OSMOTIC GUIDANCE FOR GENERATIVE FLOWS

Anonymous authors
Paper under double-blind review

ABSTRACT

This paper studies guidance for flow-based generative models from the perspective of intervention efficiency, an important aspect that is often less explicitly considered in existing guidance research. To address the question of how to achieve the same instantaneous guidance effect with minimal modification to the pretrained dynamics, we formulate guidance as a local constrained optimization problem that minimizes intervention energy subject to a prescribed objective-change rate, yielding a unique canonical minimum-energy solution. Using the marginal equivalence between deterministic and stochastic dynamics, we further show that this solution admits a relative osmotic representation. This leads to osmotic guidance (OG), a training-free framework that provides a unified view of guidance and supports diverse guidance objectives through objective-induced osmotic fields with a shared energy-calibration rule. We instantiate OG for entropy, reward, target-density, and conditional guidance, and introduce a two-anchor preview for efficient strength selection. Experiments on synthetic flow models and image datasets demonstrate effective guidance, improved intervention efficiency, and empirical support for the proposed canonical minimum-energy principle.

1 INTRODUCTION

Generative modeling aims to learn a model distribution p_θ that approximates a target data distribution p_{data} (Lipman et al., 2023; Song et al., 2021). However, the learned distribution is not necessarily suitable for every downstream task, and retraining or post-training takes extra resources (Bansal et al., 2023; Yu et al., 2023). Specifically, at inference time, one may instead prefer samples that satisfy a particular condition, maximize a reward (Song et al., 2023), predominate around high/low-density regions (Xu et al., 2026; Li et al., 2026a), or move toward another target distribution (Feng et al., 2025). This motivates guidance: modifying a pretrained generative model at sampling time to favor a desired property without retraining the model (Dhariwal & Nichol, 2021; Ho & Salimans, 2022; Bansal et al., 2023). For flow-based generative models, where generation is naturally described as probability transport through a velocity field $\boldsymbol{v}_t$ (Lipman et al., 2023; Albergo et al., 2025), guidance can therefore be introduced by adding an intervention field $\boldsymbol{g}_t$ (Feng et al., 2025) as shown in Eq. 7. Existing guidance methods construct $\boldsymbol{g}_t$ in different ways, from classifier gradients (Dhariwal & Nichol, 2021), conditional signals (Ho & Salimans, 2022), rewards (Song et al., 2023), or approximations to guided flow dynamics (Feng et al., 2025; Wang et al., 2025).

Adding a guidance field inevitably perturbs the native dynamics of the pretrained generative model. Recent studies have shown that strong guidance can cause sampling trajectories to deviate substantially from the original generative dynamics (Chung et al., 2025; Chidambaram et al., 2024). Such deviations may alter properties that are not specified by the guidance objective and can degrade sample quality (Sadat et al., 2025a; Kynkäänniemi et al., 2024); we further discuss the issue in Section 3.1 through the link between intervention energy and terminal distributional deviation, and Fig. 1 and Table 2 illustrate this effect empirically. However, much of the guidance literature focuses on improving sample quality or alignment with the desired condition, the amount of intervention required to achieve that effect is typically less explicitly explored. This motivates a complementary question: *for a prescribed guidance effect, can we realize it with a smaller intervention?*

In this work, we address this question by formulating inference-time guidance as a local constrained optimization problem. Given an objective of guidance $\mathcal{J}_t[\rho]$ for a distribution ρ at time t, we seek a guidance field that minimizes its intervention energy while enforcing a prescribed instantaneous objective-change rate. Under mild regularity conditions, this problem admits a unique canonical

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.

solution in $L^2(\rho_t)$, identifying the minimum-energy intervention for the desired instantaneous guidance effect. We then connect this canonical solution to the density-dependent structure of probability transport. By exploiting the marginal equivalence between deterministic and stochastic dynamics (Song et al., 2021; Albergo et al., 2025), we show that the canonical guidance field can be expressed as a difference between osmotic fields. This osmotic representation turns the minimum-energy guidance problem into an objective-induced relative osmotic intervention.

Based on this formulation, we introduce osmotic guidance (OG), a training-free framework that realizes guidance through objective-induced relative osmotic fields with minimum-energy calibration. OG provides a unified view of guidance, which can be applied to common guidance tasks while naturally supporting a two-anchor preview for efficient strength selection. We evaluate OG in terms of guidance effectiveness, intervention efficiency, and direct validation of its underlying minimum-energy principle.

Our main contribution can be summarized as follows:

1) We formulate inference-time guidance as a local constrained optimization problem with a canonical minimum-energy osmotic field as its solution, and provide a unified view of guidance as relative osmotic interventions.
2) We present the osmotic guidance framework, and instantiate OG for multiple guidance objectives with a two-anchor preview for efficient strength selection.
3) We demonstrate the effectiveness and intervention efficiency of OG and empirically validate its canonical minimum-energy characterization.

2 Marginal equivalence and osmotic field

Probability transport. Let $\{\pi_t\}_{t\in[0,1]}$ be a smooth probability path connecting a tractable base distribution π_0 to a target distribution π_1 (Lipman et al., 2023; Albergo et al., 2025). A deterministic realization is given by $\frac{\mathrm{d}\mathbf{x}_t}{\mathrm{d}t} = \boldsymbol{v}_t(\mathbf{x}_t), \quad \mathbf{x}_t \sim \pi_t$. Its marginals π_t satisfy the continuity equation (Ambrosio et al., 2008; Lipman et al., 2023)

$$\partial_t \pi_t = -\nabla_{\boldsymbol{x}} \cdot (\pi_t \boldsymbol{v}_t)\,. \tag{1}$$

Suppose that the same marginal path is also realized by the stochastic dynamics $\mathrm{d}\mathbf{y}_t = \boldsymbol{b}_t(\mathbf{y}_t)\,\mathrm{d}t + \sqrt{2\kappa_t}\,\mathrm{d}\mathbf{w}_t$, where $\kappa_t \geq 0$ is a time-dependent diffusivity and $\mathbf{w}_t$ is a standard Wiener process. The corresponding density evolves according to the Fokker–Planck equation (Song et al., 2021; Bakry et al., 2014)

$$\partial_t \pi_t = -\nabla_{\boldsymbol{x}} \cdot (\pi_t \boldsymbol{b}_t) + \kappa_t \Delta_{\boldsymbol{x}} \pi_t. \tag{2}$$

Using $\Delta_{\boldsymbol{x}} \pi_t = \nabla_{\boldsymbol{x}} \cdot (\pi_t \nabla_{\boldsymbol{x}} \log \pi_t)$, the Fokker–Planck equation can be rewritten as

$$\partial_t \pi_t = -\nabla_{\boldsymbol{x}} \cdot [\pi_t \left(\boldsymbol{b}_t - \kappa_t \nabla_{\boldsymbol{x}} \log \pi_t\right)]\,. \tag{3}$$

Comparing Eq. 3 with Eq. 1, the standard probability-flow construction (Song et al., 2021; Albergo et al., 2025) gives the marginally equivalent deterministic velocity

$$\boldsymbol{v}_t = \boldsymbol{b}_t - \kappa_t \nabla_{\boldsymbol{x}} \log \pi_t. \tag{4}$$

Thus, a stochastic drift and its probability-flow velocity can generate the same one-time marginals while differing by a component determined by the marginal score $\nabla_{\boldsymbol{x}} \log \pi_t$ (Song et al., 2021; Hyvärinen, 2005).

Osmotic field. Motivated by this decomposition, for any smooth positive density p, we define its osmotic field(Nelson, 1967; Léonard, 2014) as

$$\boxed{\boldsymbol{d}_t[p](\boldsymbol{x}) := \kappa_t \nabla_{\boldsymbol{x}} \log p(\boldsymbol{x}).} \tag{5}$$

For the reference marginal π_t, Eq. 4 can therefore be written as $\boldsymbol{b}_t = \boldsymbol{v}_t + \boldsymbol{d}_t[\pi_t]$. Furthermore, it encodes the local density-dependent geometry of the marginal (Nelson, 1967; Bakry et al., 2014).

For two smooth positive densities p and q, their relative osmotic field (cf. Hyvärinen, 2005) is

$$\boldsymbol{d}_t[q](\boldsymbol{x}) - \boldsymbol{d}_t[p](\boldsymbol{x}) = \kappa_t \nabla_{\boldsymbol{x}} \log \frac{q(\boldsymbol{x})}{p(\boldsymbol{x})}. \tag{6}$$

Thus, differences between osmotic fields characterize how this density-dependent component changes between two marginals. This relative view provides the basis for the objective-induced osmotic intervention developed in Section 3.2.

3 OSMOTIC GUIDANCE AS MINIMUM-ENERGY DISTRIBUTIONAL CONTROL

We think of modifying a velocity $\boldsymbol{v}_t$ at inference time through an additive guidance field $\boldsymbol{g}_t$:

$$\frac{\mathrm{d}\mathbf{x}_t}{\mathrm{d}t} = \boldsymbol{v}_t(\mathbf{x}_t) + \boldsymbol{g}_t(\mathbf{x}_t), \quad \mathbf{x}_t \sim \rho_t, \tag{7}$$

where ρ_t denotes the resulting guided marginal. We denote the contribution of $\boldsymbol{g}_t$ to the density evolution by $\partial_t \rho_t|_{\boldsymbol{g}_t} = -\nabla_{\boldsymbol{x}} \cdot (\rho_t \boldsymbol{g}_t)$.

3.1 CANONICAL GUIDANCE

Let $\mathcal{J}_t[\rho]$ be a differentiable functional of ρ that describes a desired distributional property. Under vanishing boundary terms, the instantaneous change of $\mathcal{J}_t[\rho_t]$ caused by the guidance field $\boldsymbol{g}_t$ (Ambrosio et al., 2008; Jordan et al., 1998) is

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{J}_t[\rho_t]\bigg|_{\boldsymbol{g}_t} = \mathbb{E}_{\mathbf{x}_t \sim \rho_t}\left[\boldsymbol{g}_t(\mathbf{x}_t)^\top \nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}(\mathbf{x}_t)\right]. \tag{8}$$

This shows that $\nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}$ specifies the local transport direction favored by the objective (Ambrosio et al., 2008), whereas $\boldsymbol{g}_t$ is the guidance actually applied to the pretrained dynamics.

We measure the instantaneous effort of $\boldsymbol{g}_t$ by its instantaneous guidance energy (Benamou & Brenier, 2000; Ambrosio et al., 2008)

$$\mathcal{E}_t(\boldsymbol{g}_t) := \frac{1}{2}\mathbb{E}_{\mathbf{x}_t \sim \rho_t}\left[\|\boldsymbol{g}_t(\mathbf{x}_t)\|_2^2\right]. \tag{9}$$

Proposition 1 (Energy-controlled deviation from the unguided flow)**.** *Let p_t denote the unguided marginal, with $p_0 = \rho_0$. Assume both flows are well posed with finite second moments and $\|\boldsymbol{v}_t(\boldsymbol{x}) - \boldsymbol{v}_t(\boldsymbol{y})\|_2 \le L_t\|\boldsymbol{x} - \boldsymbol{y}\|_2$, where $L_t \ge 0$ and $\int_0^T L_t\, dt < \infty$. Coupling the guided and unguided trajectories $\mathbf{x}_t$ and $\mathbf{x}_t^0$ through the same initial sample, by Gronwall's inequality, the guidance energy gives and upper bound of the Wasserstein distance between p_T and ρ_T:*

$$W_2^2(\rho_T, p_T) \le \mathbb{E}\left[\|\mathbf{x}_T - \mathbf{x}_T^0\|_2^2\right] \le \left(\int_0^T e^{\int_s^T L_r dr}\sqrt{2\mathcal{E}_s(\boldsymbol{g}_s)}ds\right)^2. \tag{10}$$

Proposition 1 shows that the intervention energy controls an upper bound on the terminal deviation from the unguided distribution, providing a theoretical motivation for minimum-energy guidance. See proof in Appendix B.4.

For a prescribed instantaneous objective-change rate γ_t, which specifies how rapidly the guidance should change the distributional objective at time t, we define the canonical guidance field $\boldsymbol{g}_t^\star$ by

$$\boldsymbol{g}_t^\star \in \arg\min_{\boldsymbol{g}_t} \mathcal{E}_t(\boldsymbol{g}_t) \qquad \text{s.t.} \qquad \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{J}_t[\rho_t]\bigg|_{\boldsymbol{g}_t} = \gamma_t. \tag{11}$$

Note that this optimization is instantaneous at the current marginal ρ_t, and it does not optimize a global path-space control objective (Berner et al., 2024; Wang et al., 2025; Pandey et al., 2025). Under this formulation, guidance is cast as a constrained optimization problem: among all guidance fields that achieve the prescribed instantaneous objective-change rate γ_t, we seek the one that minimizes the required guidance energy $\mathcal{E}_t$ (cf. Boyd & Vandenberghe, 2004; Ames et al., 2017).

Proposition 2 (Canonical minimum-energy guidance)**.** *Suppose $0 < \mathbb{E}_{\rho_t}[\|\nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}\|_2^2] < \infty$. Then the unique minimum-energy solution in $L^2(\rho_t)$ is*

$$\boxed{\boldsymbol{g}_t^\star(\boldsymbol{x}) = \frac{\gamma_t}{\mathbb{E}_{\rho_t}\left[\left\|\nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}\right\|_2^2\right]} \nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}(\boldsymbol{x}).} \tag{12}$$

Moreover,

$$\mathcal{E}_t(\boldsymbol{g}_t) \ge \frac{\gamma_t^2}{2\mathbb{E}_{\rho_t}\left[\left\|\nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}\right\|_2^2\right]}, \tag{13}$$

with equality if and only if $\boldsymbol{g}_t$ is proportional to the objective-gradient field.

Proposition 2 says that, for a prescribed objective-change rate, canonical guidance minimizes the required guidance energy; conversely, for a prescribed energy budget, the same direction maximizes the instantaneous objective-change rate that can be achieved. In this sense, $\boldsymbol{g}_t^\star$ is the locally most energy-efficient guidance field for the chosen distributional objective. See proof in Appendix B.4.

3.2 Objective-induced osmotic guidance

Eq. 12 shows that the canonical guidance field contains two components: the distributional objective determines its direction through $\nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}$, while the minimum-energy principle determines how this direction is normalized.

Objective-induced osmotic direction. To characterize the direction selected by $\mathcal{J}_t$, consider the objective-tilted marginal (cf. Levine, 2018)

$$\rho_t^{\epsilon,\mathcal{J}}(\boldsymbol{x}) \propto \rho_t(\boldsymbol{x}) \exp\left(\epsilon \frac{\delta \mathcal{J}_t}{\delta \rho}(\boldsymbol{x})\right), \qquad \epsilon \neq 0. \tag{14}$$

Its score relative to the current marginal satisfies

$$\frac{1}{\epsilon}\left(\nabla_{\boldsymbol{x}} \log \rho_t^{\epsilon,\mathcal{J}}(\boldsymbol{x}) - \nabla_{\boldsymbol{x}} \log \rho_t(\boldsymbol{x})\right) = \nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}(\boldsymbol{x}). \tag{15}$$

Thus, the direction of the minimum-energy guidance corresponds to the relative score between the current marginal and its objective-tilted counterpart.

By the definition of osmotic field in Eq. 5, and the relative osmotic field in Eq. 6, Eq. 15 implies

$$\boldsymbol{d}_t^{\mathcal{J}}(\boldsymbol{x}) := \kappa_t \nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}(\boldsymbol{x}) = \frac{1}{\epsilon}\left(\boldsymbol{d}_t[\rho_t^{\epsilon,\mathcal{J}}](\boldsymbol{x}) - \boldsymbol{d}_t[\rho_t](\boldsymbol{x})\right). \tag{16}$$

We call $\boldsymbol{d}_t^{\mathcal{J}}$ the *objective-induced osmotic field*. It characterizes the relative change in osmotic structure from the current marginal toward the distribution favored by the objective. In this form, characterizing the objective direction reduces to constructing the corresponding objective-induced osmotic field.

Energy normalization. Eq. 12 normalizes this direction according to its energy under the current marginal. We therefore define the objective energy $\mathcal{E}_t^{\mathcal{J}}$ as the guidance energy of an objective-induced osmotic field $\boldsymbol{d}_t^{\mathcal{J}}$:

$$\mathcal{E}_t^{\mathcal{J}} := \mathcal{E}_t(\boldsymbol{d}_t^{\mathcal{J}}) = \frac{1}{2}\mathbb{E}_{\rho_t}\left[\left\|\boldsymbol{d}_t^{\mathcal{J}}\right\|_2^2\right] \tag{17}$$

Osmotic guidance. Combining the objective-induced osmotic direction in Eq. 16 with the objective energy in Eq. 17, the canonical solution in Eq. 12 reads

$$\boxed{\boldsymbol{g}_t^\star(\boldsymbol{x}) = \lambda_t^{\mathcal{J}} \boldsymbol{d}_t^{\mathcal{J}}(\boldsymbol{x}), \qquad \lambda_t^{\mathcal{J}} := \frac{\gamma_t \kappa_t}{2\mathcal{E}_t^{\mathcal{J}}}.} \tag{18}$$

Thus, the two components of canonical minimum-energy guidance acquire a direct osmotic interpretation. The objective $\mathcal{J}_t$ determines the direction through the objective-induced osmotic field $\boldsymbol{d}_t^{\mathcal{J}}$, while $\mathcal{E}_t^{\mathcal{J}}$ provides its minimum-energy normalization. The prescribed rate γ_t then determines the desired strength of distributional change. Consequently, canonical guidance takes the form of an energy-normalized osmotic intervention:

$$\frac{\mathrm{d}\mathbf{x}_t^{\mathcal{J}}}{\mathrm{d}t} = \boldsymbol{v}_t(\mathbf{x}_t^{\mathcal{J}}) + \boldsymbol{g}_t^\star(\mathbf{x}_t^{\mathcal{J}}), \qquad \mathbf{x}_t^{\mathcal{J}} \sim \rho_t^{\mathcal{J}}. \tag{19}$$

3.3 Two-anchor preview for strength selection

Once the objective is fixed, osmotic guidance reduces control to the scalar rate γ_t. Choosing its strength may still require costly evaluation of multiple guided trajectories (cf. Kynkäänniemi et al., 2024; Azangulov et al., 2026). We therefore use a two-anchor preview to approximate intermediate strengths from only two solved trajectories.

Table 1: Common instances of Osmotic Guidance. The objective $\mathcal{J}_t$ determines the corresponding objective-induced osmotic field $\boldsymbol{d}_t^{\mathcal{J}}$.

