1}\exp[-(1 - e^{-4})y^2/2]$. Multiplying it by $p_1 e^{2K_1}$ gives precision $3 + (1 - e^{-4}) = 4 - e^{-4}$ and the corresponding normalizing constant, agreeing with (F.2). The discrepancy from the powered endpoint law therefore persists with a constant initial boundary factor and requires no discretization or estimated score.

G STOCHASTIC GAUSSIAN MIXTURES: CONFIGURATION AND TERMINAL-LAW EVALUATION

Mixture and initialization. The eight-component mixture has targets $A = (8, 0)$ and $B = (-8, 0)$, distractors $(-5, \pm 1.5)$ and $(-3, 0)$, and background centers $(0, \pm 7.5)$ and $(5.5, 5.5)$. All components have covariance $0.45^2 I$. Unconditional weights π_j are uniform; conditional weights q_j are $16/47$ per target, $1/47$ per distractor, and $4/47$ per background component. In forward noising time $s \in [0, 1]$,

$$\beta(s) = 0.1 + 7.9s, \quad \alpha(s) = \exp[-(0.1s + 3.95s^2)/2], \quad v_s = \alpha(s)^2(0.45)^2 + 1 - \alpha(s)^2.$$

For mixture weights a_j, the exact responsibilities and scores are

$$\gamma^a_{j,s}(x) = \frac{a_j \mathcal{N}(x; \alpha(s)\mu_j, v_s I)}{\sum_l a_l \mathcal{N}(x; \alpha(s)\mu_l, v_s I)}, \qquad s_a(x, s) = \frac{\alpha(s) \sum_j \gamma^a_{j,s}(x)\mu_j - x}{v_s}.$$

Direct reverse sampling uses Euler–Maruyama,

$$x_{s-h} = x_s + h\beta(s)\{x_s/2 + s_u(x_s, s) + w[s_c - s_u](x_s, s)\} + \sqrt{h\beta(s)}\,\zeta, \quad \zeta \sim \mathcal{N}(0, I),$$

initialized from

$$\nu_w = \sum_j m^{\text{end}}_j(w)\mathcal{N}(\alpha(1)\mu_j, v_1 I), \qquad m^{\text{end}}_j(w) \propto \pi_j^{1-w} q_j^w.$$

Since $\alpha(1) = 0.131994$, this law retains component information and generally differs from a full-density power tilt. The boundary factor in Eq. (15) cannot in general be omitted. Mode masses refer to nearest-center Voronoi regions, rather than latent component assignments.

Terminal-law estimator. Let $p_{u,s}, p_{c,s}$ denote the noising-time densities, $K_s = \log(p_{c,s}/p_{u,s})$, and $r = K_0$. Proposal endpoints $Y \sim g$ initialize exact forward VP paths $Z_0 = Y$, with

$$g = 0.1p_{u,0} + 0.3(g_1 + g_3 + g_8), \qquad g_w = \sum_j m^{\text{end}}_j(w)\mathcal{N}(\mu_j, 0.45^2 I),$$

$$Z_{s+h} = \frac{\alpha(s+h)}{\alpha(s)} Z_s + \sqrt{1 - \frac{\alpha(s+h)^2}{\alpha(s)^2}}\,\xi, \qquad \xi \sim \mathcal{N}(0, I).$$

The complete importance weight is

$$W(Y, Z) = \frac{p_{u,0}(Y)e^{wr(Y)}}{g(Y)} \frac{\nu_w(Z_1)}{p_{u,1}(Z_1)} \exp\left[-wK_1(Z_1) - \frac{w(w-1)}{2} J(Z)\right],$$

$$J(Z) = \int_0^1 \beta(s)\|s_c(Z_s, s) - s_u(Z_s, s)\|^2\, ds.$$

The prediction $\widehat{m}(B) = M^{-1} \sum_i W_i \mathbf{1}_{\{Y_i \in B\}}$ is evaluated without self-normalization; the law predicts total mass $Z = \mathbb{E}[W] = 1$. Seeds 101, 10000101, and 20000101 each provide 65,536 forward paths shared across scales and 32,768 independent reverse samples per scale, totaling 196,608 and 98,304, respectively. Energy uses 2,000-interval trapezoidal quadrature. Evaluation covers eight Voronoi regions and 441 spatial bins: 22×20 unit cells on $[-11, 11) \times [-10, 10)$ plus overflow. We report $D = \frac{1}{2} \sum_B |\widehat{m}(B) - \widehat{m}_{\text{direct}}(B)|$.

Complete law and omission controls. At $w = 3$, the complete law reduces mode discrepancy to 0.002586, compared with 0.119934 without the boundary factor and 0.086362 for the endpoint-only tilt (Table G.2 and Figure G.1). It also recovers unit mass without normalization (Table G.1). The omission controls are normalized before comparison. At $w = 1$, the boundary terms cancel and $W = p_{c,0}(Y)/g(Y)$, providing a conditional-density check.

Monte Carlo uncertainty and refinement. For bin predictions $\widehat{p}_j$, reverse fractions $\widehat{d}_j$, and $b_j = M^{-1} \sum_i W_i^2 \mathbf{1}_{\{Y_i \in B_j\}}$, uncertainty is estimated by

$$\widehat{V} = \frac{\text{diag}(b) - \widehat{p}\widehat{p}^{\mathsf{T}}}{M - 1} + \frac{\text{diag}(\widehat{d}) - \widehat{d}\widehat{d}^{\mathsf{T}}}{N - 1},$$

$$R = \frac{1}{2}\sqrt{\frac{2}{\pi}} \sum_j \sqrt{\widehat{V}_{jj}} + \sqrt{\frac{1}{2}\log(1/\alpha_{\text{MC}}) \sum_{j,k} |\widehat{V}_{jk}|}, \quad \alpha_{\text{MC}} = 0.05/9.$$

Table G.1 reports $D + R$. This Gaussian approximation is not a finite-sample confidence bound. The precision thresholds are 0.02 for $|Z - 1| + z_{1-\alpha_{\mathrm{MC}}/2}\,\mathrm{SE}(Z)$, 0.03 for the mode upper estimate, and 0.06 for the grid upper estimate, together with ESS at least 2,000 and maximum weight share at most 0.005.

At $w = 3$, paired 1,000/2,000-step calculations change the normalizer and target A/B masses by -8.00×10^{-6}, 4.63×10^{-6}, and 1.93×10^{-4}, respectively; grid discrepancy changes from 0.007054 to 0.007033. At $w = 8$, $Z = 0.963671 \pm 0.013898$ (SE), and all three primary upper errors exceed their thresholds. No step-refinement result is available at this scale. Since the normalizer uses only forward paths, its discrepancy cannot be attributed solely to reverse discretization; rare influential paths also remain a possible source of importance-sampling error. Thus the strong-guidance estimate does not yet resolve the identity.

Table G.1: Complete stochastic terminal-law evaluation. Parentheses on Z are standard errors; those on discrepancies are approximate uncertainty-inclusive upper estimates defined in the text. Precision criteria are met at $w = 1, 3$; the $w = 8$ estimate remains unresolved.

w	Z (SE)	D_{mode} (upper)	D_{grid} (upper)	ESS
1	.9993 (.0009)	.0032 (.0147)	.0090 (.0282)	167,462
3	1.0003 (.0024)	.0026 (.0129)	.0070 (.0243)	90,381
8	.9637 (.0139)	.0184 (.0497)	.0275 (.0723)	4,693

Table G.2: Mode-mass discrepancy for the complete law and normalized omission controls. The $w = 8$ complete-law estimate has insufficient precision to validate the identity at this sampling budget.

w	Complete law	Without boundary factor	Endpoint-only tilt
1	.003241	.003369	.003369
3	.002586	.119934	.086362
8	.018380	.349874	.093117

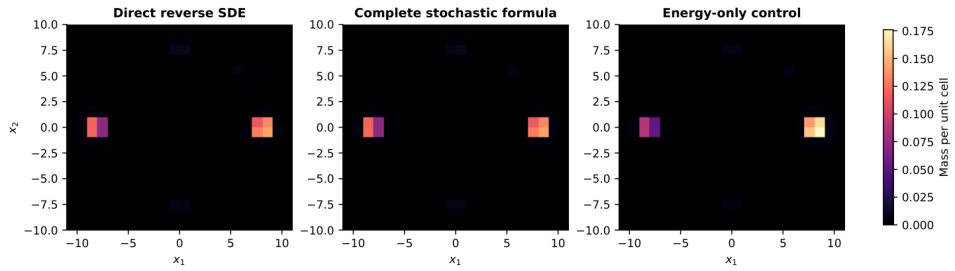

Figure G.1: Stochastic terminal masses at $w = 3$. Direct reverse-SDE samples, the complete formula, and the normalized control omitting the boundary factor use the same mass-per-cell color scale. The complete formula is evaluated without self-normalization.