Objective	$\mathcal{J}_t[\rho]$	Objective-induced osmotic field $\boldsymbol{d}_t^{\mathcal{J}}$
Entropy	$-H(\rho) = \int \rho \log \rho \, \mathrm{d}\boldsymbol{x}$	$\kappa_t \nabla_{\boldsymbol{x}} \log \rho_t = \boldsymbol{d}_t[\rho_t]$
Reward	$\mathbb{E}_\rho[R_t(\mathbf{x})]$	$\kappa_t \nabla_{\boldsymbol{x}} R_t$
Target	$-D_{\mathrm{KL}}(\rho \Vert q_t)$	$\boldsymbol{d}_t[q_t] - \boldsymbol{d}_t[\rho_t]$
Conditional	$\mathbb{E}_\rho[\log p_t(\boldsymbol{y} \mid \mathbf{x})]$	$\boldsymbol{d}_t[p_t(\cdot \mid \boldsymbol{y})] - \boldsymbol{d}_t[p_t]$

Consider a one-parameter family $\gamma_t = \gamma\,\bar{\gamma}_t, \gamma \in [\gamma_a, \gamma_b]$, where $\bar{\gamma}_t$ is a fixed schedule, and denote by $\boldsymbol{x}_1^\gamma$ the corresponding terminal state. Given two solved anchor strengths γ_a and γ_b, we approximate an intermediate strength by

$$\widehat{\boldsymbol{x}}_1^\gamma = (1-\omega)\boldsymbol{x}_1^{\gamma_a} + \omega \boldsymbol{x}_1^{\gamma_b}, \qquad \omega = \frac{\gamma - \gamma_a}{\gamma_b - \gamma_a}. \tag{20}$$

If the terminal response $\gamma \mapsto \boldsymbol{x}_1^\gamma$ is twice continuously differentiable, standard interpolation theory (Trefethen, 2019) gives

$$\|\boldsymbol{x}_1^\gamma - \widehat{\boldsymbol{x}}_1^\gamma\|_2 \le \frac{(\gamma-\gamma_a)(\gamma_b-\gamma)}{2} \sup_{\xi \in [\gamma_a, \gamma_b]} \left\| \frac{\partial^2 \boldsymbol{x}_1^\xi}{\partial \xi^2} \right\|_2. \tag{21}$$

Thus, sufficiently local anchor pairs provide a principled approximation to intermediate terminal states. Rather than replacing the final guided solve, preview is used to screen candidate strengths before evaluating the selected operating point with the full dynamics.

4 Instantiations and Implementation of Osmotic Guidance

The formulation in Section 3 separates osmotic guidance into an objective-induced field $\boldsymbol{d}_t^{\mathcal{J}}$ and an energy-calibrated scalar strength $\lambda_t^{\mathcal{J}}$. We first instantiate $\boldsymbol{d}_t^{\mathcal{J}}$ for several common guidance objectives and describe how the corresponding osmotic fields are computed in practice. We then estimate the associated guidance energy using a common procedure to obtain the calibrated strength $\lambda_t^{\mathcal{J}}$.

4.1 Instantiations of Guidance Objectives

Different choices of $\mathcal{J}_t$ induce different osmotic fields through Eq. 16. Table 1 summarizes several useful instances.

Entropy and concentration. For the negative-entropy objective $\mathcal{J}_t^{\mathrm{ent}}[\rho] = -H(\rho)$, the first variation is $1 + \log \rho$ (Jordan et al., 1998; Ambrosio et al., 2008), and therefore

$$\boldsymbol{d}_t^{\mathrm{ent}} = \kappa_t \nabla_{\boldsymbol{x}} \log \rho_t = \boldsymbol{d}_t[\rho_t]. \tag{22}$$

Thus, concentration control is the special case in which the objective-induced osmotic field coincides with the ordinary osmotic field of the current marginal. Its objective energy further reduces to

$$\mathcal{E}_t^{\mathrm{ent}} = \frac{\kappa_t^2}{2} \mathcal{I}(\rho_t), \qquad \mathcal{I}(\rho_t) := \mathbb{E}_{\rho_t}\left[\|\nabla_{\boldsymbol{x}} \log \rho_t\|_2^2 \right], \tag{23}$$

where $\mathcal{I}(\rho_t)$ coincides with the Fisher information functional (Bakry et al., 2014; Gentil et al., 2020). The corresponding objective tilt is $\rho_t^{\epsilon,\mathrm{ent}} \propto \rho_t^{1+\epsilon}$, so positive guidance increases concentration in high-density regions (Xu et al., 2026). In practice, $\boldsymbol{d}_t^{\mathrm{ent}}$ is obtained from the marginal score of the generative model (Albergo et al., 2025; Ma et al., 2024).

Reward. For a differentiable reward R_t, $\mathcal{J}_t^R[\rho] = \mathbb{E}_\rho[R_t]$ directly gives

$$\boldsymbol{d}_t^R = \kappa_t \nabla_{\boldsymbol{x}} R_t, \tag{24}$$

with the corresponding tilt $\rho_t^{\epsilon,R} \propto \rho_t \exp(\epsilon R_t)$ (Levine, 2018). Hence the objective-induced field can be evaluated directly by differentiating the reward with respect to the sample (Song et al., 2023; Bansal et al., 2023).

Target. For a target density q_t, $\mathcal{J}_t^q[\rho] = -D_{\mathrm{KL}}(\rho\|q_t)$ (cf. Jordan et al., 1998; Liu & Wang, 2016) the objective-induced osmotic field is

$$\boldsymbol{d}_t^q = \boldsymbol{d}_t[q_t] - \boldsymbol{d}_t[\rho_t], \tag{25}$$

with the corresponding tilt $\rho_t^{\epsilon,q} \propto \rho_t^{1-\epsilon} q_t^{\epsilon}$. Thus, target guidance is determined by the relative osmotic structure between the current and target marginals and can be evaluated from their respective scores (Ambrosio et al., 2008).

Conditional. Likewise, for a condition $\boldsymbol{y}$, $\mathcal{J}_t^{\boldsymbol{y}}[\rho] = \mathbb{E}_\rho[\log p_t(\boldsymbol{y} \mid \mathbf{x})]$, e.g., for class-conditioned generation and posterior inference (Dhariwal & Nichol, 2021; Ho & Salimans, 2022); Bayes' rule gives

$$\boldsymbol{d}_t^{\boldsymbol{y}} = \boldsymbol{d}_t[p_t(\cdot \mid \boldsymbol{y})] - \boldsymbol{d}_t[p_t], \tag{26}$$

revealing conditional guidance as the relative osmotic field between conditional and unconditional distributions. In practice, this field can therefore be constructed from the corresponding conditional and unconditional score estimates (Ho & Salimans, 2022; Ma et al., 2024).

4.2 Energy estimation and strength calibration

Once the objective-induced osmotic field has been obtained, the same energy-calibration procedure applies across objectives. Given a batch $\{\mathbf{x}_t^{(i)}\}_{i=1}^B$ from the current sampling trajectory, we can estimate its objective energy (Robert & Casella, 2004) as

$$\widehat{\mathcal{E}}_t^{\mathcal{J}} = \frac{1}{2B} \sum_{i=1}^{B} \left\| \boldsymbol{d}_t^{\mathcal{J}}(\boldsymbol{x}_t^{(i)}) \right\|_2^2. \tag{27}$$

The calibrated guidance strength is then

$$\widehat{\lambda}_t^{\mathcal{J}} = \frac{\gamma_t \kappa_t}{2\widehat{\mathcal{E}}_t^{\mathcal{J}} + \varepsilon}, \qquad \boldsymbol{g}_t = \widehat{\lambda}_t^{\mathcal{J}} \boldsymbol{d}_t^{\mathcal{J}}, \tag{28}$$

where $\varepsilon > 0$ is a small numerical stabilizer. The energy profile can be estimated online from the current sampling batch. When offline calibration is used, the energy profile can instead be estimated from a small calibration set and reused during inference. In this case, osmotic guidance requires no retraining of the pretrained flow and introduces only the computation of the objective-induced field and its scalar energy calibration.

5 Experiments

We evaluate osmotic guidance (OG) as a training-free, sample-time guidance method for pretrained generative flows. Our experiments examine two aspects of osmotic guidance: its effectiveness and intervention efficiency as a general guidance framework, and the empirical validation of its main theoretical principles.

5.1 Experimental setup

Datasets and models. We evaluate on 2D synthetic distributions and ImageNet-256. For 2D experiments, we transport a Gaussian base distribution to two-moons, checkerboard, and a four-component Gaussian mixture using a pretrained conditional flow matching model (Lipman et al., 2023; Tong et al., 2024). For images, we use pretrained SiT-XL/2 (Ma et al., 2024) on ImageNet-256 (Deng et al., 2009; Russakovsky et al., 2015). All methods are applied at inference time without additional training.

Objectives and baselines. We study entropy concentration, reward maximization, target-density guidance, and conditional guidance, corresponding to the instances in Table 1. For entropy control, we compare OG with temporal score rescaling (TSR; Xu et al., 2026). For reward and target-density guidance, we additionally compare direct objective-gradient guidance (Raw) and flow guidance g^{MC} (Feng et al., 2025). For conditional generation, we compare with classifier-free guidance (CFG; Ho & Salimans, 2022).

Evaluation. For 2D experiments, we report the W_2 distance (Ambrosio et al., 2008; Benamou & Brenier, 2000), with mean $\pm$ standard deviation over three seeds. For images, we report FID (Heusel et al., 2017; Chong & Forsyth, 2020), sFID (Nash et al., 2021), precision, recall (Kynkäänniemi et al., 2019), and classifier Top-1 accuracy.

Table 2: **Image generative quality at matched terminal contraction for entropy guidance.** Contraction rate approximately matched approx. at 9% and 12%. Accumulated intervention energy is computed by $E_{\text{path}} = \int_0^1 \mathcal{E}_t(\mathbf{g}_t)\, dt$.

Contraction	Method	FID↓	Top-1↑	E_{path} ↓
8.78%	TSR	8.209	**0.727**	11.89
8.79%	OG	**7.876**	0.726	**5.02**
11.72%	TSR	9.584	0.717	23.42
11.75%	OG	**8.186**	**0.723**	**9.49**

Table 3: **2D reward improvement per unit intervention energy.** Reward improvement per unit intervention energy $\Delta R/\mathcal{E}$ ↑ under seed-wise matched-energy guidance. Values are mean ± standard deviation over three seeds.

Method	$\Delta R/\mathcal{E}$ ↑		
	Moons	Checkerboard	Mixture
Raw	17.01±1.16	114.0±1.8	59.99±1.22
g^{MC}	17.16±1.31	94.7±1.5	46.98±1.66
OG	**17.18±1.16**	**123.1±2.5**	**60.78±1.32**

Table 4: **Image generative quality at matched Top-1 accuracy for conditional guidance.** Conditional ImageNet-256 generation at approximately matched classifier Top-1 accuracy: CFG 0.951, OG 0.948.

Method	Matching Top-1			
	FID↓	sFID↓	Precision↑	Recall↑
CFG	12.05	36.05	**0.889**	0.620
OG	**10.01**	**35.41**	0.857	**0.680**

Table 5: **Minimum intervention energy $E^* \times 10^4$ ↓ at matched W_2 reduction δ on the checkerboard benchmark for target-density guidance.** Values are mean ± standard deviation over three seeds.

Method	W_2 reduction δ		
	0.0265	0.0689	0.1114
g^{MC}	7.13±1.41	54.19±8.55	122.1±10.8
OG	**5.46±1.55**	**40.18±7.76**	**112.0±20.7**

5.2 Guidance effectiveness and intervention efficiency

Entropy-based concentration guidance. We first compare sample quality under matched terminal contraction. Table 2 reports FID, Top-1 accuracy, and accumulated intervention energy (E_{path}) for OG and TSR on ImageNet-256 at approximately 9% and 12% contraction. OG achieves lower FID and substantially lower E_{path} at both operating points, while maintaining comparable Top-1 accuracy. These results demonstrate that OG achieves the desired concentration effect with substantially lower intervention cost. They also support our central motivation that, for the same control effect, reducing unnecessary intervention can better preserve generative quality.

Reward guidance. We next evaluate guidance toward an external reward under matched intervention energy. Table 3 reports the reward improvement per unit intervention energy $\Delta R/\mathcal{E}$. OG performs comparably to the competing methods on two-moons and achieves the highest reward efficiency on checkerboard and the Gaussian mixture. These results demonstrate that osmotic guidance also extends effectively to reward-based objectives, achieving greater objective improvement under the same intervention budget.

Conditional guidance. We further evaluate osmotic guidance for conditional image generation. Table 4 compares OG with CFG on ImageNet-256 at approximately matched classifier Top-1 accuracy. At a similar conditioning level, OG achieves better FID, sFID, and recall, while exhibiting slightly worse precision than CFG. These results indicate that osmotic guidance can also effectively support conditional generation while maintaining competitive generative quality.

Target-density guidance. We finally study the intervention efficiency for target-density guidance. Table 5 reports the minimum energy required to achieve a given reduction in W_2 on checkerboard benchmark. Across low-to-moderate target shifts ($\delta < 0.1114$), OG consistently requires less mean intervention energy than g^{MC}. Those results further show that the minimum-energy advantage of OG extends to target-density control.

Qualitative results. Figure 1 provides qualitative samples for entropy and conditional guidance. For entropy guidance (top), increasing γ progressively reduces visual richness and fine-grained texture, illustrating controllable concentration through the guidance strength.

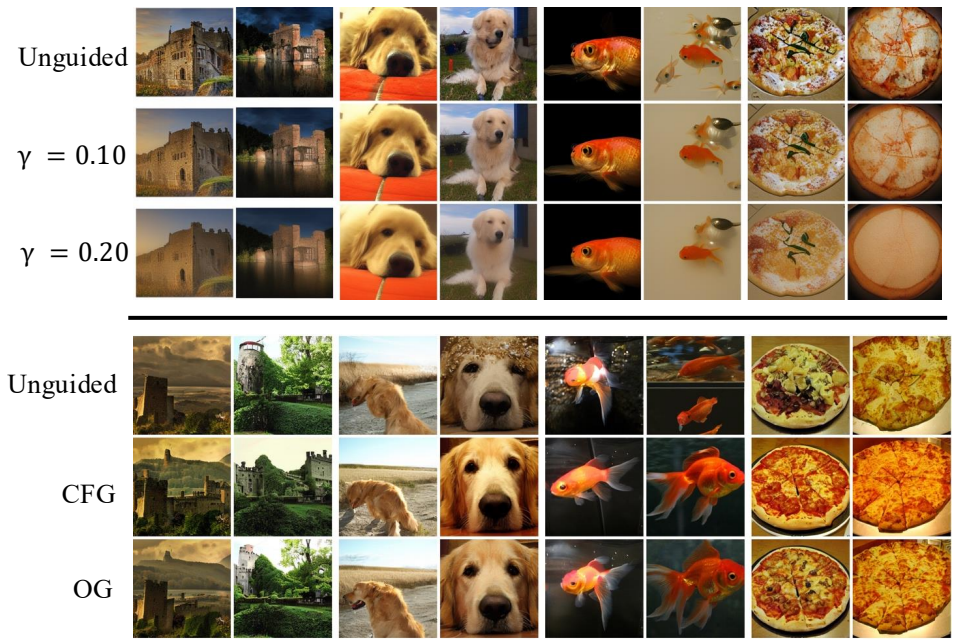

Figure 1: **Qualitative samples for entropy guidance (top three rows) and conditional guidance (bottom three rows).** For entropy guidance, we show unguided samples and OG with $\gamma = 0.10$ and $\gamma = 0.20$. For conditional guidance, we compare unguided samples, CFG, and OG at approximately matched classifier Top-1 accuracy (0.951 for CFG and 0.948 for OG).

For conditional guidance (bottom), we compare CFG and OG at approximately matched conditioning strength (Top-1: CFG 0.951, OG 0.948). Compared with the original unguided samples, CFG exhibits noticeable loss or alteration of details, including background structure, image color, object shape, and local texture. Since the two methods achieve approximately the same classifier accuracy, these additional changes are not necessary for attaining the matched level of conditional adherence. In contrast, OG better preserves these characteristics in several examples, consistent with its canonical guidance direction that suppresses intervention not contributing to the prescribed objective at first order. This observation further supports the motivation for low-energy intervention: reducing guidance that does not directly contribute to the prescribed objective can help preserve other characteristics of the pretrained generative distribution.

5.3 Principle validation

Canonical direction. We first validate the minimum-energy characterization of OG in Proposition 2 by testing whether the theoretically defined osmotic-guidance field $\boldsymbol{g}_t^\star$ maximizes instantaneous objective improvement under a fixed intervention-energy budget. Across 81 configurations, we construct 648 alternative fields $\boldsymbol{g}_t$ with matched intervention energy but different directions. $\mathcal{R}_t(\boldsymbol{g}_t)$ denotes the instantaneous objective-improvement rate induced by $\boldsymbol{g}_t$. Their normalized objective rates follow the predicted cosine-alignment relation $\mathcal{R}_t(\boldsymbol{g}_t)/\mathcal{R}_t(\boldsymbol{g}_t^\star) = \cos(\boldsymbol{g}_t, \boldsymbol{g}_t^\star)$ to numerical precision ($R^2 = 1.000000$, RMSE 4.59×10^{-8}), while the energy mismatch remains negligible (maximum relative deviation 2.32×10^{-7}). The objective improvement is therefore maximized when $\boldsymbol{g}_t$ aligns with $\boldsymbol{g}_t^\star$, supporting the canonical characterization of OG under a fixed intervention-energy budget.

Two-anchor preview. We next evaluate the two-anchor preview introduced in Section 3.3 for efficient guidance-strength selection. As shown in Table 6, replacing a five-candidate exhaustive sweep on ImageNet-256 with the two-anchor approximation yields a $1.642\times$ end-to-end screening speedup with only 0.343 percentage points of mean selection regret (bottom table). The interpolation error also increases smoothly with anchor spacing (top table). Qualitatively, Figure 2 shows that previewed samples closely reproduce the corresponding full generations at the intermediate guidance strength. Together, these results indicate that two anchors provide an accurate and substantially cheaper approximation for selecting guidance strength.

Table 6: **Two-anchor preview accuracy and efficiency.** $\Delta\gamma$ denotes the spacing between the two anchor guidance strengths, while RMSE and relative error measure the absolute and normalized discrepancy respectively.