H Deterministic Gaussian mixtures: transport and guidance timing

Mixture and numerical solvers. The one-dimensional mixture has centers $(-3, -1, 3)$, variance 0.25, unconditional weights $(1, 1, 1)/3$, and conditional weights $(4, 0.25, 4)/8.25$. We use the VP rate from Appendix G, with $v_s = 0.25\alpha(s)^2 + 1 - \alpha(s)^2$, and initialize every policy from the exact conditional marginal at noising time $s = 1$. Scores are analytic, and all computations use float64 on a

CPU. For responsibilities r_j and $d_j = (\alpha(s)\mu_j - x)/v_s$,

$$s_a = \sum_j r_j d_j, \qquad \partial_x s_a = -1/v_s + \sum_j r_j d_j^2 - s_a^2.$$

We check these derivatives by centered differences. With reverse progress $\tau = 1 - s$, the guided velocity is $v^a = \frac{1}{2}\beta(s)[x + s_u + a(\tau)(s_c - s_u)]$.

Inverse RK4 integration jointly advances the characteristic, velocity divergence, transport residual, and energy term. We compare Eq. (17) with the flow-Jacobian density

$$\rho_T^{\mathrm{Jac}}(y) = \nu(x_0^y) \exp\left[-\int_0^T \partial_x v^a(t, x_t^y)\, dt\right].$$

The base grid has 2,049 endpoints on $[-7, 7]$ with 128, 256, or 512 RK4 steps; refinement uses 4,097 endpoints with 512 or 1,024 steps. A separate forward DOP853 solver transports 257 initial CDF quantiles uniformly spaced on $[10^{-4}, 1 - 10^{-4}]$, using relative/absolute tolerances $2 \times 10^{-11}/2 \times 10^{-12}$. This comparison uses neither the density residual nor the velocity divergence.

Terminal-density accuracy. The complete formula agrees with the Jacobian density to within 2.1×10^{-11} in half-L^1 discrepancy, with raw normalization errors below 10^{-9} (Table H.1). These density calculations share numerical characteristics; the separate forward-quantile comparison gives CDF discrepancies at most 3.1×10^{-6}, limited by interpolation. These are numerical consistency checks rather than certified bounds on the exact density. In contrast, the normalized endpoint-only tilt has TV errors 0.28035 at $w = 3$ and 0.6267 at $w = 8$, demonstrating the contribution of the transport correction.

Table H.1: Deterministic terminal-density checks with 4,097 endpoints and 1,024 RK4 steps. Full D compares raw densities; omission errors compare normalized densities. Omit Z is the raw mass after removing $\dot{a}K$. CDF errors use the separate forward solver.

Policy	Full D	Endpoint D	Omit Z	Omit D	CDF error
Fixed $w = 1$	6.0×10^{-13}	.0000	1.0000	.0000	6.8×10^{-7}
Fixed $w = 3$	4.6×10^{-12}	.2803	1.0000	.0000	1.8×10^{-6}
Fixed $w = 8$	2.1×10^{-11}	.6267	1.0000	.0000	3.0×10^{-6}
Power $1 \to 3$	1.5×10^{-12}	.0669	1.6738	.0869	1.7×10^{-6}
Early pulse	3.7×10^{-12}	.1510	.8759	.1515	1.1×10^{-6}
Late pulse	9.7×10^{-12}	.1002	.8103	.1362	2.6×10^{-6}

Timing and schedule-derivative controls. Besides fixed scales $1, 3, 8$ and the power ramp $a(\tau) = 1 + 2\tau^2$, we evaluate

$$a_{\mathrm{early}}(\tau) = 1 + 2\sin^2(2\pi\tau)\mathbf{1}_{\{\tau \le 1/2\}}, \quad a_{\mathrm{late}}(\tau) = 1 + 2\sin^2(2\pi\tau)\mathbf{1}_{\{\tau \ge 1/2\}}.$$

Both pulses are continuously differentiable, start and end at one, peak at three, and have mean 1.5, the same scale histogram, and $B_{\mathrm{coef}} = 5/8$. Nevertheless, their terminal TV distance is 0.17418726; doubling spatial resolution changes it by only 2.8×10^{-6} (Figure H.1). On the finest grid, left/right Voronoi masses (boundaries $-2, 1$) are 0.450815/0.529745 for early guidance and 0.504119/0.490024 for late guidance. Their TV distances to the conditional target are 0.15095 and 0.10017, respectively. Thus matched coefficient budgets can yield different mode allocation and target discrepancy; this example does not establish a general preference for late guidance.

For the power ramp, omitting $\dot{a}K$ from Eq. (20) gives raw mass 1.6738 and normalized TV error 0.08687. The corresponding early/late errors are 0.15145/0.13622. These controls remove the derivative term while retaining the original boundary factors; an equivalent integration-by-parts representation would also change those factors.

Finite-step DDIM convergence. We apply Eq. (A.3) to the same 257 initial quantiles, using exact score-to-epsilon conversion, no clipping, and physical left-step progress i/N. Table H.2 reports endpoint errors relative to the forward ODE solution. Errors decrease with step refinement, although

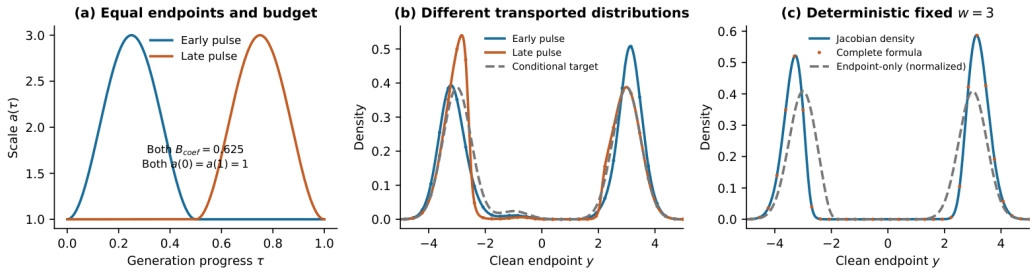

Figure H.1: Guidance timing at equal coefficient budget. (a) Early and late pulses match in endpoint, peak, mean, and scale histogram. (b) Their terminal densities differ by TV 0.17419; the complete formula matches the Jacobian-density curves. (c) At $w = 3$, the transport correction explains the discrepancy from the normalized endpoint-only tilt.

coarse integration can be inaccurate across low-density mode gaps: at $w = 8$, the maximum error falls from 5.88968 at 64 steps to 0.00259 at 2,048 steps. Across policies, the finest-grid RMS errors range from 0.000477 to 0.001009. The continuous terminal law therefore characterizes the limiting ODE, rather than an exact finite-step DDIM chain.

Table H.2: DDIM endpoint errors on 257 initial quantiles at three step counts. RMS and maximum errors use the same forward-ODE reference for every policy.

	64 steps		512 steps		2,048 steps	
Policy	RMS	Maximum	RMS	Maximum	RMS	Maximum
Fixed $w = 1$	.08729	1.35756	.00432	.05445	.00101	.01207
Fixed $w = 3$	.01404	.05190	.00185	.00687	.00048	.00178
Fixed $w = 8$	.36767	5.88968	.00205	.01079	.00053	.00259
Power $1 \to 3$	.04741	.72339	.00341	.04575	.00079	.01004
Early pulse	.02370	.18465	.00297	.02166	.00075	.00567
Late pulse	.04673	.70805	.00374	.05110	.00088	.01158

I PCS SELECTION THROUGH A DISCRETE COEFFICIENT BUDGET

We compare two prespecified 12-candidate searches within the power-cosine schedule (PCS) family against the original PCS configuration. Both searches use development FID for selection and fresh paired seeds for final evaluation. This study is independent of the Online-K comparison in Section 6.3.

I.1 DISCRETE PCS BUDGET AND ITS INVERSE

We retain the guidance convention $\widehat{\epsilon}_k = \epsilon_k^u + a_k(\epsilon_k^c - \epsilon_k^u)$, so that $a_k = 1$ gives the conditional predictor. On the actual $N = 200$ sampling grid, $\tau_k = k/(N-1)$ increases from noise to data. The PCS schedule used in the Wang family (Wang et al., 2024) is

$$r_s(\tau) = \frac{1 - \cos(\pi\tau^s)}{2}, \qquad \overline{r}_s = \frac{1}{N}\sum_{k=0}^{N-1} r_s(\tau_k),$$
$$a_k = 1 + (m-1)\frac{r_s(\tau_k)}{\overline{r}_s}, \qquad s > 0. \tag{I.1}$$

Here $m = N^{-1}\sum_k a_k$ is the discrete mean coefficient. At a fixed $m > 1$, changing s also changes the peak $W = 1 + (m-1)/\overline{r}_s$; it does not isolate timing while holding all coefficient-strength summaries fixed.

Instead, specify $W > 1$ and a discrete coefficient budget, and write

$$\begin{aligned} a_k(W, s) &= 1 + (W-1) r_s(\tau_k), \\ B_{200}(W, s) &= \frac{1}{2N} \sum_{k=0}^{N-1} (a_k^2 - a_k) \\ &= \frac{1}{2} \left[(W-1)\overline{r}_s + (W-1)^2 \overline{r_s^2} \right], \end{aligned} \tag{I.2}$$

where $\overline{r_s^2} = N^{-1} \sum_k r_s(\tau_k)^2$. All 200 points, including both endpoints, have equal weight. This budget summarizes schedule coefficients; it omits trajectory-dependent energy, diffusion weights, and the schedule-derivative term in the terminal law.