Interpolation fidelity		
$\Delta\gamma$	RMSE↓	Relative Error↓
0.02	0.00750	0.01004
0.04	0.01662	0.02235
0.06	0.02795	0.03808
0.08	0.03876	0.05311
0.10	0.05085	0.06992
Strength-screening efficiency		
Setting	Speedup↑	Selection Regret↓
SiT Preview	$\approx 1.64\times$	0.20–0.34 pp

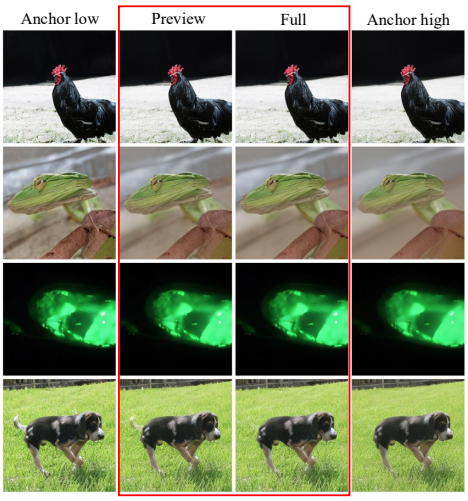

Figure 2: **Qualitative preview results.**

6 Related Work

Guidance and control for generative modeling. Inference-time guidance has been widely studied for diffusion models, from classifier and classifier-free guidance (Dhariwal & Nichol, 2021; Ho & Salimans, 2022) to training-free methods incorporating general conditioning signals or objectives (Bansal et al., 2023; Yu et al., 2023; Song et al., 2023; Chung et al., 2023). Recent methods further improve guidance geometry and stability (Chung et al., 2025; Sadat et al., 2025a), while related analyses study the behavior and limitations of classifier(-free) guidance (Zhao & Schwing, 2026; Chidambaram et al., 2024; Bradley & Nakkiran, 2024). Guidance has also been extended to flow models through multimodal, optimal-control, and general flow-matching formulations (Lin et al., 2025; Wang et al., 2025; Feng et al., 2025). More broadly, generative modeling and guidance have been connected to stochastic and optimal control (Berner et al., 2024; Pandey et al., 2025; Azangulov et al., 2026; Zhang et al., 2026; Wang et al., 2026). Unlike these approaches, osmotic guidance formulates inference-time control through a local minimum-energy principle subject to a prescribed instantaneous objective-change rate.

Energy and score-derived interventions. The score appears as the osmotic velocity in stochastic mechanics and entropic probability transport (Nelson, 1967; Gentil et al., 2020; Léonard, 2014), and modern guidance methods frequently modify score- or prediction-derived components of the sampling dynamics (Karras et al., 2024; Chung et al., 2025; Sadat et al., 2025a;b; Li et al., 2026b). Most closely related, Temporal Score Rescaling directly rescales the marginal score to control sampling concentration in diffusion and flow models (Xu et al., 2026). Energy-based guidance also uses externally defined energies as conditioning objectives (Yu et al., 2023). Osmotic guidance differs in that its guidance energy is a quadratic intervention cost rather than conditioning objective, and is used to derive the minimum-energy relative osmotic field realizing local distributional change.

7 Conclusion

We have introduced osmotic guidance (OG), a training-free framework for flow-based generative models, derived from a canonical minimum-energy distributional-control principle. By connecting the optimal guidance direction to relative osmotic fields, OG provides a unified formulation for multiple guidance objectives with a shared energy-calibration rule. Experiments on synthetic benchmarks and ImageNet-256 demonstrate effective guidance, improved intervention efficiency, and empirical support for the proposed minimum-energy characterization.

Limitations and future work. Theoretically, our minimum-energy characterization is local and instantaneous, rather than a globally optimal path-space control result. Extending osmotic guidance to trajectory-level objectives is an important direction for future work. Empirically, our evaluation remains limited in model coverage, data scale, and guidance objectives. Future work will evaluate broader flow architectures, higher-resolution and more diverse datasets, and richer objectives.

AI USE STATEMENT

In this work, we have not used generative AI tools to help develop theoretical models or conceptual frameworks, formulate mathematical claims, provide critical ingredients for proving mathematical claims, assist in the writing of proofs, propose or refine hypotheses, design or provide feedback on research methodology or experiments, implement methods, clean and reformat datasets, support qualitative and thematic data analysis, interpret results, generate synthetic datasets, or assist with translation, as these are not applicable to this work. Additionally, we used generative AI tools to create or edit software code, summarize or analyse existing literature, polish the writing (fixing grammar errors and typos), and find related works. We have reviewed all AI-assisted work. LLM-generated code was verified and tested for correctness by at least 2 authors. Literature summarized or analysed by LLM was verified and checked for correctness by at least 2 authors. The AI-polished paper was checked by at least 2 authors. Related works found by AI were all verified and checked by at least 2 authors. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

None applicable.

REPRODUCIBILITY STATEMENT

To support reproducibility of this work, we provide the following materials. For all novel formulations, algorithms, and theoretical contributions, including the canonical minimum-energy guidance principle, the objective-induced osmotic representation, and the two-anchor preview bound, complete assumptions and detailed proofs are given in Appendix B, with consistent notation summarized in Appendix A. The practical instantiations and energy-calibration procedure are described in Section 4 and Appendix C. For every dataset used in the experiments, we give a complete description of the data sources, generative backbones, sampling protocols, metrics, and evaluation settings in Section 5.1 and Appendix C, with additional tabulated results in Appendix D. Computational resources and representative wall-clock times are reported in Appendix C.7. Anonymous source code for reproducing the reported experiments is included in the supplementary material and will be made publicly available upon acceptance.

REFERENCES

Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic Interpolants: A Unifying Framework for Flows and Diffusions. *Journal of Machine Learning Research*, 26(209):1–80, 2025. URL https://jmlr.org/papers/v26/23-1605.html. arXiv:2303.08797.

Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics ETH Zürich. Birkhäuser, 2 edition, 2008. doi: 10.1007/978-3-7643-8722-8. URL https://link.springer.com/book/10.1007/978-3-7643-8722-8. DOI: 10.1007/978-3-7643-8722-8.

Aaron D. Ames, Xiangru Xu, Jessy W. Grizzle, and Paulo Tabuada. Control Barrier Function Based Quadratic Programs for Safety Critical Systems. *IEEE Transactions on Automatic Control*, 62 (8):3861–3876, 2017. doi: 10.1109/TAC.2016.2638961. URL https://arxiv.org/abs/1609.06408. arXiv:1609.06408; DOI: 10.1109/TAC.2016.2638961.

Iskander Azangulov, Peter Potaptchik, Qinyu Li, Eddie Aamari, George Deligiannidis, and Judith Rousseau. Adaptive Diffusion Guidance via Stochastic Optimal Control. In *Proceedings of the 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. arXiv:2505.19367.

Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and Geometry of Markov Diffusion Operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2014. doi: 10.1007/978-3-319-00227-9. URL https://link.springer.com/book/10.1007/978-3-319-00227-9. DOI: 10.1007/978-3-319-00227-9.

Arpit Bansal, Hong-Min Chu, Avi Schwarzschild, Soumyadip Sengupta, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Universal Guidance for Diffusion Models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 843–852, 2023. URL https://openaccess.thecvf.com/content/CVPR2023W/GCV/html/Bansal_Universal_Guidance_for_Diffusion_Models_CVPRW_2023_paper.html. arXiv:2302.07121.

Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84:375–393, 2000. doi: 10.1007/s002110050002. URL https://link.springer.com/article/10.1007/s002110050002. DOI: 10.1007/s002110050002.

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/pdf?id=oYIjw37pTP. arXiv:2211.01364.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. URL https://www.seas.ucla.edu/~vandenbe/cvxbook.html.

Arwen Bradley and Preetum Nakkiran. Classifier-Free Guidance Is a Predictor-Corrector. In *Mathematics of Modern Machine Learning (M3L) Workshop at NeurIPS*, 2024. URL https://machinelearning.apple.com/research/predictor-corrector. arXiv:2408.09000.

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. arXiv:2409.13074.

Min Jin Chong and David Forsyth. Effectively Unbiased FID and Inception Score and Where to Find Them. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6070–6079, 2020. URL https://openaccess.thecvf.com/content_CVPR_2020/html/Chong_Effectively_Unbiased_FID_and_Inception_Score_and_Where_to_Find_CVPR_2020_paper.html.

Hyungjin Chung, Jeongsol Kim, Michael T. McCann, Marc L. Klasky, and Jong Chul Ye. Diffusion Posterior Sampling for General Noisy Inverse Problems. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/pdf?id=OnD9zGAGT0k. arXiv:2209.14687.

Hyungjin Chung, Jeongsol Kim, Geon Yeong Park, Hyelin Nam, and Jong Chul Ye. CFG++: Manifold-constrained Classifier Free Guidance for Diffusion Models. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/4c9477b9e2c7ec0ad3f4f15077aaf85a-Abstract-Conference.html. arXiv:2406.08070.

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Fei-Fei Li. ImageNet: A large-scale hierarchical image database. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255. IEEE, 2009. doi: 10.1109/CVPR.2009.5206848. URL https://doi.org/10.1109/CVPR.2009.5206848. DOI: 10.1109/CVPR.2009.5206848.

Prafulla Dhariwal and Alexander Nichol. Diffusion Models Beat GANs on Image Synthesis. In *Advances in Neural Information Processing Systems*, volume 34, pp. 8780–8794, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/49ad23d1ec9fa4bd8d77d02681df5cfa-Abstract.html. arXiv:2105.05233.

Ruiqi Feng, Chenglei Yu, Wenhao Deng, Peiyan Hu, and Tailin Wu. On the Guidance of Flow Matching. In *Proceedings of the International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 16993–17029. PMLR, 2025. URL https://proceedings.mlr.press/v267/feng25s.html. arXiv:2502.02150.

Ivan Gentil, Christian Léonard, Luigia Ripani, and Luca Tamanini. An entropic interpolation proof of the HWI inequality. *Stochastic Processes and their Applications*, 130(2):907–923, 2020. doi: 10.1016/j.spa.2019.04.002. URL https://www.sciencedirect.com/science/article/pii/S0304414918303454. arXiv:1807.06893; DOI: 10.1016/j.spa.2019.04.002.

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*, volume 30, 2017. URL https://arxiv.org/abs/1706.08500. arXiv:1706.08500.

Jonathan Ho and Tim Salimans. Classifier-Free Diffusion Guidance. arXiv preprint, 2022. URL https://arxiv.org/abs/2207.12598. arXiv:2207.12598.

Aapo Hyvärinen. Estimation of Non-Normalized Statistical Models by Score Matching. *Journal of Machine Learning Research*, 6:695–709, 2005. URL https://jmlr.org/papers/volume6/hyvarinen05a/hyvarinen05a.pdf.

Richard Jordan, David Kinderlehrer, and Felix Otto. The Variational Formulation of the Fokker–Planck Equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998. doi: 10.1137/S0036141096303359. URL https://epubs.siam.org/doi/10.1137/S0036141096303359. DOI: 10.1137/S0036141096303359.

Tero Karras, Miika Aittala, Tuomas Kynkäänniemi, Jaakko Lehtinen, Timo Aila, and Samuli Laine. Guiding a Diffusion Model with a Bad Version of Itself. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://arxiv.org/abs/2406.02507. arXiv:2406.02507.

Tuomas Kynkäänniemi, Tero Karras, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Improved Precision and Recall Metric for Assessing Generative Models. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://arxiv.org/abs/1904.06991. arXiv: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*, volume 37, 2024. URL https://papers.nips.cc/paper_files/paper/2024/hash/dd540e1c8d26687d56d296e64d35949f-Abstract-Conference.html. arXiv:2404.07724.

Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete and Continuous Dynamical Systems*, 34(4):1533–1574, 2014. doi: 10.3934/dcds.2014.34.1533. URL https://arxiv.org/abs/1308.0215. arXiv:1308.0215; DOI: 10.3934/dcds.2014.34.1533.

Sergey Levine. Reinforcement Learning and Control as Probabilistic Inference: Tutorial and Review. arXiv preprint, 2018. URL https://arxiv.org/abs/1805.00909. arXiv:1805.00909.

Peizhuo Li, Emre Aksan, Alexandru-Eugen Ichim, Thabo Beeler, and Olga Sorkine-Hornung. Diversify Diffusion with Temperature Sampling and Variance-Corrective Time Shifting. arXiv preprint, 2026a. URL https://arxiv.org/abs/2607.10853. arXiv:2607.10853.

Tiancheng Li, Weijian Luo, Zhiyang Chen, Liyuan Ma, and Guo-Jun Qi. Self-Guidance: Boosting Flow and Diffusion Generation on Their Own. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 48(1):781–791, 2026b. doi: 10.1109/TPAMI.2025.3611831. URL https://arxiv.org/pdf/2412.05827. arXiv:2412.05827; DOI: 10.1109/TPAMI.2025.3611831.

Haowei Lin, Shanda Li, Haotian Ye, Yiming Yang, Stefano Ermon, Yitao Liang, and Jianzhu Ma. TFG-Flow: Training-free Guidance in Multimodal Generative Flow. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/pdf?id=GK5ni7tIHp. arXiv:2501.14216.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow Matching for Generative Modeling. In *International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2210.02747. arXiv:2210.02747.

Qiang Liu and Dilin Wang. Stein Variational Gradient Descent: A General Purpose Bayesian Inference Algorithm. In *Advances in Neural Information Processing Systems*, volume 29, 2016. URL https://arxiv.org/abs/1608.04471. arXiv:1608.04471.

Nanye Ma, Mark Goldstein, Michael S. Albergo, Nicholas M. Boffi, Eric Vanden-Eijnden, and Saining Xie. SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers. In *European Conference on Computer Vision*, 2024. URL https://www.ecva.net/papers/eccv_2024/papers_ECCV/html/9828_ECCV_2024_paper.php. arXiv:2401.08740.

Charlie Nash, Jacob Menick, Sander Dieleman, and Peter Battaglia. Generating images with sparse representations. In *Proceedings of the International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 7958–7968. PMLR, 2021. URL https://proceedings.mlr.press/v139/nash21a.html. arXiv:2103.03841.

Edward Nelson. *Dynamical Theories of Brownian Motion*. Mathematical Notes. Princeton University Press, 1967. doi: 10.2307/j.ctv15r57jg. URL https://web.math.princeton.edu/~nelson/books/bmotion.pdf. Digital edition DOI: 10.2307/j.ctv15r57jg.

Kushagra Pandey, Farrin Marouf Sofian, Felix Draxler, Theofanis Karaletsos, and Stephan Mandt. Variational Control for Guidance in Diffusion Models. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 47755–47780. PMLR, 2025. URL https://proceedings.mlr.press/v267/pandey25a.html. arXiv:2502.03686.

Christian P. Robert and George Casella. *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer, 2 edition, 2004. doi: 10.1007/978-1-4757-4145-2. URL https://link.springer.com/book/10.1007/978-1-4757-4145-2. DOI: 10.1007/978-1-4757-4145-2.

Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Fei-Fei Li. 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. arXiv:1409.0575.

Seyedmorteza Sadat, Otmar Hilliges, and Romann M. Weber. Eliminating Oversaturation and Artifacts of High Guidance Scales in Diffusion Models. In *International Conference on Learning Representations*, 2025a. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/d1450d6c10c6b6cf1b80964357f5fa08-Abstract-Conference.html. arXiv:2410.02416.

Seyedmorteza Sadat, Manuel Kansy, Otmar Hilliges, and Romann M. Weber. No Training, No Problem: Rethinking Classifier-Free Guidance for Diffusion Models. In *International Conference on Learning Representations*, 2025b. URL https://arxiv.org/abs/2407.02687. arXiv:2407.02687.

Jiaming Song, Qinsheng Zhang, Hongxu Yin, Morteza Mardani, Ming-Yu Liu, Jan Kautz, Yongxin Chen, and Arash Vahdat. Loss-Guided Diffusion Models for Plug-and-Play Controllable Generation. In *Proceedings of the International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 32483–32498. PMLR, 2023. URL https://proceedings.mlr.press/v202/song23k.html.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-Based Generative Modeling through Stochastic Differential Equations. In *International Conference on Learning Representations*, 2021. URL https://arxiv.org/abs/2011.13456. arXiv:2011.13456.

Alexander Tong, Kilian Fatras, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport. *Transactions on Machine Learning Research*, 2024. URL https://www.research.ed.ac.uk/files/472344876/TongEtalTMLR2024ImprovingAndGeneralizingFlow-based.pdf. arXiv:2302.00482.

Lloyd N. Trefethen. *Approximation Theory and Approximation Practice, Extended Edition*. Society for Industrial and Applied Mathematics, extended edition, 2019. doi: 10.1137/1.9781611975949. URL https://epubs.siam.org/doi/10.1137/1.9781611975949.fm. DOI: 10.1137/1.9781611975949.

Hanyang Wang, Yiyang Liu, Jiawei Chi, Fangfu Liu, Ran Xue, and Yueqi Duan. CFG-Ctrl: Control-Based Classifier-Free Diffusion Guidance. arXiv preprint, 2026. URL https://arxiv.org/abs/2603.03281. arXiv:2603.03281.

Luran Wang, Chaoran Cheng, Yizhen Liao, Yanru Qu, and Ge Liu. Training Free Guided Flow-Matching with Optimal Control. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/pdf?id=61ss5RA1MM. arXiv:2410.18070.

Yanbo Xu, Yu Wu, Sungjae Park, Zhizhuo Zhou, and Shubham Tulsiani. Temporal Score Rescaling for Temperature Sampling in Diffusion and Flow Models. In *Proceedings of the International Conference on Machine Learning*, 2026. URL https://arxiv.org/abs/2510.01184. arXiv:2510.01184.

Jiwen Yu, Yinhuai Wang, Chen Zhao, Bernard Ghanem, and Jian Zhang. FreeDoM: Training-Free Energy-Guided Conditional Diffusion Model. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 23174–23184, 2023. URL https://openaccess.thecvf.com/content/ICCV2023/html/Yu_FreeDoM_Training-Free_Energy-Guided_Conditional_Diffusion_Model_ICCV_2023_paper.html. arXiv:2303.09833.

Jingdong Zhang, Xinze Li, Yize Jiang, Luan Yang, Minkai Xu, and Junhong Liu. Lyapunov Guidance: A Unified Framework for Stabilizing Generative Flows. arXiv preprint, 2026. URL https://arxiv.org/html/2607.14272v1. arXiv:2607.14272.