Strict monotonicity and feasible interval. For every interior grid point $0 < \tau_k < 1$ and finite $s > 0$,

$$\partial_s r_s(\tau_k) = \frac{\pi}{2} \sin(\pi \tau_k^s) \tau_k^s \log \tau_k < 0. \tag{I.3}$$

The two endpoint derivatives are zero. Since $a_k \geq 1$ and $W > 1$,

$$\partial_s B_{200}(W, s) = \frac{W-1}{2N} \sum_{k=0}^{N-1} (2a_k - 1)\, \partial_s r_s(\tau_k) < 0. \tag{I.4}$$

As $s \to \infty$, only the last grid point retains coefficient W; as $s \to 0^+$, all points except the first approach coefficient W. Continuity and strict monotonicity therefore give the exact range

$$\frac{W(W-1)}{400} < B_{200}(W, s) < \frac{199W(W-1)}{400}, \qquad 0 < s < \infty. \tag{I.5}$$

Every interior target has a unique finite positive solution s. The positive lower limit comes from the full weight of the terminal point.

To invert the budget, we start with bounds 0 and 1, double the upper bound until its budget is no greater than the target, and perform 80 bisection iterations. We then recover

$$m = 1 + (W-1)\overline{r}_s \tag{I.6}$$

for the (m, s) PCS implementation. All 12 targets are strictly feasible. Coefficients are calculated in double precision and converted to FP32 for sampling.

I.2 Candidate sets, shared anchor, and selection

The conventional parameter grid is

$$m \in \{1.15, 1.3, 1.45\}, \qquad s \in \{1, 2, 4, 8\}. \tag{I.7}$$

The coefficient-budget grid is

$$W \in \{2, W_0, 3.5\}, \qquad B_{200} = f \lambda_0 W, \qquad f \in \{0.6, 0.8, 1, 1.25\}. \tag{I.8}$$

The shared anchor is the previously selected PCS setting $(m, s) = (1.3, 4)$:

$$\begin{aligned} W_0 &= 2.6682769268545794, \\ B_0 &= 0.31796101270447535, \\ \lambda_0 &= B_0 / W_0 = 0.11916342322057795. \end{aligned} \tag{I.9}$$

Both grids contain the anchor, giving 23 distinct configurations. The anchor and grid factors are empirical design choices. Each configuration generates 10,000 development images with seed 20261013. Q1 minimizes development FID; Q2 minimizes it subject to development accuracy of at least 98.2%. Both objectives select the same winner within each search. All candidates are specified before development evaluation, and the selected configurations are fixed before final evaluation. The development and final FID rankings of the two winners differ; final results do not affect selection.

Table I.1 lists the three evaluated configurations. The conventional search selects $(m, s) = (1.45, 8)$; the budget search selects $W = 3.5$, $f = 1$, which converts to $(m, s) = (1.314250797543, 6.034692671642)$. The candidate counts match, but the search domains differ: the budget grid includes exponents absent from the conventional grid.

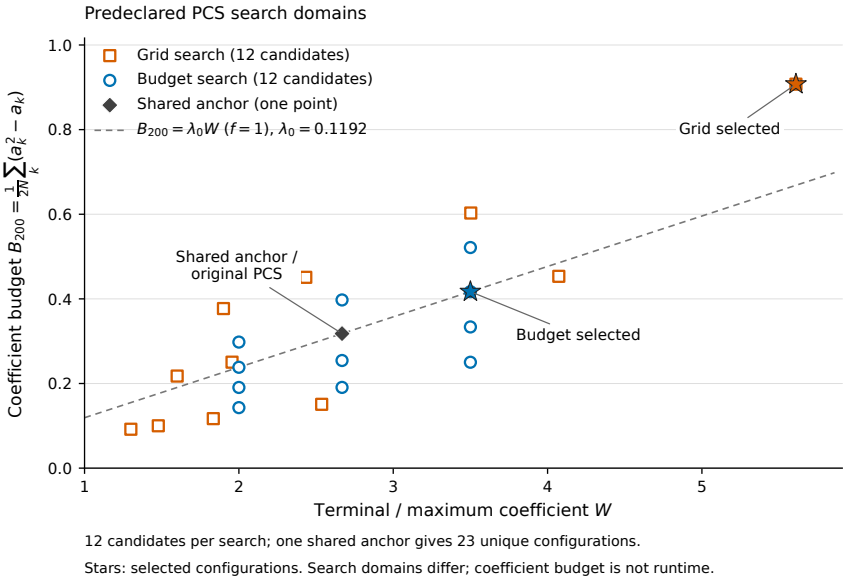

Figure I.1: Prespecified PCS candidate sets in peak–budget coordinates. Each search has 12 candidates; the shared anchor leaves 23 distinct configurations. Stars identify the selected configurations. The dashed line denotes the budget ratio $B_{200}/W = \lambda_0$.

Table I.1: PCS settings selected by Q1 and Q2; both objectives choose the same configuration within each search.

Selection	m	s	W	B_{200}
Original PCS	1.3000000	4.0000000	2.6682769	0.3179610
Parameter grid	1.4500000	8.0000000	5.6091325	0.9074266
Budget grid	1.3142508	6.0346927	3.5000000	0.4170720

m is the discrete mean coefficient, s the shape exponent, W the peak, and $B_{200} = (2N)^{-1}\sum_k (a_k^2 - a_k)$ the equal-weight coefficient budget for $N = 200$. Displayed parameters are rounded.

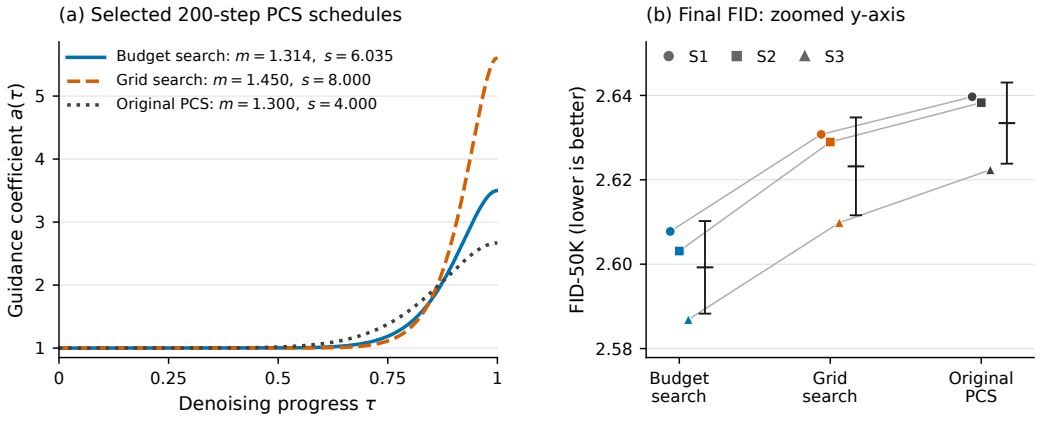

Figure I.2: PCS selection with equal candidate counts. (a) Implemented schedules. (b) Final FID on three paired seeds; lines connect the same seed and black intervals show means ± sample SD. The FID axis is zoomed. Timing and strength vary jointly, and the two candidate grids cover different domains.

I.3 Sampling and evaluation protocol

We use the epoch-1999 EMA checkpoint and evaluation assets specified in Appendix J, without retraining. Deterministic DDIM-200 uses $\eta = 0$, network indices $996, 991, \ldots, 1$, and terminal index zero, with the 1,001-point training schedule, mixed-prediction x_0 clipping, epsilon recomputation, FP16 inference, and FP32 updates. Every configuration uses 200 doubled-batch network calls, or 400 branch predictions per image.

The final seeds are 20261021, 20261022, and 20261023. Each configuration generates 50,000 images per seed, with 5,000 images per class, in batches of 128 and a final batch of 80. The noise key uses phase, seed, and global image ID; labels are IDs modulo ten. Configurations share initial noise, labels, and batch membership within each seed. The original PCS is rerun on these seeds, giving nine final runs. Q1 and Q2 share these runs. The seeds are distinct from those in Appendix J.5.

All metrics use RGB uint8 PNG images. FID uses the CIFAR-10 train50K reference and the common Inception features; precision and recall use all 50,000 generated and reference features with $k = 3$. Conditional accuracy is measured by the fixed CIFAR ResNet-18. Metrics are calculated separately for each run and reported as mean $\pm$ sample SD across three seeds. Every configuration exceeds the 98% accuracy floor in every seed.

I.4 All final runs and paired differences

Budget-based selection achieves the lowest FID in all three paired seeds against both controls. Tables I.2 and I.3 report aggregate and per-seed results.

Table I.2: PCS results over three paired seeds, with 50,000 images per configuration and seed. Accuracy is in percent; feature precision and recall use $k = 3$.