Xiaoming Zhao and Alex Schwing. Studying Classifier(-Free) Guidance from a Classifier-Centric Perspective. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(16):13271–13279, 2026. doi: 10.1609/aaai.v40i16.38329. URL https://ojs.aaai.org/index.php/AAAI/article/view/38329. DOI: 10.1609/aaai.v40i16.38329.

Appendix

Contents

- A Notation
  - A.1 Operators and conventions
  - A.2 Symbols
  - A.3 Key expressions
- B Derivations and proofs
  - B.1 Regularity and boundary conditions
  - B.2 Continuity, Fokker–Planck, and marginal equivalence
  - B.3 First variation and guidance-induced rate
  - B.4 Canonical minimum-energy solution and uniqueness
  - B.5 Objective tilt and osmotic representation
  - B.6 Derivation of objective instances
  - B.7 Model proxies and calibration error
  - B.8 Score extraction for independent affine Gaussian paths
  - B.9 Two-anchor interpolation bound
- C Implementation and experimental details
  - C.1 Implemented fields and calibrated coefficients
  - C.2 Energy units, solvers, and logging
  - C.3 2D protocols, baselines, and metrics
  - C.4 Image sampling and conditional guidance
  - C.5 Alignment validation and finite differences
  - C.6 Two-anchor preview, selection, and costs
  - C.7 Computational resources and wall-clock time
- D Additional results and limitations
  - D.1 Matched contraction in 2D
  - D.2 Reward and target-observable efficiency
  - D.3 Target frontiers: overlap and uncertainty
  - D.4 Image energy screens and corrected confirmation
  - D.5 Conditional quality and energy logs
  - D.6 Alignment and finite-difference diagnostics
  - D.7 Preview regret, timing, and anchor-width diagnostic
  - D.8 Historical 50k entropy quality
  - D.9 Historical sensitivity and within-class diagnostics
  - D.10 Unfinished matched-contraction ImageNet extension
  - D.11 Full current target sweeps

A NOTATION

Unless stated otherwise, Euclidean norms sum over coordinates and expectations are taken under the indicated density. The tables below are designed to break across pages.

A.1 OPERATORS AND CONVENTIONS

Table 7: Common operators and conventions used in the main text.

Symbol	Meaning
$\|\cdot\|_2$	Euclidean ℓ_2 norm of a vector.
$\|\cdot\|_2^2$	Squared Euclidean norm.
$\langle\cdot,\cdot\rangle$	Euclidean inner product (when written without a density).
$\mathbb{E}[\cdot]$	Expectation.
$\mathbb{E}_{\rho_t}[\cdot]$	Expectation under the guided marginal ρ_t.
$\mathbb{E}_{\mathbf{x}_t\sim\rho_t}[\cdot]$	Same expectation with the sampling variable written explicitly.
$\nabla_{\boldsymbol{x}}$	Spatial gradient with respect to $\boldsymbol{x}$.
$\nabla_{\boldsymbol{x}}\cdot$	Spatial divergence.
$\Delta_{\boldsymbol{x}}$	Laplacian $\nabla_{\boldsymbol{x}}\cdot\nabla_{\boldsymbol{x}}$.
∂_t	Partial derivative with respect to time.
$\dfrac{\mathrm{d}}{\mathrm{d}t}$	Total / material time derivative along a path.
$\dfrac{\mathrm{d}}{\mathrm{d}t}(\cdot)\Big\|_{\boldsymbol{g}_t}$	Instantaneous contribution attributable to the guidance field $\boldsymbol{g}_t$.
$\propto$	Equality up to a positive normalizing constant.
$\sim$	Distributed according to.
$L^2(\rho_t)$	Space of square-integrable vector fields under ρ_t.
$\mathbb{R}^D$	Ambient Euclidean space of dimension D.

A.2 SYMBOLS

Table 8: Main-paper symbols.

Symbol	Meaning
$t \in [0,1]$	Continuous time index along the generative path.
D	Ambient dimension.
B	Batch size used for empirical energy estimation.
π_t	Reference marginal of the pretrained model at time t.
π_0	Tractable base distribution at $t = 0$.
π_1	Target data-side distribution at $t = 1$ for the reference path.
ρ_t	Guided sampling marginal induced by adding $\boldsymbol{g}_t$.
ρ	Generic probability density argument of a functional.
$\boldsymbol{x}$	State variable in ambient space.
$\mathbf{x}_t$	Random state of the reference/guided ODE at time t.
$\mathbf{y}_t$	Random state of the reference SDE at time t.
$\mathbf{w}_t$	Standard Wiener process.
$\boldsymbol{v}_t$	Pretrained (probability-flow) ODE velocity field.
$\boldsymbol{b}_t$	Stochastic drift of the reference SDE.
$\boldsymbol{g}_t$	Additive guidance / intervention field.
$\boldsymbol{g}_t^\star$	Canonical minimum-energy guidance field.
$\kappa_t \geq 0$	Spatially constant time-dependent diffusivity.
$\nabla_{\boldsymbol{x}} \log \pi_t$	Marginal score of the reference density.
p	Generic smooth positive density.
q	Second generic smooth positive density (relative osmotic identities).
$\boldsymbol{d}_t[p]$	Osmotic field of a positive density p (Eq. 5).
$\mathcal{J}_t[\rho]$	Distributional guidance objective.
$\frac{\delta \mathcal{J}_t}{\delta \rho}$	First variation of the objective.

Continued on next page

Symbol	Meaning
$\nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}$	Objective-gradient field (canonical direction).
γ_t	Prescribed instantaneous objective-change rate.
$\mathcal{E}_t(\boldsymbol{g}_t)$	Instantaneous guidance energy (Eq. 9).
$C(\boldsymbol{g})$	Accumulated guidance energy $\int_0^1 \mathcal{E}_t(\boldsymbol{g}_t)\,\mathrm{d}t$.
ϵ	Exponential-tilt parameter in the objective-tilted density.
$\rho_t^{\epsilon,\mathcal{J}}$	Objective-tilted marginal (Eq. 14).
$\boldsymbol{d}_t^{\mathcal{J}}$	Objective-induced osmotic field (Eq. 16).
$\mathcal{E}_t^{\mathcal{J}}$	Objective energy used for calibration (Eq. 17).
$\lambda_t^{\mathcal{J}}$	Energy-calibrated strength (Eq. 18).
$\mathbf{x}_t^{\mathcal{J}}$	Guided state under osmotic guidance for objective $\mathcal{J}$.
$\rho_t^{\mathcal{J}}$	Guided marginal under osmotic guidance for objective $\mathcal{J}$.
γ	Scalar guidance strength in the family $\gamma_t = \gamma\bar{\gamma}_t$.
$\bar{\gamma}_t$	Fixed time schedule of the guidance-rate family.
γ_a	Lower anchor strength for two-anchor preview.
γ_b	Upper anchor strength for two-anchor preview.
$\boldsymbol{x}_1^{\gamma}$	Terminal state at strength γ.
$\widehat{\boldsymbol{x}}_1^{\gamma}$	Two-anchor linear preview of the terminal state.
ω	Interpolation weight $(\gamma - \gamma_a)/(\gamma_b - \gamma_a)$.
ξ	Dummy strength variable in the preview error bound.
$H(\rho)$	Differential entropy of ρ.
$\mathcal{I}(\rho_t)$	Fisher information of ρ_t.
$\boldsymbol{d}_t^{\mathrm{ent}}$	Entropy / concentration osmotic field.
R_t	Differentiable reward function.
$\boldsymbol{d}_t^{R}$	Reward osmotic field.

Continued on next page

Symbol	Meaning
q_t	Target density for target-density guidance.
$D_{\mathrm{KL}}(\rho\|q_t)$	Kullback–Leibler divergence from ρ to q_t.
$\boldsymbol{d}_t^q$	Target relative osmotic field.
$\boldsymbol{y}$	Conditioning variable.
$p_t(\boldsymbol{y} \mid \boldsymbol{x})$	Conditional likelihood of $\boldsymbol{y}$ given $\boldsymbol{x}$.
p_t	Unconditional marginal used in conditional guidance.
$p_t(\cdot \mid \boldsymbol{y})$	Conditional generative distribution given $\boldsymbol{y}$.
$\boldsymbol{d}_t^{\boldsymbol{y}}$	Conditional relative osmotic field.
$\mathbf{x}_t^{(i)}$	i-th sample in the current batch.
$\widehat{\mathcal{E}}_t^{\mathcal{J}}$	Empirical estimate of the objective energy.
$\widehat{\lambda}_t^{\mathcal{J}}$	Empirically calibrated guidance strength.
ε	Small positive numerical stabilizer in the denominator.

A.3 KEY EXPRESSIONS

Table 9: Key definitions and identities from the main text.

Expression	Role
$\boldsymbol{d}_t[p](\boldsymbol{x}) := \kappa_t \nabla_{\boldsymbol{x}} \log p(\boldsymbol{x})$	Osmotic field (Eq. 5).
$\boldsymbol{d}_t[q] - \boldsymbol{d}_t[p] = \kappa_t \nabla_{\boldsymbol{x}} \log \frac{q}{p}$	Relative osmotic field (Eq. 6).
$\boldsymbol{v}_t = \boldsymbol{b}_t - \kappa_t \nabla_{\boldsymbol{x}} \log \pi_t$	Marginal equivalence (Eq. 4).
$\boldsymbol{b}_t = \boldsymbol{v}_t + \boldsymbol{d}_t[\pi_t]$	Osmotic decomposition of the SDE drift.
$\mathcal{E}_t(\boldsymbol{g}_t) := \frac{1}{2}\mathbb{E}_{\rho_t}\|\boldsymbol{g}_t\|_2^2$	Instantaneous guidance energy (Eq. 9).
$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{J}_t[\rho_t]\Big\|_{\boldsymbol{g}_t} = \mathbb{E}_{\rho_t}\Big[\boldsymbol{g}_t^\top \nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}\Big]$	Guidance-induced objective rate (Eq. 8).
$\boldsymbol{g}_t^\star = \frac{\gamma_t}{\mathbb{E}_{\rho_t}\big\|\nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}\big\|_2^2} \nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}$	Canonical minimum-energy solution (Eq. 12).

Continued on next page

Expression	Role
$\mathcal{E}_t(\boldsymbol{g}_t) \geq \dfrac{\gamma_t^2}{2\mathbb{E}_{\rho_t}\left\|\nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}\right\|_2^2}$	Minimum-energy lower bound (Eq. 13).
$\boldsymbol{d}_t^{\mathcal{J}} := \kappa_t \nabla_{\boldsymbol{x}} \dfrac{\delta \mathcal{J}_t}{\delta \rho} = \dfrac{1}{\epsilon}\big(\boldsymbol{d}_t[\rho_t^{\epsilon,\mathcal{J}}] - \boldsymbol{d}_t[\rho_t]\big)$	Objective-induced osmotic field (Eq. 16).
$\mathcal{E}_t^{\mathcal{J}} := \mathcal{E}_t(\boldsymbol{d}_t^{\mathcal{J}}) = \frac{1}{2}\mathbb{E}_{\rho_t}\|\boldsymbol{d}_t^{\mathcal{J}}\|_2^2$	Objective energy (Eq. 17).
$\boldsymbol{g}_t^\star = \lambda_t^{\mathcal{J}} \boldsymbol{d}_t^{\mathcal{J}}, \lambda_t^{\mathcal{J}} := \dfrac{\gamma_t \kappa_t}{2\mathcal{E}_t^{\mathcal{J}}}$	Energy calibration (Eq. 18).
$\gamma_t = \gamma \bar{\gamma}_t$	Guidance-rate family (Eq. 3.3).
$\widehat{\boldsymbol{x}}_1^\gamma = (1-\omega)\boldsymbol{x}_1^{\gamma_a} + \omega \boldsymbol{x}_1^{\gamma_b}$	Two-anchor preview (Eq. 20).
$\boldsymbol{d}_t^{\mathrm{ent}} = \kappa_t \nabla_{\boldsymbol{x}} \log \rho_t = \boldsymbol{d}_t[\rho_t]$	Entropy instance (Eq. 22).
$\mathcal{I}(\rho_t) := \mathbb{E}_{\rho_t}\|\nabla_{\boldsymbol{x}} \log \rho_t\|_2^2$	Fisher information (Eq. 23).
$\mathcal{E}_t^{\mathrm{ent}} = \dfrac{\kappa_t^2}{2}\mathcal{I}(\rho_t)$	Entropy objective energy (Eq. 23).
$\boldsymbol{d}_t^R = \kappa_t \nabla_{\boldsymbol{x}} R_t$	Reward instance.
$\boldsymbol{d}_t^q = \boldsymbol{d}_t[q_t] - \boldsymbol{d}_t[\rho_t]$	Target instance (Eq. 25).
$\boldsymbol{d}_t^{\boldsymbol{y}} = \boldsymbol{d}_t[p_t(\cdot \mid \boldsymbol{y})] - \boldsymbol{d}_t[p_t]$	Conditional instance (Eq. 26).
$\widehat{\mathcal{E}}_t^{\mathcal{J}} = \dfrac{1}{2B}\sum_{i=1}^{B}\big\|\boldsymbol{d}_t^{\mathcal{J}}(\boldsymbol{x}_t^{(i)})\big\|_2^2$	Empirical objective energy (Eq. 27).
$\widehat{\lambda}_t^{\mathcal{J}} = \dfrac{\gamma_t \kappa_t}{2\widehat{\mathcal{E}}_t^{\mathcal{J}} + \varepsilon},$ $\boldsymbol{g}_t = \widehat{\lambda}_t^{\mathcal{J}} \boldsymbol{d}_t^{\mathcal{J}}$	Empirical strength calibration (Eq. 28).

Local abbreviations used only in proofs. For readability in Appendix B, we occasionally write

$$\boldsymbol{a}_t := \nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho}, \qquad Q_t := \mathbb{E}_{\rho_t}\|\boldsymbol{a}_t\|_2^2, \qquad \mathcal{R}_t(\boldsymbol{g}) := \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{J}_t[\rho_t]\Big|_{\boldsymbol{g}}.$$

These identities are definitional; they do not introduce quantities beyond the main text.

B Derivations and proofs

B.1 Regularity and boundary conditions

We state the regularity assumptions under which the subsequent integrations by parts and instantaneous identities are valid.

Assume sufficiently smooth positive densities on $\mathbb{R}^D$, with integrability and decay that justify differentiating under the integral and discarding boundary terms. Periodic domains are also admissible. On bounded domains, transport and diffusion fluxes must obey compatible boundary conditions so that the discarded boundary integrals vanish.

For the local optimization problem we require $\boldsymbol{a}_t \in L^2(\rho_t; \mathbb{R}^D)$ and $0 < Q_t < \infty$. Field identities hold ρ_t-almost everywhere. These conditions justify instantaneous first-order identities; they do not by themselves guarantee global existence of a density-dependent guided flow. We assume a regular guided solution exists on the interval under consideration. Exponential tilts additionally require finite, strictly positive normalizers.

B.2 Continuity, Fokker–Planck, and marginal equivalence

We derive the continuity and Fokker–Planck equations used in Section 2, and prove the probability-flow marginal equivalence $\boldsymbol{v}_t = \boldsymbol{b}_t - \kappa_t \nabla_{\boldsymbol{x}} \log \pi_t$.

For a smooth compactly supported test function φ, the ODE $\mathrm{d}\mathbf{x}_t/\mathrm{d}t = \boldsymbol{v}_t(\mathbf{x}_t)$ yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \int \varphi \, \pi_t = \int \nabla_{\boldsymbol{x}} \varphi^\top \boldsymbol{v}_t \, \pi_t = - \int \varphi \, \nabla_{\boldsymbol{x}} \cdot (\pi_t \boldsymbol{v}_t). \tag{B.1}$$

This proves Eq. 1.

For the SDE $\mathrm{d}\mathbf{y}_t = \boldsymbol{b}_t(\mathbf{y}_t)\,\mathrm{d}t + \sqrt{2\kappa_t}\,\mathrm{d}\mathbf{w}_t$, Itô's formula gives

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}[\varphi(\mathbf{y}_t)] = \mathbb{E}\big[\boldsymbol{b}_t(\mathbf{y}_t)^\top \nabla_{\boldsymbol{x}} \varphi(\mathbf{y}_t) + \kappa_t \Delta_{\boldsymbol{x}} \varphi(\mathbf{y}_t)\big] \,. \tag{B.2}$$

Integration by parts proves Eq. 2. Since $\pi_t \nabla_{\boldsymbol{x}} \log \pi_t = \nabla_{\boldsymbol{x}} \pi_t$,

$$-\nabla_{\boldsymbol{x}} \cdot (\pi_t \boldsymbol{b}_t) + \kappa_t \Delta_{\boldsymbol{x}} \pi_t = -\nabla_{\boldsymbol{x}} \cdot \big\{ \pi_t \big(\boldsymbol{b}_t - \kappa_t \nabla_{\boldsymbol{x}} \log \pi_t \big) \big\} \,. \tag{B.3}$$

Thus $\boldsymbol{v}_t = \boldsymbol{b}_t - \kappa_t \nabla_{\boldsymbol{x}} \log \pi_t$ is a probability-flow velocity, proving Eqs. 3 and 4. With the same initial law and uniqueness of the density evolution, the ODE and SDE share one-time marginals.

Marginal equivalence does not imply uniqueness of velocity: adding a field $\boldsymbol{h}_t$ with $\nabla_{\boldsymbol{x}} \cdot (\pi_t \boldsymbol{h}_t) = 0$ preserves the continuity equation. The main text uses the standard probability-flow choice. Equivalence concerns one-time marginals, not path laws. The calculation assumes κ_t depends only on time. The decomposition $\boldsymbol{b}_t = \boldsymbol{v}_t + \boldsymbol{d}_t[\pi_t]$ is then a rearrangement of Eq. 4.

For positive densities p and q,

$$\boldsymbol{d}_t[q] - \boldsymbol{d}_t[p] = \kappa_t \big(\nabla_{\boldsymbol{x}} \log q - \nabla_{\boldsymbol{x}} \log p \big) = \kappa_t \nabla_{\boldsymbol{x}} \log \frac{q}{p}, \tag{B.4}$$

which proves Eq. 6.

B.3 First variation and guidance-induced rate

We derive the instantaneous objective-change rate $\mathcal{R}_t(\boldsymbol{g}_t)$ in Eq. 8, and clarify what the rate constraint does and does not control.

At fixed t, the first variation is defined by

$$\frac{\mathrm{d}}{\mathrm{d}\eta} \mathcal{J}_t[\rho_t + \eta h] \bigg|_{\eta=0} = \int \frac{\delta \mathcal{J}_t}{\delta \rho} \, h, \qquad \int h = 0. \tag{B.5}$$

An additive constant in the first variation does not affect $\boldsymbol{a}_t = \nabla_{\boldsymbol{x}} (\delta \mathcal{J}_t / \delta \rho)$.

The guided continuity equation is $\partial_t \rho_t = -\nabla_{\boldsymbol{x}} \cdot [\rho_t (\boldsymbol{v}_t + \boldsymbol{g}_t)]$. Therefore the contribution of $\boldsymbol{g}_t$ alone is

$$\mathcal{R}_t(\boldsymbol{g}_t) = \int \frac{\delta \mathcal{J}_t}{\delta \rho} \big[-\nabla_{\boldsymbol{x}} \cdot (\rho_t \boldsymbol{g}_t) \big] = \int \rho_t \, \boldsymbol{g}_t^\top \nabla_{\boldsymbol{x}} \frac{\delta \mathcal{J}_t}{\delta \rho} = \mathbb{E}_{\rho_t} \big[\boldsymbol{g}_t^\top \boldsymbol{a}_t \big] \,, \tag{B.6}$$

which proves Eq. 8.

The full time derivative also contains the explicit objective-time term and the base transport,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{J}_t[\rho_t] = \partial_t \mathcal{J}_t[\rho]\big|_{\rho=\rho_t} + \mathbb{E}_{\rho_t}[\boldsymbol{v}_t^\top \boldsymbol{a}_t] + \mathcal{R}_t(\boldsymbol{g}_t). \tag{B.7}$$

Setting $\mathcal{R}_t(\boldsymbol{g}_t) = \gamma_t$ therefore constrains only the guidance contribution. Integrating γ_t alone does not generally equal the improvement of a guided run relative to an independently evolving unguided run.

B.4 Canonical minimum-energy solution and uniqueness

Proof of Proposition 1: Couple the guided and unguided trajectories through the same initial sample and define $\Delta_t := \mathbf{x}_t - \mathbf{x}_t^0$. By the Lipschitz assumption on $\boldsymbol{v}_t$,

$$\frac{d}{dt}\|\Delta_t\|_2 \le L_t\|\Delta_t\|_2 + \|\boldsymbol{g}_t(\mathbf{x}_t)\|_2.$$

Since $\Delta_0 = 0$, Gronwall's inequality gives

$$\|\Delta_T\|_2 \le \int_0^T e^{\int_s^T L_r\, dr}\|\boldsymbol{g}_s(\mathbf{x}_s)\|_2\, ds.$$

Using the synchronous coupling, Minkowski's inequality, and $\mathcal{E}_s(\boldsymbol{g}_s) = \frac{1}{2}\mathbb{E}\|\boldsymbol{g}_s(\mathbf{x}_s)\|_2^2$,

$$W_2(\rho_T, p_T) \le \int_0^T e^{\int_s^T L_r\, dr}\sqrt{2\mathcal{E}_s(\boldsymbol{g}_s)}\, ds.$$

Finally, since $L_t \ge 0$, we have $e^{\int_s^T L_r dr} \le e^{\int_0^T L_t dt}$; applying Cauchy–Schwarz yields

$$W_2^2(\rho_T, p_T) \le 2Te^{2\int_0^T L_t\, dt}\int_0^T \mathcal{E}_t(\boldsymbol{g}_t)\, dt.$$

Proof of Proposition 2: among all fields attaining rate γ_t, the unique minimum-energy solution in $L^2(\rho_t)$ is $\boldsymbol{g}_t^\star = (\gamma_t/Q_t)\boldsymbol{a}_t$, and we record the dual fixed-energy characterization used in the principle-validation experiments.

Fix t and write $\langle \boldsymbol{u}, \boldsymbol{w}\rangle_\rho = \mathbb{E}_{\rho_t}[\boldsymbol{u}^\top \boldsymbol{w}]$. For every feasible $\boldsymbol{g}$, Cauchy–Schwarz gives

$$\gamma_t^2 = \langle \boldsymbol{g}, \boldsymbol{a}_t\rangle_\rho^2 \le \|\boldsymbol{g}\|_\rho^2\, Q_t, \qquad \mathcal{E}_t(\boldsymbol{g}) \ge \frac{\gamma_t^2}{2Q_t}. \tag{B.8}$$

The field $\boldsymbol{g}^\star = (\gamma_t/Q_t)\boldsymbol{a}_t$ is feasible and attains the bound, proving Eqs. 12 and 13.

For any feasible $\boldsymbol{g} = \boldsymbol{g}^\star + \boldsymbol{h}$, feasibility forces $\langle \boldsymbol{h}, \boldsymbol{a}_t\rangle_\rho = 0$, and therefore

$$\mathcal{E}_t(\boldsymbol{g}) = \frac{\gamma_t^2}{2Q_t} + \frac{1}{2}\|\boldsymbol{h}\|_\rho^2. \tag{B.9}$$

Equality holds if and only if $\boldsymbol{h} = 0$ in $L^2(\rho_t)$, proving uniqueness. Among feasible fields, proportionality fixes the coefficient to γ_t/Q_t. Negative prescribed rates are covered by the same argument. If $Q_t = 0$, a nonzero rate is infeasible; for zero rate the unique minimum-energy field is zero. Cosines involving zero fields are undefined.

For an energy budget $E_0 > 0$, the dual statement is

$$\max_{\mathcal{E}_t(\boldsymbol{g})\le E_0} \mathcal{R}_t(\boldsymbol{g}) = \sqrt{2E_0 Q_t}, \qquad \boldsymbol{g}_{\mathrm{asc}} = \sqrt{2E_0/Q_t}\,\boldsymbol{a}_t. \tag{B.10}$$

The negative field minimizes the signed rate. For nonzero $\boldsymbol{g}^\star$ and any $\boldsymbol{g}$ with the same energy,

$$\frac{\mathcal{R}_t(\boldsymbol{g})}{\mathcal{R}_t(\boldsymbol{g}^\star)} = \frac{\langle \boldsymbol{g}, \boldsymbol{g}^\star\rangle_\rho}{\|\boldsymbol{g}\|_\rho\|\boldsymbol{g}^\star\|_\rho} = \cos_\rho(\boldsymbol{g}, \boldsymbol{g}^\star). \tag{B.11}$$

This is the alignment identity used in the principle-validation experiments of Section 5.3.

Scope of the orthogonal remainder. The remainder $\boldsymbol{h}$ contributes zero to the chosen objective at first order at the same fixed marginal. It need not be pointwise orthogonal to $\boldsymbol{a}_t$, nor preserve another functional with gradient $\boldsymbol{b}_t$, since $\mathbb{E}_{\rho_t}[\boldsymbol{h}^\top \boldsymbol{b}_t]$ may be nonzero. Eq. B.9 does not by itself prove preservation of image details, nor minimum accumulated energy when competing controls induce different marginals. Matched terminal Top-1 accuracy is likewise not the same as matching the local rate constraint at a common marginal.

B.5 Objective tilt and osmotic representation

We show that the canonical direction equals the relative score of the objective-tilted density, and that energy normalization recovers Eq. 18.

Evaluate the first variation at ρ_t and hold it fixed when forming

$$Z_{t,\epsilon} = \int \rho_t \exp\Big(\epsilon \frac{\delta \mathcal{J}_t}{\delta \rho}\Big), \qquad \rho_t^{\epsilon,\mathcal{J}} = Z_{t,\epsilon}^{-1} \rho_t \exp\Big(\epsilon \frac{\delta \mathcal{J}_t}{\delta \rho}\Big). \tag{B.12}$$

Whenever $0 < Z_{t,\epsilon} < \infty$ and $\epsilon \neq 0$,

$$\nabla_{\boldsymbol{x}} \log \rho_t^{\epsilon,\mathcal{J}} = \nabla_{\boldsymbol{x}} \log \rho_t + \epsilon\, \boldsymbol{a}_t, \tag{B.13}$$

because the normalizer is independent of $\boldsymbol{x}$. This proves Eqs. 15 and 16 exactly, without a small-tilt limit. The tilt is an auxiliary representation; it need not be a marginal attained by finite-time integration. Normalizability does not follow from square-integrability of $\boldsymbol{a}_t$ alone.

Since $\boldsymbol{d}_t^{\mathcal{J}} = \kappa_t \boldsymbol{a}_t$,

$$2\mathcal{E}_t^{\mathcal{J}} = \kappa_t^2 Q_t, \qquad \frac{\gamma_t \kappa_t}{2\mathcal{E}_t^{\mathcal{J}}} \boldsymbol{d}_t^{\mathcal{J}} = \frac{\gamma_t}{Q_t} \boldsymbol{a}_t \quad (\kappa_t > 0). \tag{B.14}$$

This establishes Eqs. 17 and 18; substitution gives Eq. 19. The ideal field is independent of positive κ_t. At $\kappa_t = 0$ the osmotic quotient is indeterminate; if $Q_t > 0$, use the gradient formula $\boldsymbol{g}_t^\star = (\gamma_t / Q_t)\boldsymbol{a}_t$ directly.

B.6 Derivation of objective instances

We derive the four objective-induced osmotic fields in Table 1, and note one closely related objective that is *not* negative KL.

Negative entropy. For $\mathcal{J}_t^{\mathrm{ent}}[\rho] = -H(\rho) = \int \rho \log \rho$,

$$\frac{\mathrm{d}}{\mathrm{d}\eta} \int (\rho + \eta h) \log(\rho + \eta h) \Big|_{\eta=0} = \int (1 + \log \rho)\, h. \tag{B.15}$$

Thus $\boldsymbol{a} = \nabla_{\boldsymbol{x}} \log \rho$ and $\boldsymbol{d}_t^{\mathrm{ent}} = \kappa_t \nabla_{\boldsymbol{x}} \log \rho_t = \boldsymbol{d}_t[\rho_t]$, proving Eqs. 22 and 23. The tilt is proportional to $\rho^{1+\epsilon}$ when integrable. Positive ϵ amplifies high-density regions relative to low-density ones. Exact positive-rate guidance contributes $\mathrm{d}H/\mathrm{d}t|_{\boldsymbol{g}} = -\gamma_t$; this does not by itself imply decreasing total entropy under the base flow, nor improved perceptual quality.

Reward. For $\mathcal{J}_t^R[\rho] = \mathbb{E}_\rho[R_t]$, the first variation is R_t, so $\boldsymbol{a}_t = \nabla_{\boldsymbol{x}} R_t$ and $\boldsymbol{d}_t^R = \kappa_t \nabla_{\boldsymbol{x}} R_t$, with tilt proportional to $\rho_t \exp(\epsilon R_t)$. If $R_t(\boldsymbol{x}) = R(\widehat{\boldsymbol{x}}_1(\boldsymbol{x}, t))$, the chain rule requires the Jacobian of the terminal map; omitting it is an approximation.

Negative KL. For a target q_t held fixed under variations of ρ,

$$\mathcal{J}_t^q[\rho] = -D_{\mathrm{KL}}(\rho \| q_t) = -\int \rho \log(\rho / q_t), \tag{B.16}$$

$$\frac{\delta \mathcal{J}_t^q}{\delta \rho} = \log q_t - \log \rho - 1, \qquad \boldsymbol{a}_t = \nabla_{\boldsymbol{x}} \log q_t - \nabla_{\boldsymbol{x}} \log \rho_t. \tag{B.17}$$

Hence $\boldsymbol{d}_t^q = \boldsymbol{d}_t[q_t] - \boldsymbol{d}_t[\rho_t]$, proving Eq. 25. The tilt is proportional to $\rho_t^{1-\epsilon} q_t^\epsilon$. For positive densities and $0 \leq \epsilon \leq 1$, Hölder's inequality gives a finite normalizer; outside this interval, integrability must be checked separately.

Mean target log-density is a different objective. For $\mathcal{J}_t^{\log q}[\rho] = \mathbb{E}_\rho[\log q_t]$, one obtains $\boldsymbol{a}_t = \nabla_{\boldsymbol{x}} \log q_t$ with no subtraction of the guided score. In particular,

$$\mathcal{R}_t^{-\mathrm{KL}}(\boldsymbol{g}) = \mathbb{E}_{\rho_t}[\boldsymbol{g}^\top \nabla_{\boldsymbol{x}} \log q_t] - \mathbb{E}_{\rho_t}[\boldsymbol{g}^\top \nabla_{\boldsymbol{x}} \log \rho_t]. \tag{B.18}$$

A finite difference of mean $\log q_t$ measures only the first term. Several empirical target tables below report this observable; they should not be read as exact negative-KL derivatives.

Conditional likelihood. Holding the likelihood fixed while varying ρ and assuming $p_t(\boldsymbol{y}) > 0$, Bayes' rule gives, for $\mathcal{J}_t^{\boldsymbol{y}}[\rho] = \mathbb{E}_\rho[\log p_t(\boldsymbol{y} \mid \boldsymbol{x})]$,

$$\boldsymbol{a}_t = \nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{y} \mid \boldsymbol{x}) = \nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x} \mid \boldsymbol{y}) - \nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x}). \tag{B.19}$$

Thus $\boldsymbol{d}_t^{\boldsymbol{y}} = \boldsymbol{d}_t[p_t(\cdot \mid \boldsymbol{y})] - \boldsymbol{d}_t[p_t]$, proving Eq. 26. The tilt is proportional to $\rho_t(\boldsymbol{x})\, p_t(\boldsymbol{y} \mid \boldsymbol{x})^\epsilon$. Unlike the entropy and negative-KL instances, this direction does not require the guided marginal score; its energy is still evaluated under the guided marginal.

B.7 Model proxies and calibration error

We quantify the effect of replacing the true objective gradient by a model proxy (e.g., the pretrained score evaluated on guided samples), and relate the empirical stabilizer to the ideal calibration formula.

Evaluating $\nabla_{\boldsymbol{x}} \log \pi_t$ at a guided sample does not make it equal to $\nabla_{\boldsymbol{x}} \log \rho_t$. Using the pretrained score for entropy therefore controls $\mathbb{E}_{\rho_t}[\log \pi_t]$ with π_t held fixed, rather than the true entropy of ρ_t. Likewise, $\nabla_{\boldsymbol{x}} \log q_t - \nabla_{\boldsymbol{x}} \log \pi_t$ corresponds to $\mathbb{E}_{\rho_t}[\log(q_t/\pi_t)]$, not generally to $-D_{\mathrm{KL}}(\rho_t \| q_t)$.

Let $\widehat{\boldsymbol{a}}_t$ approximate $\boldsymbol{a}_t$, write $\widehat{Q}_t = \mathbb{E}_{\rho_t}\|\widehat{\boldsymbol{a}}_t\|_2^2 > 0$, and set $e_t = \|\boldsymbol{a}_t - \widehat{\boldsymbol{a}}_t\|_{L^2(\rho_t)}$. For the proxy-calibrated field $\boldsymbol{g}_t = \gamma_t \widehat{\boldsymbol{a}}_t / \widehat{Q}_t$, Cauchy–Schwarz yields

$$\left|\mathcal{R}_t^{\mathrm{true}}(\boldsymbol{g}_t) - \gamma_t\right| \le |\gamma_t|\, e_t / \sqrt{\widehat{Q}_t}. \tag{B.20}$$

The error must be controlled under the guided law, not only the training law.

Let $Q_B = B^{-1} \sum_{i=1}^B \|\widehat{\boldsymbol{a}}_t(\boldsymbol{x}_t^{(i)})\|_2^2$ and $\boldsymbol{d}_t^{\mathcal{J}}(\boldsymbol{x}_t^{(i)}) = \kappa_t \widehat{\boldsymbol{a}}_t(\boldsymbol{x}_t^{(i)})$. The empirical calibration in Eqs. 27–28 gives

$$\boldsymbol{g}_t(\boldsymbol{x}_t^{(i)}) = \frac{\gamma_t \kappa_t^2}{\kappa_t^2 Q_B + \varepsilon} \widehat{\boldsymbol{a}}_t(\boldsymbol{x}_t^{(i)}), \qquad \widehat{\mathcal{R}}_B = \gamma_t \frac{\kappa_t^2 Q_B}{\kappa_t^2 Q_B + \varepsilon}. \tag{B.21}$$

The stabilizer biases the achieved rate toward zero. At $\varepsilon = 0$ and positive κ_t, Q_B, the formula is the exact empirical minimum-energy solution. Equivalence with a gradient-space stabilizer ε_a requires $\varepsilon = \kappa_t^2 \varepsilon_a$.

Replacing Q_B by a frozen calibration profile $Q_{\mathrm{cal}}(t)$ yields the evaluation proxy rate

$$\mathcal{R}_t^{\mathrm{proxy}} = \gamma_t \frac{\kappa_t^2 Q_{\mathrm{eval}}(t)}{\kappa_t^2 Q_{\mathrm{cal}}(t) + \varepsilon}. \tag{B.22}$$

Frozen replay is therefore generally approximate. An RMS limit $\|\boldsymbol{g}\|_B \le G$ permits the exact prescribed rate only if $|\gamma_t| \le G\sqrt{Q_B}$. Online normalization couples particles within a batch; independent batches or probes are needed to assess calibration transfer.

B.8 Score extraction for independent affine Gaussian paths

We record the velocity-to-score conversion for independent affine Gaussian paths, which underlies the conditional score-contrast interpretation used in the image experiments.

Let $\mathbf{x}_t = \alpha_t \mathbf{x}_{\mathrm{data}} + \sigma_t \mathbf{z}$ with independent $\mathbf{z} \sim \mathcal{N}(0, I)$ and $\sigma_t > 0$. Write $\boldsymbol{m}_t = \mathbb{E}[\mathbf{x}_{\mathrm{data}} \mid \mathbf{x}_t = \boldsymbol{x}]$ and $\boldsymbol{z}_t = \mathbb{E}[\mathbf{z} \mid \mathbf{x}_t = \boldsymbol{x}]$. Differentiating the Gaussian convolution gives $\boldsymbol{s}_t = \nabla_{\boldsymbol{x}} \log \pi_t = -\boldsymbol{z}_t / \sigma_t$. Since $\boldsymbol{x} = \alpha_t \boldsymbol{m}_t + \sigma_t \boldsymbol{z}_t$ and $\boldsymbol{v}_t = \dot{\alpha}_t \boldsymbol{m}_t + \dot{\sigma}_t \boldsymbol{z}_t$, eliminating $\boldsymbol{m}_t$ yields

$$\boldsymbol{s}_t(\boldsymbol{x}) = \frac{\dot{\alpha}_t\, \boldsymbol{x} - \alpha_t \boldsymbol{v}_t(\boldsymbol{x})}{\sigma_t(\alpha_t \dot{\sigma}_t - \dot{\alpha}_t \sigma_t)} \tag{B.23}$$

whenever the denominator is nonzero. For $\alpha_t = t$ and $\sigma_t = 1-t$, this becomes $\boldsymbol{s}_t = (t\boldsymbol{v}_t - \boldsymbol{x})/(1-t)$ for $t < 1$, and the estimation error scales as $t(\widehat{\boldsymbol{v}}_t - \boldsymbol{v}_t)/(1-t)$. The formula is singular at $t = 1$ and does not apply unchanged to arbitrary coupled transport paths or reversed time conventions. Under the same independent linear path,

$$\boldsymbol{s}_t(\boldsymbol{x} \mid \boldsymbol{y}) - \boldsymbol{s}_t(\boldsymbol{x}) = \frac{t}{1-t}\big(\boldsymbol{v}_t(\boldsymbol{x} \mid \boldsymbol{y}) - \boldsymbol{v}_t(\boldsymbol{x})\big).$$

The sampler convention determines the overall sign and scale.