Selection	FID ↓	Acc. (%)	Precision	Recall
Original PCS	2.6334 ± 0.0096	98.349 ± 0.076	0.6877 ± 0.0015	0.5760 ± 0.0034
Parameter grid	2.6232 ± 0.0116	98.931 ± 0.007	0.6889 ± 0.0014	0.5716 ± 0.0022
Budget grid	2.5993 ± 0.0110	98.459 ± 0.055	0.6856 ± 0.0014	0.5760 ± 0.0014

$\pm$ denotes sample SD. All configurations satisfy the 98% accuracy floor in every seed.

Table I.3: All nine final runs, with 50,000 generated images per row and three paired seeds per configuration.

Selection	Seed	FID ↓	Acc. (%)	Precision	Recall
Original PCS	S1	2.639683	98.360	0.68786	0.57402
	S2	2.638276	98.268	0.68618	0.58000
	S3	2.622367	98.418	0.68910	0.57404
Parameter grid	S1	2.630768	98.938	0.68840	0.56990
	S2	2.628961	98.924	0.68772	0.57402
	S3	2.609842	98.930	0.69050	0.57088
Budget grid	S1	2.607769	98.450	0.68662	0.57442
	S2	2.603113	98.410	0.68402	0.57700
	S3	2.586870	98.518	0.68604	0.57670

S1–S3 denote 20261021, 20261022, and 20261023, respectively.

For each final seed j, define the paired difference

$$\Delta_j = \mathrm{FID}_{\mathrm{budget},j} - \mathrm{FID}_{\mathrm{control},j}. \tag{I.10}$$

Negative values favor budget-based selection. The paired differences against the conventional grid are $(-0.022999, -0.025848, -0.022972)$ in seed order; against the original PCS, they are $(-0.031914, -0.035163, -0.035497)$. Table I.4 summarizes these differences, with SD computed across pairs. Relative FID reductions are 0.913% and 1.298%, respectively.

Compared with the conventional-grid winner, budget-based selection has higher mean recall (0.57604 versus 0.57160), with lower accuracy (98.459% versus 98.931%) and precision (0.68556 versus

Table I.4: Paired final FID differences: budget grid minus control, with three paired seeds and 50,000 images per configuration and seed.

Control	Δ FID (mean $\pm$ SD)	Reduction	Lower-FID seeds
Parameter grid	-0.023940 ± 0.001653	0.913%	3/3
Original PCS	-0.034191 ± 0.001979	1.298%	3/3

Negative differences favor the budget grid. SD is computed across paired differences. Reduction is $-100\overline{\Delta}/\overline{\mathrm{FID}}_{\mathrm{control}}$.

0.68887). These directions hold in every paired seed. Relative to the original PCS, accuracy is higher and precision lower in every seed, while mean recall is essentially unchanged (0.57604 versus 0.57602). The FID gains therefore accompany an accuracy–precision–recall trade-off rather than uniform improvement across metrics.

I.5 Interpretation

The selected budget schedule raises peak and budget by 31.17% over the original PCS, while changing the mean from 1.3 to approximately 1.31425; its peak and budget remain below those of the conventional-grid winner. The improvement thus reflects a jointly selected allocation of guidance strength and timing. It supports the proposed coordinates for these prespecified searches on CIFAR-10/DDIM-200, without isolating individual schedule effects or establishing superiority across arbitrary search domains. The three paired seeds measure sampling variability for one fixed checkpoint.

J CIFAR-10: Online-K implementation and evaluation

This appendix specifies the guidance methods and CIFAR-10 evaluations in Section 6.3. We provide the four baseline families, the Online-K estimator, final results, and evidence-update diagnostics. The original baseline evaluation and the paired FP32 Online-K/PCS evaluation share the model, grid, and metric pipeline but use different precision and seed groups. Appendix I describes the separate PCS selection study.

J.1 Sampling and evaluation protocol

Frozen model and sampling grid. All configurations use the same ten-class U-Net with 128 base channels, channel multipliers $(1, 2, 2, 2)$, two residual blocks per level, attention at the second resolution level, and dropout parameter 0.1. Evaluation mode disables dropout. We load the EMA weights of `model_1999.pth` strictly using the `ema_model.nn_model.` prefix; no model is retrained. The numerical VP schedule consists of 1,001 FP32 values $\beta = \mathrm{linspace}(10^{-4}, 0.02, 1001)$, with $\bar{\alpha} = \exp(\mathrm{cumsum}(\log(1-\beta)))$. The DDIM sampler (Song et al., 2021a) uses 200 transitions and $\eta = 0$. The network indices are $t_j = 996 - 5j$, $j = 0, \ldots, 199$, and network time is $t_j/1000$. The final transition reaches index $t_{200} = 0$, where $\bar{\alpha}_0 \simeq 0.9999$, rather than exactly one.

Write $A_j = \bar{\alpha}_{t_j}$ and $B_j = \bar{\alpha}_{t_{j+1}}$. The coefficient convention is

$$\widehat{\epsilon}_j = \epsilon_j^u + a_j(\epsilon_j^c - \epsilon_j^u), \tag{J.1}$$

so $a = 0$ is unconditional prediction and $a = 1$ is conditional prediction. Every configuration applies the same clipped DDIM map:

$$h_j(e) = \mathrm{clip}_{[-1,1]}\left(\frac{x_j - \sqrt{1-A_j}\,e}{\sqrt{A_j}}\right), \qquad e'_j(e) = \frac{x_j - \sqrt{A_j}\,h_j(e)}{\sqrt{1-A_j}},$$
$$T_j(e) = \sqrt{B_j}\,h_j(e) + \sqrt{1-B_j}\,e'_j(e), \qquad x_{j+1} = T_j(\widehat{\epsilon}_j). \tag{J.2}$$

Thus predictions are mixed before the clean-image estimate is clipped, and epsilon is recomputed after clipping. Online-K evaluates its raw scores at the reached states x_j before this clipping operation.

Paired randomness and numerical precision. For global sample index i, the target label is $i \bmod 10$. Initial noise is generated on CPU with a per-image seed formed from the first eight SHA256 bytes, interpreted as a little-endian unsigned integer, of `phase:seed:i`. Methods evaluated within the same seed group therefore share initial noise and labels. Trace probes use the run seed plus 10,000,000; each probe is keyed by this trace seed, sample index, endpoint index, and probe index. This scheme preserves noise and probe identities across batch boundaries and resumed runs. Online-K is deterministic conditional on both its noise and probe streams, even though DDIM itself uses $\eta = 0$. In the paired Online-K/PCS evaluation, network inference and input vector–Jacobian products (VJPs) use FP32, without autocast or TF32. Evidence reductions and accumulation use FP64. cuDNN benchmarking is disabled and deterministic mode is enabled. All final runs use NVIDIA A100 80GB PCIe GPUs and batches of 128 images; a 50,000-image run ends with a batch of 80 after 390 full batches. These precision settings apply to both Online-K and its paired PCS reference. The original baseline precision and seeds are specified in Appendix J.3.

Image representation and FID. The final tensors are mapped from $[-1, 1]$ to $[0, 255]$, rounded by adding 0.5 before uint8 conversion, and saved as RGB PNG files. All image metrics read these files. FID (Heusel et al., 2017) compares generated images with all 50,000 CIFAR-10 training images, using the 2,048-dimensional Inception features from `pytorch-fid 0.3.0` and the fixed `pt_inception-2015-12-05-6726825d.pth` weights. RGB uint8 images are converted to $[0, 1]$, resized bilinearly to 299×299 with `align_corners=False`, and mapped by $2x - 1$. Feature extraction uses FP32 with TF32 disabled; means, covariances, and FID are computed in FP64. FID is evaluated on the mixture of all classes.

Secondary metrics. Classifier agreement is the fraction of PNG images whose predicted label under a fixed CIFARResNet18 matches the requested label. The classifier uses no resizing, channel means $(0.4914, 0.4822, 0.4465)$, and standard deviations $(0.247, 0.2435, 0.2616)$. Agreement is reported in percent and measures consistency with this evaluator. Feature precision and recall (Kynkäänniemi et al., 2019) use the same Inception features as FID and all 50,000 real and generated images in each final run. Each feature ball has radius equal to the distance to the third nearest other point in its own set ($k = 3$, self excluded). Precision is the fraction of generated points covered by real feature balls; recall is the fraction of real points covered by generated feature balls. Squared Euclidean distances use FP32 with TF32 disabled and row blocks of size 512.

J.2 Guidance settings and development selection

Table J.1 summarizes development selection and final evaluation. Online-K uses Eq. (23), with $\gamma = 4$ and one trace probe per endpoint ($R = 1$); minimum development FID selects $(\lambda, \kappa) = (2, 4)$. Its paired PCS reference is held at the previously selected $(m, s) = (1.3, 4)$, so the two policies have different search histories. All selections are fixed before final evaluation. The additional agreement-constrained configuration and complete Online-K development scores appear in Appendix J.9.

Table J.1: Selection and evaluation protocol. Development evaluations use 10,000 images per candidate; final evaluations use 50,000 per configuration and seed. All groups use the common DDIM-200 protocol.