B.9 Two-anchor interpolation bound

We prove the local interpolation error bound Eq. 21 for the two-anchor preview.

Fix the initial noise, labels, and all other randomness across strengths. For a batch-dependent controller, fix the whole logical batch. Let $F(\gamma) = \boldsymbol{x}_1^\gamma$ (or the full-batch terminal vector), assume $F \in C^2([\gamma_a, \gamma_b]; \mathbb{R}^D)$ with $\gamma_a < \gamma_b$, and let L be the linear interpolant of the two anchors. Integrating F'' twice and matching endpoints gives

$$F(\gamma) - L(\gamma) = -\int_{\gamma_a}^{\gamma_b} K_\gamma(u) F''(u)\,\mathrm{d}u, \tag{B.24}$$

where

$$K_\gamma(u) = \begin{cases} (u-\gamma_a)(\gamma_b-\gamma)/(\gamma_b-\gamma_a), & \gamma_a \le u \le \gamma, \\ (\gamma-\gamma_a)(\gamma_b-u)/(\gamma_b-\gamma_a), & \gamma < u \le \gamma_b. \end{cases} \tag{B.25}$$

Because $K_\gamma \ge 0$ and $\int_{\gamma_a}^{\gamma_b} K_\gamma(u)\,\mathrm{d}u = (\gamma-\gamma_a)(\gamma_b-\gamma)/2$,

$$\|F(\gamma) - L(\gamma)\|_2 \le \frac{(\gamma-\gamma_a)(\gamma_b-\gamma)}{2} \sup_u \|F''(u)\|_2 \le \frac{(\gamma_b-\gamma_a)^2}{8} \sup_u \|F''(u)\|_2. \tag{B.26}$$

This proves Eq. 21 for vector-valued states without requiring a common scalar mean-value point across coordinates.

For finite-batch dynamics $\dot{X} = V(t, X, \gamma)$, sufficient local conditions include twice continuous differentiability in (X, γ), bounded relevant derivatives along solutions, and existence throughout the interval. Clipping transitions, degenerate denominators, or branching solvers can invalidate smoothness. Numerical anchor errors e_a, e_b add at most $(1-\omega)e_a + \omega e_b$ to the bound. This is an interpolation result inside the anchor interval, not an extrapolation guarantee. The selected strength is always evaluated by a full guided solve.

C Implementation and experimental details

C.1 Implemented fields and calibrated coefficients

We specify the concrete objective directions used in the 2D experiments and the empirical floor-based calibration used by the controller.

Write $\boldsymbol{h}_t$ for the implemented objective direction and $\boldsymbol{g}_t$ for the field actually added to the velocity. For the 2D experiments the three directions are

$$\boldsymbol{h}_t^{\mathrm{ent}}(\boldsymbol{x}) = \boldsymbol{d}_\theta(t, \boldsymbol{x})/\kappa, \tag{C.1}$$

$$\boldsymbol{h}_t^{R}(\boldsymbol{x}) = -(\boldsymbol{x} - \boldsymbol{c}), \quad \boldsymbol{c} = (1.5, 0), \quad R(\boldsymbol{x}) = -\tfrac{1}{2}\|\boldsymbol{x} - \boldsymbol{c}\|_2^2, \tag{C.2}$$

$$\boldsymbol{h}_t^{q}(\boldsymbol{x}) = \nabla_{\boldsymbol{x}} \log q_{\mathrm{mix}}(\boldsymbol{x}) - \boldsymbol{d}_\theta(t, \boldsymbol{x})/\kappa. \tag{C.3}$$

The reward and target code directions carry no outer κ factor. The target is an equal-weight four-component isotropic Gaussian mixture shared across the three source datasets. For component centers $\boldsymbol{\mu}_j$ and common standard deviation σ,

$$\nabla_{\boldsymbol{x}} \log q_{\mathrm{mix}}(\boldsymbol{x}) = -\sum_{j=1}^{4} w_j(\boldsymbol{x})\,(\boldsymbol{x} - \boldsymbol{\mu}_j)/\sigma^2, \qquad w_j(\boldsymbol{x}) = \frac{e^{-\|\boldsymbol{x}-\boldsymbol{\mu}_j\|_2^2/(2\sigma^2)}}{\sum_\ell e^{-\|\boldsymbol{x}-\boldsymbol{\mu}_\ell\|_2^2/(2\sigma^2)}}. \tag{C.4}$$

The same target-relative direction is used by the general guidance, target-frontier, and canonical-alignment studies. Their reported scalar target observable is mean $\log q_{\mathrm{mix}}$, whose derivative differs from that of negative KL (Eq. B.18).

The exported 2D controller averages over samples and coordinates and uses a floor rather than an additive stabilizer:

$$m_B(t) = \frac{1}{BD}\sum_{i=1}^{B} \|\boldsymbol{h}_t(\boldsymbol{x}_i)\|_2^2, \qquad \boldsymbol{g}_t(\boldsymbol{x}) = \frac{\gamma_{\mathrm{code}}}{\max\{m_B(t), \varepsilon\}}\boldsymbol{h}_t(\boldsymbol{x}), \quad \varepsilon = 10^{-12}. \tag{C.5}$$

The reported 2D suites replay frozen calibration calendars rather than recomputing the norm at every guided solver stage; in that case replace m_B by the interpolated calibration value $m_{\mathrm{cal}}(t)$. The general suite uses 1024 calibration samples and 31 time bins.

With the empirical floor and an on-batch norm, the achieved per-coordinate proxy rate and paper-convention energy are

$$\widehat{\mathcal{R}}_B/D = \gamma_{\mathrm{code}}\frac{m_B}{\max(m_B, \varepsilon)}, \qquad \widehat{\mathcal{E}}_B = \frac{D\,\gamma_{\mathrm{code}}^2\, m_B}{2\max(m_B, \varepsilon)^2}. \tag{C.6}$$

For frozen calibration the numerator remains the evaluation norm, while the denominator is the calibration norm. The floor leaves the proxy rate unbiased when inactive; the additive stabilizer in Eq. 28 instead produces the bias in Eq. B.21.

Importantly, the learned reference score is evaluated at the current sample, but no guided-density estimator is present. Consequently, the entropy and target experiments test reference-score proxies. The local variational theorem remains exact for the specified density and objective; it does not turn these proxies into the true entropy or negative-KL derivative of the guided law.

C.2 Energy units, solvers, and logging

We document how logged energies convert to the paper convention $\mathcal{E}_t = \frac{1}{2}\mathbb{E}\|\boldsymbol{g}_t\|_2^2$, and which ODE solvers are used.

The 2D logger integrates $\mathbb{E}\|\boldsymbol{g}_t\|_2^2/D$ without a factor of $1/2$, using trapezoidal integration of saved profiles. Therefore

$$C_{\mathrm{paper}} = \frac{D}{2}\,C_{\text{2D logger}}. \tag{C.7}$$

For $D = 2$ the numerical values coincide. With the same direction and an inactive floor, on-batch calibration would require $\gamma_{\mathrm{paper}} = D\gamma_{\mathrm{code}}$.

Image energy logs sum all latent coordinates and omit $1/2$: $C_{\mathrm{paper}} = C_{\text{image logger}}/2$ for the same trajectory and quadrature. The latent dimension is $4 \times 32 \times 32 = 4096$. This energy convention must not be confused with the flat-mean norm used by a reference-Fisher coefficient; the conditional coefficient below uses a summed norm.

The 2D production studies use explicit midpoint with step size 0.02 (50 steps, 100 velocity evaluations). Image studies use 50 Heun steps (100 ODE velocity evaluations). Heun averaging of velocity stages does not determine how a separate diagnostic energy is integrated: the audited SiT running total is left-Riemann at the first stage, whereas the saved profile helper uses a trapezoidal integral. We report both when both are available. A doubled conditional/unconditional network workload is distinct from per-path ODE NFE.

C.3 2D protocols, baselines, and metrics

We summarize sample sizes, seeds, strength grids, and matching protocols for the 2D suites reported in the main text and Appendix D.

The general guidance suite uses 2048 samples, pairwise subsampling to 1500, seeds $0, 1, 2$, and $\kappa = 0.1$. Its OG strengths are $\{.02, .05, .1, .2\}$ and raw strengths are $\{.25, .5, 1, 2\}$.

The matched-contraction suite uses 8192 evaluation samples, a 2048-sample selection screen, pairwise subsampling to 4096, and a 1024-sample exact transport comparison (POT 0.9.7.post1). Samples share initial randomness across methods within each seed. The spread statistic is the sum of

coordinate-wise unbiased sample variances, and contraction is

$$c = 1 - \frac{\mathrm{spread}_{\mathrm{guided}}}{\mathrm{spread}_{\mathrm{unguided}}}. \tag{C.8}$$

All 27 method/dataset/target cell means lie within 0.3 percentage points of the requested 5%, 9%, or 12% contraction (26 within 0.2). The constant baseline has no terminal cutoff in this 2D suite; the TSR data-scale parameter is 1.

The reward/target baseline comparison uses 2048 samples and seeds $0, 1, 2$. The Monte Carlo baseline is the adapted g^{MC} construction with a 512-sample bank and an independent Gaussian VP path. To match OG at $\gamma = 0.1$, it selects the nearest energy on a strength grid and, when the relative gap exceeds 25%, applies one square-root energy-ratio refinement. Matching is approximate; actual energies and paired gains are therefore reported together. Ratios are formed within each seed before taking their mean and sample standard deviation.

The target-frontier study uses 8192 samples, pairwise subset 4096, W_2 subset 1024, and the same seeds and midpoint step. Its strength grids are OG $\{.02, .05, .08, .1, .15, .2, .3, .45, .7\}$, Raw $\{.005, .01, .02, .035, .05, .08, .12, .2\}$, and MC guide scale $\{.01, .02, .05, .1, .15, .25, .4, .6\}$ with outer scale 1. We distinguish a seed-mean envelope from per-seed Pareto interpolation; for the latter, improvements are relative to each seed's unguided W_2, and only the common improvement range is used.

C.4 Image sampling and conditional guidance

We specify the ImageNet-256 conditional protocol and how the executed CFG/OG velocities relate to the ideal relative-osmotic field.

The conditional protocol uses SiT-XL/2, the `stabilityai/sd-vae-ft-ema` decoder, latent scaling 0.18215, null class 1000, generation batch 48, decode batch 32, $t_{\mathrm{eps}} = 10^{-4}$, and normalization stabilizer 10^{-12}. The one-image-per-class screen has 1000 samples; the confirmation has five images per class (5000 samples). The base seed is 20260919, with per-class seed $20260919 + 1000003\,y$.

Let P project onto the first three of four latent channels, and define $\Delta\boldsymbol{v} = \boldsymbol{v}_{\mathrm{cond}} - \boldsymbol{v}_{\mathrm{uncond}}$. The executed conditional velocity is

$$V_\beta = P\big(\boldsymbol{v}_{\mathrm{uncond}} + \beta\Delta\boldsymbol{v}\big) + (I - P)\boldsymbol{v}_{\mathrm{cond}}. \tag{C.9}$$

CFG sets $\beta = w$. The calibrated protocol uses $\beta(t) = \gamma/(m_{\mathrm{cond}}(t) + \varepsilon)$, where m_{cond} pools summed squared norms of $P\Delta\boldsymbol{v}$ along conditional Heun trajectories (256 calibration samples, 24 time bins). It is a projected velocity-contrast calibration. Under the ideal independent linear path of Appendix B.8, the score contrast is $t\Delta\boldsymbol{v}/(1 - t)$; thus the executed direction is not literally the full likelihood score, and even with an exact online norm the corresponding projected score-rate parameter is $\gamma t/(1 - t)$ rather than a constant γ.

Logged conditional energy is relative to $P\boldsymbol{v}_{\mathrm{uncond}} + (I - P)\boldsymbol{v}_{\mathrm{cond}}$, so the intervention is $\beta\, P\Delta\boldsymbol{v}$. CFG with $w = 1$ therefore has nonzero logged energy. Matched terminal Top-1 is not matched instantaneous rate.

C.5 Alignment validation and finite differences

We detail the matched-energy cosine-alignment sweep used to validate Proposition 2.

The validation uses 1024 samples per configuration, seeds $0, 1, 2$, and times $0.2, 0.5, 0.8$. For direction $\boldsymbol{\phi}$ and a normalized orthogonal field $\boldsymbol{z}$, it constructs

$$\boldsymbol{g}_c = A\Big(c\,\boldsymbol{\phi}/\|\boldsymbol{\phi}\|_\rho + \sqrt{1 - c^2}\,\boldsymbol{z}\Big), \qquad A = \|\boldsymbol{\phi}\|_\rho,$$

so $\boldsymbol{g}_1 = \boldsymbol{\phi}$. Norms sum coordinates and energy includes the factor $1/2$ in this diagnostic. There are 648 alignment rows and 1728 finite-difference rows. The finite difference uses $\boldsymbol{x} \mapsto \boldsymbol{x} + \eta\boldsymbol{g}_c$ with $\eta \in \{10^{-4}, 3\cdot10^{-4}, 10^{-3}, 3\cdot10^{-3}\}$. Its normalized rate divides by $\max(|\langle\boldsymbol{g}_1, \boldsymbol{\phi}\rangle|, 10^{-12})$. Reward and mean target log-density are tested; entropy finite differences were not run.

C.6 TWO-ANCHOR PREVIEW, SELECTION, AND COSTS

We describe the ImageNet two-anchor preview protocol, the contraction selection regret, and the cost model behind the reported speedups.

The ImageNet confirmation uses 16 batches of 64 samples (32 classes with two samples per class), class offset 19, and seeds 20261101 through 20261116. All use Heun 50 steps, CFG 1, and a frozen reference-Fisher schedule. Fisher interpolation is linear in γ between 0 and 0.1; Constant interpolation is linear in λ between 0 and 0.35 with a 0.95 terminal window; TSR interpolation is linear in $\log k$ between $k = 1$ and 1.15. The five screened candidates are $\{.02, .04, .06, .08, .10\}$, $\{.10, .15, .20, .25, .35\}$, and $\{1.05, 1.08, 1.10, 1.12, 1.15\}$, respectively. Interpolation mixes terminal latents from common-noise anchors. The selected candidate receives a full solve unless it is already an anchor.

For requested latent contraction $c_\star$, selection regret is

$$\left|c_{\mathrm{full}}(\widehat{\theta}_{\mathrm{preview}}) - c_\star\right| - \min_{\theta \in \mathcal{G}} \left|c_{\mathrm{full}}(\theta) - c_\star\right|. \tag{C.10}$$

We multiply by 100 to report percentage points. Targets are $\{.05, .07, .09, .11\}$; the .03 target was excluded because the calibration grid lacked a TSR candidate within .02.

The deployment cost model includes two anchor solves, five latent mix/decode previews, and the final solve when needed. The naive comparison includes five full candidates plus deployment. A one-time calibration cost of 2120.523 seconds (55.537 for the schedule and 2064.985 for pilots) is excluded from repeated-session speedups and reported separately. Break-even is computed from ODE savings under this reuse assumption; cold-start use is therefore not claimed to be faster.

C.7 COMPUTATIONAL RESOURCES AND WALL-CLOCK TIME

All reported experiments are training-free: they reuse pretrained flow checkpoints and modify only the inference-time velocity. Table 10 lists the GPUs used by the main suites, and Table 11 summarizes representative wall-clock times from the logged runs. Times are single-GPU wall clocks unless noted; they exclude host idle time between jobs.

Table 10: Hardware used for the main experimental suites.

Suite	Hardware / software
2D matched contraction, target frontier, general guidance	NVIDIA GeForce RTX 5090; PyTorch 2.7.0+cu128; POT 0.9.7.post1
2D reward / target baseline with g^{MC}	NVIDIA RTX 6000 Ada Generation
ImageNet conditional guidance (SiT-XL/2)	NVIDIA A800 40GB (directory name contains `h100`; logged device is A800)
ImageNet preview / entropy screens	Single NVIDIA GPU (same SiT inference stack; Heun 50)

Osmotic guidance adds only the objective-induced field and a scalar energy normalization on top of the pretrained velocity solve. In 2D, the OG overhead relative to the unguided midpoint solve is typically a factor of about 1.8–2 at matched sample size. On ImageNet, the dominant cost is the SiT Heun trajectory itself (and, for conditional guidance, the paired conditional/unconditional forwards); the calibration schedule is computed once and reused. No experiment in this paper retrains the generative backbone.

D ADDITIONAL RESULTS AND LIMITATIONS

This section supplements the main-paper figures with complete operating-point tables. Unless specified otherwise, 2D uncertainty is the sample standard deviation over seeds $0, 1, 2$. Image quality metrics computed on a single pooled set have no implied seed uncertainty. Energy columns retain their labeled source conventions (Appendix C.2).

Table 11: Representative computational cost. ODE NFE counts velocity evaluations along one path; conditional methods also evaluate conditional and unconditional networks.