Evaluation group	Development selection	Final evaluation
Online-K	Six candidates: $\lambda \in \{1, 2\}$, $\kappa \in \{1, 2, 4\}$; $\gamma = 4$, $R = 1$. Seed 20260927; minimum FID selects $(2, 4)$.	FP32; seeds 20260928–30, paired with PCS.
Paired PCS	Fixed $(m, s) = (1.3, 4)$ from the earlier baseline selection.	Same FP32 seeds and initial noise as Online-K.
Four baseline families	18 candidates per family; seed 20260914; minimum FID within each family. Selected parameters appear in Appendix J.3.	FP16 inference, FP32 updates; seeds 20260921–23, paired within this group.

J.3 FID-SELECTED BASELINE CONFIGURATIONS

The four baseline rows in Table 2 use the development and final groups in Table J.1. Linear, Clamp-linear, and PCS share the 18-candidate Wang-family search, which selects PCS. Exact FID ties are resolved by configuration ID. All four families use the common checkpoint, DDIM grid, clipping rule, and metric pipeline in Appendix J.1; FBG observer calculations use FP64.

The FP16 baseline group and FP32 Online-K/PCS group have different seeds. Cross-group FID differences summarize the selected configurations, while the FP32 group provides the paired comparison. Both PCS evaluations use $(m, s) = (1.3, 4)$ and are reported separately.

Fixed guidance and Wang schedules. Fixed CFG uses $a_j = 1.125$. For the Wang family (Wang et al., 2024), write $\tau_j = j/199$ and let m denote the discrete mean coefficient:

$$\begin{aligned} a_j^{\text{linear}} &= 1 + 2(m-1)\tau_j, \quad a_j^{\text{clamp}} = \max\{1.01, 1 + 2b\tau_j\}, \\ r_j &= \tfrac{1}{2}[1 - \cos(\pi\tau_j^s)], \quad a_j^{\text{PCS}} = 1 + (m-1)r_j \Big/ \Big(\tfrac{1}{200}\sum_{q=0}^{199} r_q\Big). \end{aligned} \tag{J.3}$$

Clamp-linear calibrates b by 64 bisection iterations to attain mean m. The selected PCS has $(m, s) = (1.3, 4)$ and peak coefficient 2.668277.

β-CFG adaptation. The VP/DDIM adaptation of Malarz et al. (2025) uses the area-normalized Beta density and the prediction difference over all 3,072 image coordinates:

$$a_{j,i} = \omega \frac{f_{\text{Beta}(u,v)}(\tau_j)}{\max(\|\epsilon^c_{j,i} - \epsilon^u_{j,i}\|_2, 10^{-8})^\rho}, \qquad \omega = \frac{m_{\text{pilot}}}{C}. \tag{J.4}$$

Here C averages the fraction in Eq. (J.4) over a 500-image conditional-only pilot with seed 20260913. The selected $(u, v) = (3, 2)$, $\rho = 0.5$, and $m_{\text{pilot}} = 1.15$ give $C = 1.1207457511$ and $\omega = 1.0261024848$. The coefficient has no additive one or cap; both endpoints have $a = 0$. Sampling uses the common clipped DDIM map, not a CFG++ update.

Feedback Guidance adaptation. FBG (Koulischer et al., 2025) initializes $L_{0,i} = 0$ for each batch. With $\pi = 0.95$ and $L_{\min} = \log\{(1-\pi)W/(W-1)\}$, it clips L to $[L_{\min}, 3]$ and sets $a_{j,i} = \{1 - (1-\pi)\exp(-L_{j,i})\}^{-1}$. Using A_j, B_j, h_j, and T_j from Eq. (J.2), define

$$\begin{aligned} \mu_j(e) &= \frac{\sqrt{B_j}(1 - A_j/B_j)}{1 - A_j} h_j(e) + \frac{\sqrt{A_j/B_j}(1 - B_j)}{1 - A_j} x_j, \\ v_j^{\text{VP}} &= (1 - A_j/B_j)\frac{1 - B_j}{1 - A_j}, \qquad x_{j+1} = T_j(\widehat{\epsilon}_j). \end{aligned} \tag{J.5}$$

For $j < 199$, the realized next state updates the observer through

$$\begin{aligned} \ell_{j,i} &= \frac{\langle \mu_j(\epsilon^c_{j,i}) - \mu_j(\epsilon^u_{j,i}),\, x_{j+1,i} - [\mu_j(\epsilon^u_{j,i}) + \mu_j(\epsilon^c_{j,i})]/2 \rangle}{\max(v_j^{\text{VP}}, 10^{-12})}, \\ L_{j+1,i} &= \text{clip}_{[L_{\min}, 3]}(L_{j,i} + \theta\ell_{j,i} + b_{\text{offset}}). \end{aligned} \tag{J.6}$$

The inner product sums over coordinates without dimension normalization. This positive-variance Gaussian observer is a VP approximation, not the transition density of deterministic DDIM. The shared conditional pilot gives the median positive ℓ over $\tau \in [0.2, 0.8]$ as $S = 0.034792823666$. Calibration uses $b_{\text{offset}} = \log\{(1-\pi)3/2\}/(200f)$ and $\theta = g|b_{\text{offset}}|/S$. The selected $(W, f, g) = (3, 0.5, 2.5)$ gives $b_{\text{offset}} = -0.025902671654$ and $\theta = 1.8612079249$.

The accompanying `baselines/` directory preserves the selected parameters, full-precision summaries, all 12 seed records, and the four baseline implementations. Its validation script recomputes the reported means and sample standard deviations from the seed records.

Table J.2: Auxiliary metrics for the four FID-selected baseline configurations in their common FP16 evaluation group. Entries are means ± sample SD over three 50,000-image runs. These metrics supplement the FID comparison.

Method	Agreement (%)	Precision	Recall
Fixed CFG	95.72 ± 0.12	0.6844 ± 0.0022	0.5778 ± 0.0016
Wang PCS	98.34 ± 0.06	0.6873 ± 0.0033	0.5784 ± 0.0022
β-CFG	94.25 ± 0.06	0.6950 ± 0.0028	0.5648 ± 0.0048
FBG	98.08 ± 0.04	0.7301 ± 0.0010	0.5154 ± 0.0019

Table J.3: All 12 final runs for the four selected baselines. Each row uses 50,000 images; seeds are paired within this table. Network inference uses FP16, with FP32 guidance and DDIM updates.

Method	Seed	FID ↓	Agreement (%)	Precision	Recall
Fixed CFG	20260921	2.9510	95.59	0.6822	0.5760
Fixed CFG	20260922	2.9146	95.80	0.6842	0.5793
Fixed CFG	20260923	2.9765	95.78	0.6867	0.5780
Wang PCS	20260921	2.6012	98.28	0.6837	0.5778
Wang PCS	20260922	2.5793	98.39	0.6880	0.5809
Wang PCS	20260923	2.6339	98.35	0.6902	0.5766
β-CFG	20260921	3.0640	94.24	0.6933	0.5657
β-CFG	20260922	3.0691	94.21	0.6934	0.5690
β-CFG	20260923	3.1267	94.32	0.6982	0.5596
FBG	20260921	4.5163	98.12	0.7313	0.5141
FBG	20260922	4.5125	98.06	0.7296	0.5176
FBG	20260923	4.5505	98.05	0.7294	0.5146

J.4 Online-K estimator and execution order

Online-K estimates this evidence by integrating the density-ratio differential along the sampled trajectory, including its divergence contribution. For variance-preserving diffusion, let α_j denote the cumulative signal power at the jth network evaluation and let x_j be the actual DDIM state. The raw conditional and unconditional predictions define the learned scores and correction

$$\begin{aligned} s_{u,j} &= -\epsilon_{u,j}/\sqrt{1-\alpha_j}, \quad s_{c,j} = -\epsilon_{c,j}/\sqrt{1-\alpha_j}, \\ d_j &= s_{c,j} - s_{u,j}, \qquad \widehat{C}_j = \widehat{\mathrm{div}} d_j + \langle x_j + s_{c,j} + s_{u,j}, d_j \rangle. \end{aligned} \tag{J.7}$$

For exact VP scores, $dK = \langle d, dx \rangle - \frac{1}{2} C\, d \log \bar{\alpha}$. Applying endpoint trapezoidal quadrature gives

$$\begin{aligned} \widehat{K}_j = \widehat{K}_{j-1} &+ \tfrac{1}{2}\langle d_{j-1} + d_j, x_j - x_{j-1} \rangle \\ &- \tfrac{1}{4} \log\left(\frac{\alpha_j}{\alpha_{j-1}}\right) (\widehat{C}_{j-1} + \widehat{C}_j), \qquad j \geq 1. \end{aligned} \tag{J.8}$$

We initialize $\widehat{K}_0 = 0$ as a finite-noise approximation. Inner products sum over all image coordinates, and $\widehat{K}$ is neither normalized nor clipped. A fresh Rademacher probe z estimates each endpoint divergence as $z^\top J_d z$, where $J_d = \partial d / \partial x$, using one input vector–Jacobian product (VJP). The raw scores are evaluated before the shared DDIM clipping operation; the displacement in (J.8) is taken from the resulting sampled states.

At endpoint (x_j, u_j), where $u_j = \log A_j$, the raw epsilon predictions define scores $s_j^c = -\epsilon_j^c/\sqrt{1-A_j}$ and $s_j^u = -\epsilon_j^u/\sqrt{1-A_j}$. Write $d_j = s_j^c - s_j^u$ and let $J_j = \partial d_j / \partial x_j$. The divergence estimate uses independent Rademacher probes $z_{j,r} \in \{-1, +1\}^{3072}$:

$$\widehat{D}_{R,j} = \frac{1}{R} \sum_{r=1}^{R} z_{j,r}^\top J_j z_{j,r}, \qquad J_j^\top z_{j,r} = \mathrm{VJP}(d_j, x_j; z_{j,r}). \tag{J.9}$$

This estimate enters $\widehat{C}_j = \widehat{D}_{R,j} + \langle x_j + s_j^c + s_j^u, d_j \rangle$ and the trapezoidal update in Eq. (J.8). Network parameters remain frozen; differentiation is only with respect to the current input. The

implementation stores detached endpoint states, score differences, and $\widehat{C}_j$, without differentiating through the sampling trajectory or constructing a Hessian. VJP products and spatial inner products are reduced in FP64 after the FP32 network and VJP calculations. The accumulated evidence is used directly, without clipping, dimensional normalization, or batch standardization.

Sampling procedure. For each image, the following order completes the previous sampling interval before choosing the current guidance coefficient:

1. Initialize x_0 from the noise stream, $\widehat{K}_0 = 0$, and an empty previous-endpoint cache.
2. At $j = 0, \ldots, 199$, evaluate the raw conditional and unconditional predictions at x_j and compute $d_j, \widehat{C}_j$ with fresh probes.
3. If $j > 0$, update $\widehat{K}_j$ by Eq. (J.8), using the cached endpoint and the actual displacement $x_j - x_{j-1}$.
4. Evaluate a_j from the updated $\widehat{K}_j$ using Eq. (23). Mix the epsilon predictions and take the clipped DDIM step in Eq. (J.2).
5. Cache the detached current endpoint and proceed to the next step.

No additional endpoint is evaluated after the final transition. Consequently, the last recorded evidence is $\widehat{K}_{199}$, the value used for the final gate, rather than an estimate at the output image. Execution checks verify fresh probe streams, image-ID invariance across batching, frozen parameters, detached caches, and this update-before-gate order.

Computational cost. Each trajectory uses 200 joint conditional/unconditional forward calls, equivalent to 400 branch predictions, for both methods. Online-K adds $200R$ input VJPs and 199 evidence updates; the final experiments use $R = 1$. PCS adds neither operation. Thus equal forward-call counts do not imply equal computational cost. The final runs shared GPU resources, so their recorded generation times do not establish a controlled speed ratio.

J.5 Paired FP32 results

The FID-selected Online-K configuration and PCS are evaluated on three paired seeds under the protocol in Table J.1. Table J.4 gives the six runs. Online-K lowers mean FID from 2.611 to 2.562, a 1.90% reduction, with paired difference -0.0497 ± 0.0033 and lower FID in every seed (Table J.5). Means and sample standard deviations describe sampling variation for the frozen checkpoint; relative FID changes divide the difference of means by the PCS mean. The additional parameter setting is reported in Appendix J.9.

Table J.4: Paired final runs for the FID-selected Online-K configuration and PCS. Each row uses 50,000 generated images and the CIFAR-10 train50K reference. Accuracy is classifier agreement.

Method	Seed	FID	Acc. (%)	Precision	Recall
Online-K ($\lambda = 2, \kappa = 4$)	20260928	2.556290	97.396	0.68504	0.57774
Online-K ($\lambda = 2, \kappa = 4$)	20260929	2.551244	97.488	0.68792	0.57806
Online-K ($\lambda = 2, \kappa = 4$)	20260930	2.577686	97.542	0.68548	0.57822
PCS	20260928	2.609664	98.212	0.68802	0.57576
PCS	20260929	2.598284	98.374	0.69028	0.57418
PCS	20260930	2.626385	98.416	0.69062	0.57626

Table J.6 reports classifier agreement and feature precision/recall alongside FID. Online-K has agreement of 97.475%, versus 98.334% for PCS, with slightly lower precision and higher recall.

Table J.5: Paired Online-K minus PCS FID differences across the three shared seeds, reported as mean $\pm$ sample SD.

Method − PCS	ΔFID (mean ± SD)	Relative change	Seeds with lower FID
Online-K ($\lambda = 2, \kappa = 4$)	-0.0497 ± 0.0033	-1.90%	3/3

Table J.6: Secondary metrics for the paired comparison, mean $\pm$ sample SD across three seeds. Accuracy is in percent. Mean a averages guidance over images, steps, and seeds.

Method	Acc. (%)	Precision	Recall	Mean a
Online-K ($\lambda = 2, \kappa = 4$)	97.475 ± 0.074	0.6861 ± 0.0016	0.5780 ± 0.0002	1.319
PCS	98.334 ± 0.108	0.6896 ± 0.0014	0.5754 ± 0.0011	1.300

J.6 Guidance trajectories

Figure J.1 compares the guidance trajectories of FID-selected Online-K and PCS. Both concentrate additional guidance toward the data end of the trajectory. Within each run, batch means are weighted by image count; the three equally sized runs are then averaged. The faint curves show seed means.

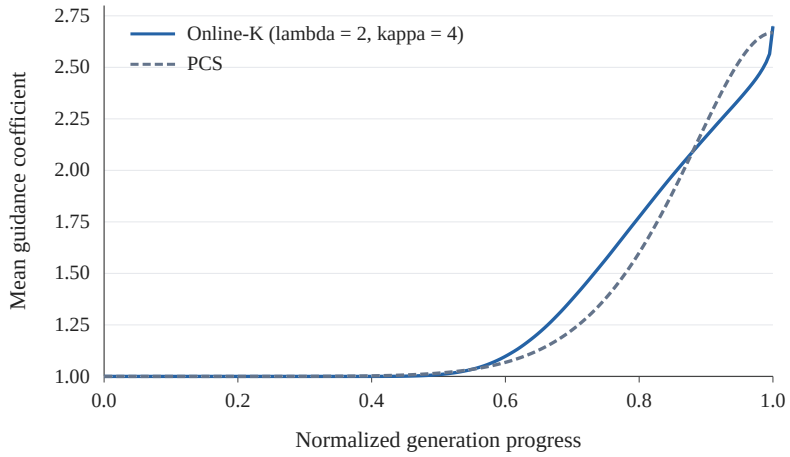

Figure J.1: Mean guidance coefficients along the 200 evaluation endpoints. Faint lines show individual-seed means; thick lines average the three seeds. Each seed mean uses all 50,000 images.

J.7 Estimator error bounds and analytic validation

J.7.1 Density-ratio identity and error decomposition

For consistent conditional and unconditional VP densities, let $K^*(x, u) = \log[p_c(x, u)/p_u(x, u)]$, where $u = \log \bar{\alpha}$. Subtracting the conditional and unconditional VP log-density equations gives $\partial_u K^* = -C^*/2$. The chain rule therefore yields

$$dK^* = \langle d^*, dx \rangle - \tfrac{1}{2} C^* \, du, \quad d^* = s_c^* - s_u^*, \quad C^* = \operatorname{div} d^* + \langle x + s_c^* + s_u^*, d^* \rangle. \tag{J.10}$$

Equation (J.8) discretizes the spatial and temporal terms at the actual DDIM endpoints. The following decomposition separates initialization, quadrature, learned-field, random-trace, and floating-point contributions.

On the same realized trajectory, let $I_j(d, C)$ denote the exact accumulated line integral and $Q_j(d, C)$ the accumulated endpoint quadrature. Then

$$\widehat{K}_j - K^*(x_j, u_j) = \underbrace{\widehat{K}_0 - K^*(x_0, u_0)}_{\text{initialization}} + \underbrace{Q_j(d^*, C^*) - I_j(d^*, C^*)}_{\text{quadrature}} + \underbrace{Q_j(d_\theta, C_\theta) - Q_j(d^*, C^*)}_{\text{learned scores}} + \underbrace{Q_j(d_\theta, \widehat{C}_\theta) - Q_j(d_\theta, C_\theta)}_{\text{random trace}} + E_{\text{fp}}. \tag{J.11}$$

Here C_θ uses the exact Jacobian trace of the learned field and $\widehat{C}_\theta$ uses its probe estimate. The initial value $\widehat{K}_0 = 0$ is approximate at finite noise, and learned conditional and unconditional scores need not correspond to a common VP mixture. The analytic mixture below isolates quadrature error with known densities; Appendix J.8 reports fixed-trajectory diagnostics for the learned image model.