Workload	Solver / NFE	Wall-clock time
2D ODE solve (N=8192, matched contraction)	Midpoint, step 0.02 ($\approx$ 100 NFE)	Unguided 0.07 ± 0.04 s; OG 0.13 ± 0.03 s; Constant 0.12 ± 0.00 s; TSR 0.15 ± 0.00 s
2D target-frontier solve (N=8192)	Midpoint, $\approx$ 100 NFE	Unguided ≈ 0.07 s; OG ≈ 0.14 s; Raw ≈ 0.13 s; $g^{\mathrm{MC}} \approx 0.12$ s (mean over strengths)
2D reward/target baseline (N=2048)	Midpoint, $\approx$ 100 NFE	OG ≈ 0.09 s; Raw ≈ 0.06 s; $g^{\mathrm{MC}} \approx 0.12$ s
ImageNet conditional m_{cond} calibration (256 samples, 24 bins)	Heun 50 (100 NFE)	416 s
ImageNet conditional screen (1000 images)	Heun 50 (100 NFE)	$\approx 3.3\times10^3$ s ($\approx$ 55 min) per strength; CFG w=1 $\approx 2.2\times10^3$ s
ImageNet conditional confirm (5000 images)	Heun 50 (100 NFE)	$\approx$ 1.3–1.6$\times10^4$ s ($\approx$ 3.7–4.6 h) per strength
ImageNet two-anchor preview (repeated session, cost model)	Heun 50 (100 NFE)	One-time calibration 2.121×10^3 s; deployed session $\approx$ 333 s ($1.642\times$ vs naive ≈ 547 s)
Historical ImageNet entropy 5k sensitivity	Heun 50 (100 NFE)	$\approx 8.2\times10^3$ s ($\approx$ 2.3 h) per newly generated strength

D.1 Matched contraction in 2D

Complete selected strengths, achieved contractions, energies, and distributional discrepancies (SW, MMD, W_2) for entropy guidance on moons, checkerboard, and mixture at target contractions of 5%, 9%, and 12%.

The following tables include every selected operating point. Mean and sample SD are over three paired seeds, with 8192 samples per seed. Negative small MMD estimates are retained as recorded. Matching concerns contraction and does not force equal distributional quality. OG uses less energy than TSR in all nine guided comparisons, but does not dominate every quality metric or Constant at every point.

Table 12: moons: selected strengths, achieved contraction, and energy.

Method	Target (%)	Strength	Actual (%)	Energy
Unguided	0	0	0 ± 0	0 ± 0
OG	5	0.07279	5.005 ± 0.37	$0.00381 \pm 7.7 \times 10^{-5}$
Constant	5	0.5781	5.04 ± 0.053	0.008458 ± 0.0012
TSR	5	1.136	5.045 ± 0.13	0.01489 ± 0.00054
OG	9	0.1332	9.004 ± 0.57	0.01212 ± 0.00025
Constant	9	1.119	9.07 ± 0.15	0.02666 ± 0.0034
TSR	9	1.289	9.103 ± 0.2	0.05137 ± 0.0019
OG	12	0.1809	12.02 ± 0.87	0.02149 ± 0.00047
Constant	12	1.571	12.07 ± 0.39	0.04625 ± 0.0052
TSR	12	1.443	12.25 ± 0.22	0.09568 ± 0.0034

D.2 Reward and target-observable efficiency

Seed-wise matched-energy comparisons for reward maximization and mean target log-density, including residual energy mismatch when matching is approximate.

Gain is paired with the unguided seed; gain/energy is averaged after forming individual ratios. The target observable is mean log-density, not negative KL. Approximate energy matching leaves residual differences: for checkerboard reward, MC uses about 20% less energy than OG.

Table 13: moons: distributional discrepancies at the same operating points.

Method	Target (%)	SW	MMD	W_2
Unguided	0	0.02598 ± 0.0064	$3.6 \times 10^{-5} \pm 0.00018$	0.09871 ± 0.017
OG	5	0.03158 ± 0.0052	0.0002128 ± 0.00014	0.09484 ± 0.014
Constant	5	0.03146 ± 0.0042	0.000193 ± 0.00011	0.09468 ± 0.01
TSR	5	0.03226 ± 0.0035	$0.0002163 \pm 8.2 \times 10^{-5}$	0.09572 ± 0.0077
OG	9	0.04655 ± 0.0079	0.0008033 ± 0.00036	0.1143 ± 0.045
Constant	9	0.04749 ± 0.0064	0.0007816 ± 0.00027	0.1154 ± 0.043
TSR	9	0.04754 ± 0.0051	0.0008013 ± 0.00021	0.113 ± 0.035
OG	12	0.05968 ± 0.0098	0.001508 ± 0.0006	0.1317 ± 0.055
Constant	12	0.06171 ± 0.0083	0.00149 ± 0.00048	0.1383 ± 0.054
TSR	12	0.06217 ± 0.0065	0.001548 ± 0.00038	0.1341 ± 0.047

Table 14: checkerboard: selected strengths, achieved contraction, and energy.

Method	Target (%)	Strength	Actual (%)	Energy
Unguided	0	0	0 ± 0	0 ± 0
OG	5	0.07206	5.002 ± 0.35	0.005554 ± 0.00012
Constant	5	0.8129	5.017 ± 0.42	0.005538 ± 0.00029
TSR	5	1.503	5.056 ± 0.43	0.05535 ± 0.0022
OG	9	0.1323	9.021 ± 0.61	0.01727 ± 0.00039
Constant	9	1.491	9.024 ± 0.77	0.01739 ± 0.00089
TSR	9	2.471	9.077 ± 1.1	0.2051 ± 0.008
OG	12	0.1793	12.07 ± 0.83	0.0299 ± 0.0007
Constant	12	2.023	12.09 ± 1.1	0.03044 ± 0.0015
TSR	12	5.268	12.17 ± 2.5	0.506 ± 0.026

Table 15: checkerboard: distributional discrepancies at the same operating points.

Method	Target (%)	SW	MMD	W_2
Unguided	0	0.1265 ± 0.035	0.0002709 ± 0.00033	0.4501 ± 0.049
OG	5	0.1438 ± 0.03	0.000458 ± 0.00032	0.4567 ± 0.027
Constant	5	0.1439 ± 0.028	0.0004548 ± 0.0003	0.4572 ± 0.025
TSR	5	0.1501 ± 0.032	0.0004348 ± 0.00034	0.4694 ± 0.041
OG	9	0.1768 ± 0.032	0.0009103 ± 0.00045	0.4815 ± 0.014
Constant	9	0.1772 ± 0.027	0.0009069 ± 0.00037	0.4825 ± 0.012
TSR	9	0.1991 ± 0.026	0.0008423 ± 0.00038	0.5212 ± 0.04
OG	12	0.21 ± 0.036	0.001466 ± 0.00061	0.5084 ± 0.012
Constant	12	0.2112 ± 0.029	0.001469 ± 0.00046	0.5111 ± 0.0063
TSR	12	0.2662 ± 0.015	0.001395 ± 0.00023	0.6131 ± 0.047

Table 16: mixture: selected strengths, achieved contraction, and energy.

Method	Target (%)	Strength	Actual (%)	Energy
Unguided	0	0	0 ± 0	0 ± 0
OG	5	0.06305	5.018 ± 0.18	0.003383 ± 0.00018
Constant	5	0.5924	5.018 ± 0.062	0.003943 ± 0.00028
TSR	5	1.199	5.107 ± 0.21	0.01886 ± 0.00012
OG	9	0.1166	9.037 ± 0.3	0.01099 ± 0.00058
Constant	9	1.106	9.05 ± 0.13	0.01296 ± 0.00085
TSR	9	1.443	9.178 ± 0.4	0.07024 ± 0.00046
OG	12	0.1587	12.06 ± 0.37	0.01958 ± 0.001
Constant	12	1.514	12.06 ± 0.17	0.0232 ± 0.0014
TSR	12	1.721	12.14 ± 0.55	0.1408 ± 0.00091

Table 17: mixture: distributional discrepancies at the same operating points.

Method	Target (%)	SW	MMD	W_2
Unguided	0	0.09086 ± 0.026	0.0001758 ± 0.00025	0.3277 ± 0.048
OG	5	0.1146 ± 0.019	0.0004608 ± 0.00029	0.3403 ± 0.04
Constant	5	0.1144 ± 0.02	0.0004507 ± 0.00029	0.3409 ± 0.041
TSR	5	0.1195 ± 0.022	0.0004402 ± 0.00028	0.3531 ± 0.042
OG	9	0.1444 ± 0.016	0.0009863 ± 0.0004	0.3627 ± 0.032
Constant	9	0.145 ± 0.017	0.0009636 ± 0.00038	0.3649 ± 0.034
TSR	9	0.1625 ± 0.017	0.0009169 ± 0.00035	0.3994 ± 0.035
OG	12	0.1705 ± 0.015	0.001579 ± 0.00052	0.3839 ± 0.025
Constant	12	0.172 ± 0.015	0.001542 ± 0.00047	0.3879 ± 0.029
TSR	12	0.2038 ± 0.015	0.001448 ± 0.00041	0.4505 ± 0.03

Table 18: Reward: terminal observable, paired gain, and efficiency.

Dataset	Method	Observable	Energy	Gain	Gain/energy
moons	OG	-0.9562 ± 0.031	0.005703 ± 0.00012	0.0979 ± 0.0049	17.18 ± 1.2
moons	Raw	-0.9579 ± 0.031	0.005661 ± 0.00012	0.09622 ± 0.0049	17.01 ± 1.2
moons	g^{MC}	-0.9553 ± 0.031	0.005774 ± 0.00023	0.09888 ± 0.0038	17.16 ± 1.3
checkerboard	OG	-7.216 ± 0.25	0.003216 ± 0.0001	0.396 ± 0.013	123.1 ± 2.5
checkerboard	Raw	-7.249 ± 0.25	0.003178 ± 0.0001	0.3624 ± 0.012	114 ± 1.8
checkerboard	g^{MC}	-7.367 ± 0.24	0.002582 ± 0.00014	0.2446 ± 0.014	94.72 ± 1.5
mixture	OG	-3.627 ± 0.032	$0.003812 \pm 7.2 \times 10^{-5}$	0.2317 ± 0.0077	60.78 ± 1.3
mixture	Raw	-3.627 ± 0.033	$0.003853 \pm 7.2 \times 10^{-5}$	0.2311 ± 0.0075	59.99 ± 1.2
mixture	g^{MC}	-3.679 ± 0.03	0.003824 ± 0.00016	0.1795 ± 0.0064	46.98 ± 1.7

Table 19: Mean target log-density (source task key: kl): terminal observable, paired gain, and efficiency.

Dataset	Method	Observable	Energy	Gain	Gain/energy
moons	OG	-4.185 ± 0.016	$0.001479 \pm 4 \times 10^{-5}$	0.06605 ± 0.0039	44.72 ± 3.7
moons	Raw	-4.185 ± 0.016	$0.001487 \pm 4.2 \times 10^{-5}$	0.06587 ± 0.0043	44.36 ± 4.1
moons	g^{MC}	-4.158 ± 0.016	$0.001685 \pm 1 \times 10^{-5}$	0.09329 ± 0.003	55.37 ± 2.1
checkerboard	OG	-11.87 ± 0.45	$0.001043 \pm 3.9 \times 10^{-5}$	0.2945 ± 0.026	282.4 ± 23
checkerboard	Raw	-11.85 ± 0.45	0.00113 ± 0.00016	0.3179 ± 0.021	284.6 ± 38
checkerboard	g^{MC}	-11.88 ± 0.49	0.001124 ± 0.0001	0.2807 ± 0.026	249.7 ± 6
mixture	OG	-2.77 ± 0.017	0.003733 ± 0.00019	0.06338 ± 0.0092	17.06 ± 3
mixture	Raw	-2.799 ± 0.0095	0.003702 ± 0.00019	0.03431 ± 0.0028	9.309 ± 1.2
mixture	g^{MC}	-2.736 ± 0.0076	0.003861 ± 0.00015	0.09663 ± 0.0023	25.07 ± 1.5

D.3 Target frontiers: overlap and uncertainty

Per-seed and seed-mean frontier analyses for target-density guidance, including cases where OG and MC lack a common reduction interval.

The per-seed analysis uses $\delta = W_{2,\mathrm{base}} - W_2$ and a monotone Pareto envelope before interpolation. Only checkerboard has a common OG–MC reduction interval across all three seeds; moons and mixture do not. For checkerboard, the unguided baseline is 1.863 ± 0.113.

Table 20: Checkerboard per-seed reduction analysis. SDs are across three seeds; ratios are paired seed statistics.

δ	Target W_2	OG energy	MC energy	MC/OG
.0265	$1.836 \pm .113$	$.000546 \pm .000155$	$.000713 \pm .000141$	$1.326 \pm .123$
.0689	$1.794 \pm .113$	$.004018 \pm .000776$	$.005419 \pm .000855$	$1.361 \pm .178$
.1114	$1.751 \pm .113$	$.01120 \pm .002069$	$.01221 \pm .001084$	$1.110 \pm .186$
.1539	$1.709 \pm .113$	$.02335 \pm .004906$	$.02407 \pm .003324$	$1.055 \pm .228$
.1964	$1.666 \pm .113$	$.04008 \pm .006944$	$.03848 \pm .003400$	$.976 \pm .162$

The separate seed-mean envelope below is descriptive. It does not provide seed uncertainty at fixed absolute target distance and must not be substituted for the paired reduction analysis above.

D.4 Image energy screens and corrected confirmation

Complete ImageNet entropy energy screens, the corrected matched-energy confirmation corresponding to Table 2, and a TSR–osmotic equivalence diagnostic.

Energy here is the summed-coordinate trapezoidal profile integral, without $1/2$. Screens have 128 samples and are distinct from 1024-sample evaluations and 50k quality metrics. The corrected closure comparison targets OG energy 5.36946.

The corrected confirmation achieves 9.23% contraction for OG, 5.45% for Constant, and 6.14% for TSR, at nearby rather than identical energies. The TSR–osmotic equivalence diagnostic at $k = 1.1$ uses 32 samples: maximum latent difference 9.798×10^{-4} (threshold 10^{-3}), and relative energy-curve error 7.899×10^{-6}.

Table 21: All exported absolute-target Pareto operating points from the seed-mean envelope.

Dataset	Target W_2	OG energy	Raw energy	MC energy
moons	1.671	0.001161	0.00073	0.006646
moons	1.671	0.0009182	0.0005977	0.005103
moons	1.671	0.0007331	0.0004653	0.00356
moons	1.672	0.0005479	0.000333	0.002018
moons	1.672	0.0003655	0.0002006	0.001319
checkerboard	1.648	0.04681	0.0411	0.04461
checkerboard	1.696	0.02798	0.02366	0.02852
checkerboard	1.743	0.01311	0.01209	0.01349
checkerboard	1.791	0.004345	0.004387	0.005922
checkerboard	1.838	0.000475	0.0004956	0.0005896

Table 22: moons: interpolated matched-energy distance, mean ± SD with available seed count in parentheses.

Energy	OG W_2 (n)	Raw W_2 (n)	MC W_2 (n)
0.0002006	1.672 ± 0.015 (3)	1.676 ± 0.018 (2)	1.673 ± 0.014 (3)
0.0008778	1.671 ± 0.014 (3)	1.671 ± 0.015 (3)	1.672 ± 0.014 (3)
0.00384	1.668 ± 0.015 (3)	1.667 ± 0.015 (3)	1.671 ± 0.013 (3)
0.0168	1.664 ± 0.016 (3)	1.661 ± 0.016 (3)	1.672 ± 0.01 (3)
0.07349	1.668 ± 0.024 (2)	1.652 ± 0.02 (3)	1.683 ± 0.011 (3)

Table 23: checkerboard: interpolated matched-energy distance, mean ± SD with available seed count in parentheses.

Energy	OG W_2 (n)	Raw W_2 (n)	MC W_2 (n)
0.0005896	1.835 ± 0.11 (3)	1.836 ± 0.11 (3)	1.775 ± 0.04 (2)
0.00176	1.817 ± 0.11 (3)	1.817 ± 0.11 (3)	1.821 ± 0.11 (3)
0.005254	1.785 ± 0.11 (3)	1.782 ± 0.11 (3)	1.795 ± 0.11 (3)
0.01568	1.733 ± 0.12 (3)	1.723 ± 0.11 (3)	1.734 ± 0.11 (3)
0.04681	1.703 ± 0.1 (2)	1.636 ± 0.12 (3)	1.642 ± 0.11 (3)

Table 24: mixture: interpolated matched-energy distance, mean ± SD with available seed count in parentheses.

Energy	OG W_2 (n)	Raw W_2 (n)	MC W_2 (n)
0.0001535	0.3988 (1)	0.3997 ± 0.058 (3)	0.3996 ± 0.058 (3)
0.0006102	0.4004 ± 0.058 (3)	0.4027 ± 0.06 (3)	0.4014 ± 0.062 (3)
0.002426	0.4052 ± 0.06 (3)	0.4135 ± 0.061 (3)	0.4069 ± 0.055 (3)
0.009647	0.4256 ± 0.061 (3)	0.4497 ± 0.061 (3)	0.4061 ± 0.05 (3)
0.03836	0.4858 ± 0.058 (3)	0.5559 ± 0.057 (3)	0.5377 (1)

Table 25: Complete 128-sample entropy screen.

Method	Strength	Energy	Contraction (%)
unguided	0	0	0
fisher	0.02	0.6053	3.306
fisher	0.04	2.408	6.378
fisher	0.06	5.388	9.252
fisher	0.08	9.524	11.94
fisher	0.1	14.79	14.56
const	0.1	2.572	3.948
const	0.15	5.854	5.804
const	0.2	10.5	7.558
const	0.25	16.47	9.234
const	0.35	32.25	12.34
const	0.5	64.53	16.49
tsr	1.05	4.465	5.788
tsr	1.08	11.21	8.737
tsr	1.1	17.26	10.51
tsr	1.12	24.47	12.16
tsr	1.15	37.31	14.39
tsr	1.2	63.39	17.64
tsr_osmotic	1.1	17.26	10.51

Table 26: All 1024-sample energy evaluations; these are not all matched-energy settings.

Method	Strength	Energy	Contraction (%)
unguided	0	0	0
fisher	0.06	5.369	8.792
const	0.25	16.39	8.738
const	0.15	5.848	5.482
tsr	1.08	11.16	8.253
tsr_osmotic	1.1	17.17	9.936
tsr	1.1	17.17	9.936

Table 27: Corrected 1024-sample matched-energy confirmation, with its own paired baseline.

Method	Strength	Energy	Contraction (%)
unguided	0	0	0
fisher	0.06	5.381	9.235
const	0.1437	5.271	5.45
tsr	1.055	5.325	6.138

D.5 Conditional quality and energy logs

Full conditional 5k quality metrics and both energy-log conventions for CFG and calibrated OG at multiple strengths.