J.7.2 Quadrature and learned-field contributions

Write $z_i = (x_i, u_i)$, $h_i = u_i - u_{i-1} > 0$, $U = \sum_i h_i$, and $h_{\max} = \max_i h_i$. Suppose $G(z) = K^*(x, u)$ is C^3 along the endpoint line segments, its third derivative has operator norm at most M_3, and $\|x_i - x_{i-1}\| \le L h_i$. The line-segment trapezoidal rule gives

$$|e_{\text{quad},i}| \le \frac{M_3}{12}\|z_i - z_{i-1}\|^3, \qquad |E_{\text{quad}}| \le \frac{M_3}{12}(1 + L^2)^{3/2} U h_{\max}^2. \tag{J.12}$$

The smoothness condition concerns the density ratio along the reached endpoint segments; it does not require differentiability of the clipping map. If, in addition, the endpoint fields satisfy $\|d_\theta - d^*\| \le \varepsilon_d$ and $|C_\theta - C^*| \le \varepsilon_C$, then

$$|E_{\text{model}}| \le U\left(L\varepsilon_d + \frac{\varepsilon_C}{2}\right). \tag{J.13}$$

These bounds control evidence error under the stated field and smoothness assumptions. Their constants have not been estimated for the CIFAR-10 model.

Analytic integration test. We test the quadrature on 4,096 trajectories of the VP mixture $\frac{1}{2}\mathcal{N}(\sqrt{\alpha}, 1) + \frac{1}{2}\mathcal{N}(-\sqrt{\alpha}, 1)$, using exact scores and divergence, analytic initial evidence, constant guidance 3, and no clean-image clipping. Over $u \in [-3, -1]$, increasing the number of intervals from 25 to 400 reduces endpoint RMSE from 8.60×10^{-5} to 3.31×10^{-7}, with observed orders between 2.001 and 2.011 (Table J.7). With exact initial evidence, scores, and divergence, this experiment isolates the quadrature contribution in Eq. (J.11). The observed order agrees with the $O(h_{\max}^2)$ bound in Eq. (J.12). Figure J.2 shows the convergence curve.

Table J.7: Endpoint integration error for the analytic VP mixture. Scores, divergence, and initial evidence are exact.

DDIM intervals N	Endpoint RMSE	Observed order
25	8.597×10^{-5}	–
50	2.133×10^{-5}	2.0106
100	5.315×10^{-6}	2.0051
200	1.326×10^{-6}	2.0025
400	3.313×10^{-7}	2.0012

J.7.3 Random trace contribution

At a reached state independent of its fresh probe, let $S = (J + J^\top)/2$. Independent Rademacher probes satisfy

$$\mathbb{E}[\widehat{D}_R \mid x, u] = \operatorname{tr} J, \qquad \operatorname{Var}(\widehat{D}_R \mid x, u) = \frac{1}{R}\sum_{i<\ell}(J_{i\ell} + J_{\ell i})^2 = \frac{2}{R}\|\operatorname{offdiag} S\|_F^2. \tag{J.14}$$

On a deterministic time grid, write $\epsilon_i = \widehat{D}_{R,i} - \operatorname{tr} J_i$. Across n completed intervals, its contribution to the evidence is

$$E_{\text{trace}} = -\sum_{i=0}^{n} w_i \epsilon_i, \qquad w_0 = \frac{h_1}{4}, \quad w_n = \frac{h_n}{4}, \quad w_i = \frac{h_i + h_{i+1}}{4} \quad (0 < i < n). \tag{J.15}$$

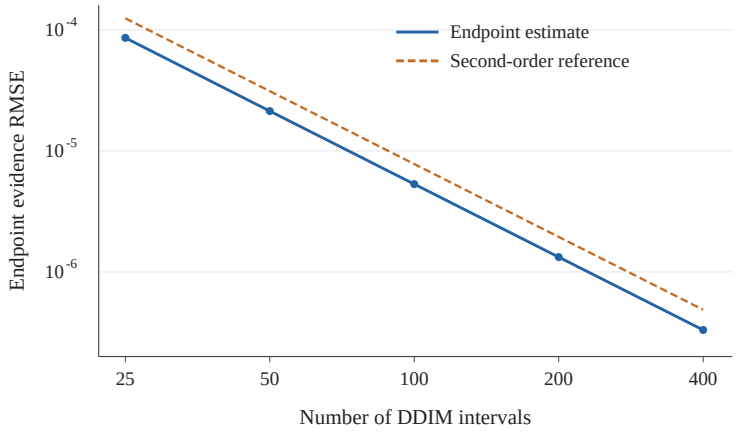

Figure J.2: Evidence-integration RMSE on the analytic VP mixture, with a second-order reference slope. Each grid uses 4,096 trajectories with exact scores, divergence, and initial evidence.

With fresh independent probes, these centered errors are martingale differences: previous probes may affect future states, but the probe at each reached state remains independent of its past. If the conditional single-probe variance is uniformly bounded by V over reached states, then

$$\mathbb{E}[E_{\mathrm{trace}}] = 0, \qquad \mathrm{Var}(E_{\mathrm{trace}}) \le \frac{V}{R}\sum_i w_i^2 \le \frac{VUh_{\max}}{4R}, \qquad \mathrm{RMS}(E_{\mathrm{trace}}) \le \frac{1}{2}\sqrt{\frac{VUh_{\max}}{R}}. \tag{J.16}$$

Equation (J.16) controls the additive trace term in Eq. (J.11); the uniform variance bound V has not been estimated for the image runs. Within-endpoint sample variance requires $R > 1$ and is recorded as unavailable for $R = 1$.

Trace validation. The trace check enumerates all eight Rademacher sign vectors in dimension three. The exact trace and the estimated mean agree to approximately 2.24×10^{-8}. The population variance is approximately 0.77, consistent with Eq. (J.14); the Bessel-corrected variance of the eight enumerated values is approximately 0.88 because of the factor $8/7$ (Table J.8). This checks the trace implementation without estimating its variance on CIFAR-10.

Table J.8: Three-dimensional trace validation. Population variance and Bessel-corrected sample variance use different denominators.

Validation quantity	Reference	Observed	Absolute difference
Trace: full 8-probe enumeration	0.900000036	0.900000058	2.235×10^{-8}
Single-probe population variance	0.769999862	0.769999981	1.192×10^{-7}
Returned variance ($ddof = 1$)	0.879999842	0.879999893	5.143×10^{-8}
Maximum absolute error in C	0	6.043×10^{-8}	6.043×10^{-8}

J.8 Learned-evidence calibration on fixed trajectories

We evaluate $\widehat{K}$ on 128 fixed PCS trajectories with nested probe sets $R = 1, 2, 4$. The guidance policy is fixed, so all settings produce identical image tensors. At the last gating endpoint, mean evidence lies between 18.1 and 18.2, and 99.219–100% of estimates exceed $\log 10$ (Table J.9). Increasing the probe count leaves this discrepancy in place.

For a consistent uniform ten-class mixture, $K = \log p(c \mid x, t) - \log p(c) \le \log 10$. The learned estimate therefore serves as an uncalibrated control signal, and the selected $\kappa = 4$ is an empirical threshold on its scale, rather than a posterior-probability cutoff. The gate remains bounded as specified in Appendix D.3.

Figure J.3 reports RMS disagreement with the random $R = 4$ estimate on the same trajectories. It measures sensitivity to probe count; $R = 4$ is a comparison reference, not a ground-truth density ratio. The separate contributions in Eq. (J.11) are unavailable for these learned densities.

Table J.9: Evidence diagnostics on 128 fixed PCS trajectories, without feedback. The last logged value precedes the final DDIM transition.

R	Last-gate mean K	Last-gate median K	$K > \log 10$ (%)	RMS difference vs. $R = 4$
1	18.2287	13.8969	99.219	1.4835
2	18.1946	13.3777	100.000	0.8510
4	18.1221	13.3040	100.000	0.0000

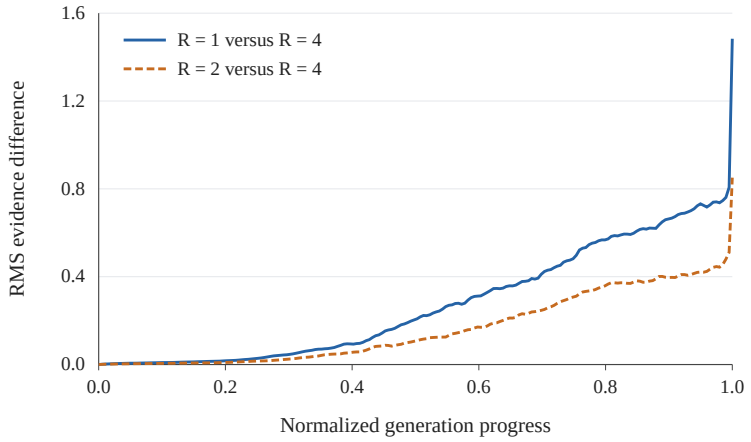

Figure J.3: RMS disagreement of $R = 1, 2$ evidence with the $R = 4$ estimate on 128 identical fixed PCS trajectories. Probe sets are nested and evidence is observed without feedback.

J.9 Additional selection results

Table J.10 contains all six Online-K development settings and the fixed PCS reference. Minimizing development FID selects $(\lambda, \kappa) = (2, 4)$. Adding the requirement of at least 98.2% development classifier agreement selects $(1, 1)$ instead. Both use $\gamma = 4$ and $R = 1$. The latter setting is reported here under its distinct selection objective.

Table J.10: Complete development results: 10,000 images per setting, seed 20260927. The two Online-K selections use FID and agreement-constrained FID, respectively; PCS is fixed.

Family	Parameters	FID	Acc. (%)	Selection
Online-K	$\lambda = 1,\ \gamma = 4,\ \kappa = 1$	4.8645	98.44	FID with agreement $\geq 98.2\%$
Online-K	$\lambda = 1,\ \gamma = 4,\ \kappa = 2$	4.6044	97.71	
Online-K	$\lambda = 1,\ \gamma = 4,\ \kappa = 4$	4.6657	96.66	
Online-K	$\lambda = 2,\ \gamma = 4,\ \kappa = 1$	6.0184	99.19	
Online-K	$\lambda = 2,\ \gamma = 4,\ \kappa = 2$	4.9305	98.65	
Online-K	$\lambda = 2,\ \gamma = 4,\ \kappa = 4$	4.5428	97.68	Minimum FID
PCS	$m = 1.3,\ s = 4$	4.5878	98.43	fixed PCS reference

The mean guidance coefficient is 1.487. The paired FID difference from PCS is $+0.3909 \pm 0.0281$ ($+14.97\%$), with higher FID in all three seeds. These results document the alternative selection objective; the FID-selected comparison is given in Appendix J.5.

Table J.11: Final results for the agreement-constrained selection, Online-K $(\lambda, \kappa) = (1, 1)$. Each seed uses 50,000 images and the same FP32 protocol and paired PCS runs as Table J.4. The aggregate row reports mean $\pm$ sample SD.

Seed	FID	Acc. (%)	Precision	Recall
20260928	2.969201	98.002	0.70964	0.55640
20260929	2.997515	98.184	0.70956	0.55472
20260930	3.040322	98.212	0.71054	0.55714
Aggregate	3.0023 ± 0.0358	98.133 ± 0.114	0.7099 ± 0.0005	0.5561 ± 0.0012

J.10 Paired qualitative examples

Figure J.4 shows the first 20 global sample indices of seed 20260928, giving two examples per class. Identical indices share initial noise and target labels across methods. The display uses a fixed index rule; the quantitative comparison uses every generated image in all three final seeds.



Figure J.4: Paired generated samples, two per class, from the first 20 sample indices of seed 20260928. Columns follow the ten CIFAR-10 classes; rows identify the configuration and example.

K ImageNet: DiT guidance evaluation

K.1 Protocol and method settings

We use the pretrained DiT-XL/2 checkpoint for class-conditional ImageNet generation at 256×256 resolution (Peebles & Xie, 2023). All final runs use deterministic DDIM with 50 sampling steps ($\eta = 0$), drawn from the original 1,000-step diffusion grid. CFG is applied to the first three latent epsilon channels, following the DiT implementation. The coefficient convention is $\widehat{\epsilon} = \epsilon_u + a(\epsilon_c - \epsilon_u)$, so fixed $w = 1$ is conditional prediction.

PCS and Online-K use the EMA VAE. The Online-K variant uses a batch-shared expected-energy proxy to set the guidance scale, with peak $W = 1 + \lambda$ and gate parameters (γ, κ). It does not use the per-sample divergence/VJP estimator of Appendix J.4. The paired 50K comparison uses seed 1 and batch size eight. Parameters were frozen after the paper's 1K-image development

screens: PCS searched 12 settings, $W \in \{3, 3.8, 5\}$ and $p \in \{1, 2, 4, 6\}$; Online-K searched 12 settings, $W \in \{2, 3, 5\}$, $\gamma \in \{1, 2\}$, and $\kappa \in \{1, 1.5\}$. The selected settings are $(W, p) = (3, 2)$ and $(W, \gamma, \kappa) = (3, 1, 1.5)$, respectively. Both reported screens contain 12 settings, with different parameter grids. The repository's 36-setting Online-K smoke test is a separately logged run and is not the 12-setting search count reported here.

The Fixed CFG rows in Table K.1 use the EMA VAE, seed 1, batch size eight, and the same DiT checkpoint and DDIM protocol as PCS and Online-K. They report $w \in \{1, 1.5, 2\}$. Thus, all rows in Table K.1 share the seed and evaluation protocol. The separate 1K fixed-guidance development screen in Table K.2 used the MSE VAE and is reported only as screening; it is not the Fixed CFG protocol used for the 50K table. All final metrics use 50,000 generated images and ImageNet validation 50K as reference, with `pytorch-fid` Inception-2048 features and feature precision/recall at $k = 3$. The extreme-pixel fraction is the fraction of extreme RGB components, rather than HSV saturation.

K.2 Final evaluation and fixed-guidance screening

Table K.1: ImageNet DiT-XL/2, EMA VAE, DDIM-50, 50K generated images, seed 1, batch size 8. All rows use the same seed and evaluation protocol.

Method	Parameters	FID ↓	Prec. ↑	Rec. ↑	Extreme frac. ↓
PCS	$W = 3,\ p = 2$	3.693297	0.70212	0.61776	0.034722
Online-K	$W = 3,\ \gamma = 1,\ \kappa = 1.5$	3.515127	0.68026	0.61138	0.037808
Fixed CFG	$w = 1$	8.047363	0.56856	0.67948	0.032823
Fixed CFG	$w = 1.5$	3.785053	0.71184	0.58892	0.039982
Fixed CFG	$w = 2$	7.874961	0.78584	0.50064	0.049111

In the paired EMA comparison, the selected Online-K variant has FID 3.515127, a 4.82% reduction relative to PCS at 3.693297. Its feature precision and recall are lower, locating a different operating point across these metrics. ResNet50 V2 agreement against requested labels is available for both methods: Online-K has top-1/top-5 accuracy of $83.462\%/95.452\%$, and PCS has $88.518\%/97.530\%$. These are single-seed results, not multi-seed averages. Among the three seed-1 EMA Fixed CFG rows in Table K.1, $w = 1.5$ gives the lowest FID. Separately, the MSE-VAE 1K screening in Table K.2 gives the lowest FID at $w = 1.5$ among its five tested scales; its values should not be compared directly with the EMA 50K table.

Table K.2: Fixed CFG development screening with the MSE VAE: 1,000 generated images per scale, with ImageNet validation 50K as reference. These screening values are separate from the EMA 50K results.

w	FID ↓	Precision ↑	Recall ↑	Extreme fraction ↓
1	49.860651	0.583000	0.711060	0.027390
1.5	44.017629	0.715000	0.731380	0.034151
2	48.507447	0.777000	0.727700	0.043650
3	56.759925	0.850000	0.680780	0.065359
4	60.279377	0.828000	0.710500	0.088136

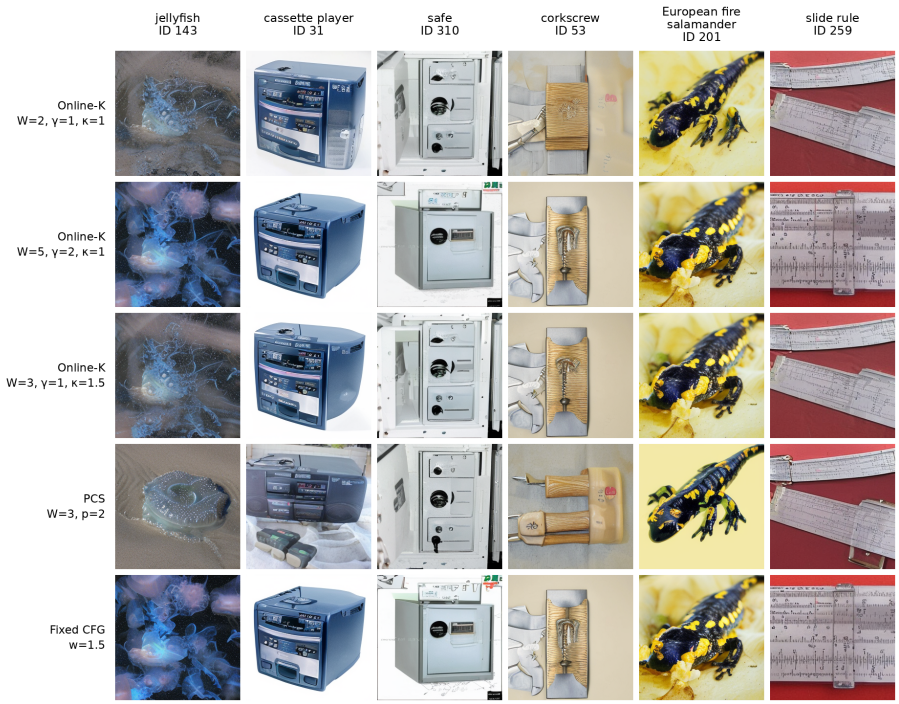

Figure K.1: Qualitative ImageNet comparison. Rows show three Online-K configurations, PCS, and Fixed CFG with $w = 1.5$. Columns show six distinct seed-0 sample IDs, selected from the 1K display pool for high cross-method pixel differences; the examples are illustrative rather than randomly sampled and do not replace the 50K evaluation in Table K.1.