All six confirmation rows contain 5000 samples. At similar high Top-1, calibrated guidance improves FID relative to CFG 2.5 but uses more logged energy under either available quadrature. CFG 1.5 has lower FID and lower adherence. These are operating-point trade-offs, not evidence that lower energy explains the quality improvement.

Table 28: Conditional 5k quality. No independent-seed error bars are available.

Method	Strength	FID	sFID	Prec.	Recall	Top-1	Conf.
cal	148.1	10.01	35.41	0.8572	0.6803	0.9478	0.439
cal	185.2	12.63	35.41	0.8818	0.6294	0.953	0.4505
cfg	1	15.92	37.63	0.6714	0.7876	0.7078	0.2848
cfg	1.5	7.731	35.19	0.808	0.7496	0.8696	0.3816
cfg	2	9.139	34.98	0.8608	0.6891	0.9236	0.4213
cfg	2.5	12.05	36.05	0.8888	0.62	0.9508	0.4408

D.6 Alignment and finite-difference diagnostics

Cosine-alignment validation of the canonical direction and finite-difference checks for reward and target-density observables.

Across 648 constructed directions, the normalized inner-product rate agrees with cosine at RMSE 4.591×10^{-8}; the aligned direction is maximal in all 81 configurations. This checks the same-field algebra, not an independent density-score estimate.

Reward agrees closely, while target log-density has a persistent direction/observable mismatch: the field subtracts the learned reference score, but the scalar observable is $\log q$ (Eq. B.18). Entropy finite differences were not run.

D.7 Preview regret, timing, and anchor-width diagnostic

Two-anchor preview selection regret, end-to-end timing/speedup, and a finite-range anchor-width RMSE diagnostic.

The 16-batch ImageNet confirmation measures latent-contraction selection regret, not Top-1 regret. All mean regrets are below two percentage points.

The naive complete-session cost model gives 546.642 seconds. Speedup averages per-session ratios and need not equal the ratio of displayed mean times. ODE-only deployment times are 304.740,

Table 29: Conditional 5k diagnostics. Both energy columns omit $1/2$ and use the hybrid unconditional reference in Eq. C.9. Negative contraction denotes expansion relative to CFG 1.

Method	Strength	Contraction (%)	Left-Riemann energy	Profile energy
cal	148.1	−9.25	749.4	630.6
cal	185.2	−15.89	1121	1240
cfg	1	0	61.16	63.05
cfg	1.5	−5.204	129.7	133.3
cfg	2	−11.84	222.9	228.3
cfg	2.5	−18.96	339.5	348.5

Table 30: Finite-difference diagnostics recomputed from the raw export.

Observable	η	RMSE	Mean ratio at $c = 1$	Rows
reward	0.0001	0.0003334	0.9999	216
reward	0.0003	0.000194	0.9998	216
reward	0.001	0.0005045	0.9995	216
reward	0.003	0.001502	0.9985	216
target_density	0.0001	0.08671	1.076	216
target_density	0.0003	0.08659	1.076	216
target_density	0.001	0.08626	1.075	216
target_density	0.003	0.08529	1.074	216

324.819, and 324.851 seconds, versus 539.984 seconds for the naive model. The one-time calibration cost is 2120.523 seconds; ODE-saving break-even estimates are 9.014, 9.855, and 9.857 sessions.

The available width sweep fixes the lower Fisher anchor at .02, uses 128 samples and interpolation fractions $\{.25, .5, .75\}$, and varies the upper anchor:

The log–log slope recomputed from these five midpoint values is 1.193. This finite-range diagnostic does not demonstrate quadratic empirical scaling; the C^2 bound in Appendix B.9 remains a theorem under its stated assumptions.

The separate 2D preview study uses 2048 samples, seeds 0–2, 100 midpoint steps (200 NFE), and a 20-step low-cost baseline. Its anchors are Fisher $(0, .15)$, Constant $(0, 2)$, and TSR $(1, 1.5)$. Paired Fisher-minus-baseline regret differences in percentage points are .001 $[-.038, .040]$ against Constant, $-.384$ $[-.499, -.269]$ against raw-parameter TSR, and $-.082$ $[-.148, -.016]$ against log-parameter TSR.

D.8 Historical 50k entropy quality

A historical ImageNet-256 entropy benchmark (50k samples) under reference-Fisher calibration, retained for completeness and uncertainty diagnostics.

The dated numerical export identifies this experiment as reference-Fisher calibration. Its selected quality values agree with those later summarized under OG. Reference-Fisher and Constant have nearly matched contraction; TSR is more contracted. Energy from a separate sweep must not be attached to these rows.

Table 31: ImageNet preview confirmation: means [reported 95% intervals]. Repeated-session times include final deployment as defined in the source cost model.

Method	Regret (pp)	Complete time (s)	Speedup
Constant	.195 [.11, .28]	312.75 [305.83, 319.67]	1.751 [1.714, 1.789]
Ref.-Fisher	.343 [.21, .48]	333.09 [329.14, 337.05]	1.642 [1.624, 1.660]
TSR-log	.220 [.14, .30]	333.09 [329.07, 337.11]	1.642 [1.624, 1.660]

Table 32: Midpoint latent interpolation error versus anchor width.

$\Delta\gamma$	RMSE	Relative error
0.02	0.007501	0.01004
0.04	0.01662	0.02235
0.06	0.02795	0.03808
0.08	0.03876	0.05311
0.1	0.05085	0.06992

Table 33: Historical ImageNet-256 entropy benchmark, $50{,}000$ generated samples per condition. These are pooled metrics, not seed means.

Method	FID	sFID	Precision	Recall	Top-1	Contr. (%)
Unguided	9.609	6.497	.66788	.6830	.70670	0
Ref.-Fisher, $\gamma = .06$	7.883	4.964	.69136	.6642	.72610	9.0691
Constant, $\lambda = .25$	8.314	5.920	.69314	.6510	.72762	9.0197
TSR, $k = 1.1$	8.773	6.924	.69054	.6442	.72230	10.2588

D.9 Historical sensitivity and within-class diagnostics

Historical strength sensitivity and within-class diagnostics for the frozen reference-Fisher variant; these are supporting sensitivity data, not new tests of online guided-marginal calibration.

Contraction increases throughout, but FID and Top-1 degrade from $.06$ to $.10$.

In an older five-seed moons sweep, reference-Fisher $\gamma = -.5$ produced extreme spread despite high recall, showing that high recall alone can conceal failure modes.

D.10 Unfinished matched-contraction ImageNet extension

Status of an unfinished 50k matched-contraction ImageNet extension; no finalized FID/Top-1 results are claimed.

As of the latest supplied status snapshot (24 September 2026), generation of 50k samples was in progress (TSR $k = 1.118$ at 12% contraction; job index $9/9$). Eight conditions were packed and the metrics directory was empty in the export. No finalized 50k FID or Top-1 result for this extension is claimed.

D.11 Full current target sweeps

Complete seed-aggregated target-density sweeps for checkerboard, mixture, and moons (all 78 rows), including unguided baselines.

Source energies omit $1/2$ and average coordinates, so they numerically equal paper energy in two dimensions.

Table 34: Paired reference-Fisher minus Constant differences from the historical 50k analysis. Entries are means [reported 95% intervals] for separately computed 10k block metrics.

Metric	All five blocks	New blocks 1–4 only
FID	$-.4154$ $[-.4982, -.3326]$	$-.4376$ $[-.5285, -.3468]$
Precision	$-.0018$ $[-.0035, -.0001]$	$-.0019$ $[-.0041, .0003]$
Recall	$.0078$ $[.0067, .0088]$	$.0083$ $[.0081, .0084]$
Top-1	$-.0015$ $[-.0035, .0005]$	$-.0014$ $[-.0039, .0011]$

Table 35: Historical 5k entropy sensitivity: same paired noise/labels across strengths, five images per class, CFG= 1, 100 NFE.

γ	FID	sFID	Precision	Recall	Top-1	Contr. (%)
0	15.783	37.214	.6636	.7805	.7024	0
.02	14.721	35.664	.6748	.7778	.7122	3.2202
.04	14.174	34.802	.6782	.7797	.7222	6.2524
.06	14.169	34.685	.6858	.7725	.7280	9.1011
.10	15.112	37.187	.6874	.7631	.7150	14.3370

Table 36: Historical within-class diagnostics, eight classes with 64 samples each.

γ	Top-1	Confidence	DINO pair	LPIPS	Latent variance
0	.814	.330	1.1054	.7206	.6107
.02	.820	.329	1.1084	.7303	.5899
.04	.812	.329	1.1120	.7397	.5705
.06	.803	.326	1.1146	.7484	.5519
.10	.758	.311	1.1222	.7622	.5175

Table 37: checkerboard: complete target sweep, part 1.

Method	Strength	Energy	W_2	SW	Mean log q
feng	0.01	$0.0005896 \pm 3.7 \times 10^{-5}$	1.838 ± 0.11	1.098 ± 0.058	-12.05 ± 0.47
feng	0.02	0.002344 ± 0.00014	1.813 ± 0.11	1.082 ± 0.057	-11.85 ± 0.46
feng	0.05	0.01443 ± 0.00091	1.737 ± 0.1	1.034 ± 0.055	-11.26 ± 0.42
feng	0.1	0.056 ± 0.0037	1.615 ± 0.096	0.9557 ± 0.05	-10.33 ± 0.35
feng	0.15	0.121 ± 0.0079	1.494 ± 0.088	0.8791 ± 0.045	-9.453 ± 0.29
feng	0.25	0.3075 ± 0.023	1.267 ± 0.072	0.734 ± 0.037	-7.889 ± 0.19
feng	0.4	0.662 ± 0.047	0.9587 ± 0.049	0.5417 ± 0.028	-6.001 ± 0.056
feng	0.6	1.115 ± 0.071	0.6589 ± 0.032	0.3596 ± 0.025	-4.309 ± 0.06
og	0.02	$4.235 \times 10^{-5} \pm 6.4 \times 10^{-7}$	1.855 ± 0.11	1.11 ± 0.059	-12.19 ± 0.48
og	0.05	$0.0002636 \pm 4 \times 10^{-6}$	1.843 ± 0.11	1.104 ± 0.058	-12.11 ± 0.48
og	0.08	$0.0006721 \pm 1 \times 10^{-5}$	1.833 ± 0.11	1.099 ± 0.058	-12.03 ± 0.47
og	0.1	$0.001047 \pm 1.6 \times 10^{-5}$	1.826 ± 0.11	1.095 ± 0.058	-11.97 ± 0.47
og	0.15	$0.002339 \pm 3.6 \times 10^{-5}$	1.809 ± 0.11	1.086 ± 0.058	-11.83 ± 0.45

Table 38: checkerboard: complete target sweep, part 2.

Method	Strength	Energy	W_2	SW	Mean log q
og	0.2	$0.004129 \pm 6.3 \times 10^{-5}$	1.792 ± 0.11	1.077 ± 0.058	-11.7 ± 0.45
og	0.3	0.009154 ± 0.00014	1.759 ± 0.11	1.06 ± 0.058	-11.44 ± 0.44
og	0.45	0.02013 ± 0.00031	1.716 ± 0.12	1.036 ± 0.058	-11.09 ± 0.42
og	0.7	0.04681 ± 0.00071	1.648 ± 0.12	0.9994 ± 0.057	-10.56 ± 0.39
raw	0.005	$0.0003375 \pm 9.9 \times 10^{-6}$	1.841 ± 0.11	1.102 ± 0.058	-12.08 ± 0.47
raw	0.01	$0.001337 \pm 4 \times 10^{-5}$	1.821 ± 0.11	1.092 ± 0.058	-11.91 ± 0.47
raw	0.02	0.005241 ± 0.00015	1.782 ± 0.11	1.07 ± 0.057	-11.58 ± 0.44
raw	0.035	0.01557 ± 0.00047	1.723 ± 0.11	1.039 ± 0.057	-11.11 ± 0.42
raw	0.05	0.03084 ± 0.00095	1.671 ± 0.11	1.01 ± 0.057	-10.67 ± 0.4
raw	0.08	0.0744 ± 0.0023	1.574 ± 0.12	0.9546 ± 0.057	-9.874 ± 0.35
raw	0.12	0.1551 ± 0.005	1.471 ± 0.12	0.8898 ± 0.056	-8.97 ± 0.31
raw	0.2	0.375 ± 0.013	1.334 ± 0.11	0.7981 ± 0.051	-7.588 ± 0.25
unguided	0	0 ± 0	1.863 ± 0.11	1.114 ± 0.059	-12.25 ± 0.49

Table 39: mixture: complete target sweep, part 1.

Method	Strength	Energy	W_2	SW	Mean log q
feng	0.01	$1.456 \times 10^{-5} \pm 9.1 \times 10^{-7}$	0.3977 ± 0.057	0.1022 ± 0.0032	-2.827 ± 0.021
feng	0.02	$6.111 \times 10^{-5} \pm 5.8 \times 10^{-6}$	0.3993 ± 0.056	0.1029 ± 0.0031	-2.82 ± 0.02
feng	0.05	$0.0003382 \pm 2.1 \times 10^{-5}$	0.4002 ± 0.06	0.1036 ± 0.0032	-2.803 ± 0.021
feng	0.1	0.001378 ± 0.00013	0.4044 ± 0.067	0.1067 ± 0.0064	-2.774 ± 0.022
feng	0.15	0.003187 ± 0.00025	0.407 ± 0.047	0.1056 ± 0.011	-2.743 ± 0.018
feng	0.25	0.007714 ± 0.001	0.3962 ± 0.056	0.1049 ± 0.013	-2.693 ± 0.011
feng	0.4	0.01919 ± 0.00079	0.4473 ± 0.043	0.1306 ± 0.0097	-2.617 ± 0.015
feng	0.6	0.03836 ± 0.0034	0.453 ± 0.093	0.1475 ± 0.018	-2.532 ± 0.025
og	0.02	$0.0001535 \pm 7.2 \times 10^{-6}$	0.3992 ± 0.057	0.1047 ± 0.00083	-2.819 ± 0.019
og	0.05	$0.0009492 \pm 4.4 \times 10^{-5}$	0.4015 ± 0.058	0.1087 ± 0.0035	-2.799 ± 0.017
og	0.08	0.002394 ± 0.00011	0.4053 ± 0.06	0.1142 ± 0.0064	-2.779 ± 0.017
og	0.1	0.003699 ± 0.00017	0.4087 ± 0.06	0.1187 ± 0.0081	-2.768 ± 0.017
og	0.15	0.008167 ± 0.00039	0.4224 ± 0.062	0.1326 ± 0.013	-2.74 ± 0.019

Table 40: mixture: complete target sweep, part 2.

Method	Strength	Energy	W_2	SW	Mean log q
og	0.2	0.01426 ± 0.0007	0.4363 ± 0.062	0.1473 ± 0.016	-2.713 ± 0.023
og	0.3	0.0309 ± 0.0017	0.4721 ± 0.06	0.1826 ± 0.02	-2.668 ± 0.031
og	0.45	0.06549 ± 0.0034	0.5367 ± 0.062	0.241 ± 0.022	-2.609 ± 0.044
og	0.7	0.1461 ± 0.0083	0.6691 ± 0.066	0.343 ± 0.028	-2.536 ± 0.063
raw	0.005	$0.0001256 \pm 2.1 \times 10^{-6}$	0.3995 ± 0.058	0.106 ± 0.0012	-2.826 ± 0.019
raw	0.01	$0.0004999 \pm 8.3 \times 10^{-6}$	0.402 ± 0.06	0.1105 ± 0.0037	-2.819 ± 0.017
raw	0.02	$0.001978 \pm 3.2 \times 10^{-5}$	0.411 ± 0.061	0.1226 ± 0.0078	-2.807 ± 0.015
raw	0.035	$0.005956 \pm 9.3 \times 10^{-5}$	0.4333 ± 0.062	0.146 ± 0.013	-2.793 ± 0.014
raw	0.05	0.01195 ± 0.00018	0.4598 ± 0.06	0.173 ± 0.015	-2.784 ± 0.012
raw	0.08	0.0295 ± 0.00042	0.5269 ± 0.056	0.2338 ± 0.018	-2.767 ± 0.0082
raw	0.12	0.06314 ± 0.00077	0.6371 ± 0.055	0.3194 ± 0.021	-2.758 ± 0.0045
raw	0.2	0.1577 ± 0.0016	0.8877 ± 0.045	0.4991 ± 0.021	-2.735 ± 0.009
unguided	0	0 ± 0	0.3984 ± 0.056	0.1026 ± 0.0029	-2.833 ± 0.021

Table 41: moons: complete target sweep, part 1.

Method	Strength	Energy	W_2	SW	Mean log q
feng	0.01	$6.804 \times 10^{-5} \pm 1.6 \times 10^{-6}$	1.673 ± 0.015	1.077 ± 0.0082	-4.231 ± 0.033
feng	0.02	$0.0002711 \pm 8.6 \times 10^{-6}$	1.673 ± 0.014	1.078 ± 0.0083	-4.212 ± 0.032
feng	0.05	$0.001679 \pm 4.8 \times 10^{-5}$	1.672 ± 0.014	1.079 ± 0.0085	-4.155 ± 0.031
feng	0.1	0.006646 ± 0.00016	1.671 ± 0.012	1.083 ± 0.0087	-4.061 ± 0.03
feng	0.15	0.01489 ± 0.00019	1.671 ± 0.01	1.087 ± 0.0083	-3.966 ± 0.032
feng	0.25	0.0396 ± 0.00093	1.674 ± 0.011	1.097 ± 0.0097	-3.785 ± 0.029
feng	0.4	0.09687 ± 0.0026	1.688 ± 0.011	1.119 ± 0.0085	-3.52 ± 0.024
feng	0.6	0.1981 ± 0.0071	1.73 ± 0.0075	1.156 ± 0.0061	-3.208 ± 0.011
og	0.02	$5.903 \times 10^{-5} \pm 6.5 \times 10^{-7}$	1.673 ± 0.015	1.077 ± 0.008	-4.237 ± 0.033
og	0.05	$0.0003695 \pm 4.3 \times 10^{-6}$	1.672 ± 0.015	1.078 ± 0.0079	-4.217 ± 0.033
og	0.08	$0.0009476 \pm 1.1 \times 10^{-5}$	1.671 ± 0.014	1.08 ± 0.0078	-4.196 ± 0.033
og	0.1	$0.001482 \pm 1.8 \times 10^{-5}$	1.671 ± 0.015	1.08 ± 0.0079	-4.183 ± 0.033
og	0.15	$0.003343 \pm 4.3 \times 10^{-5}$	1.669 ± 0.015	1.082 ± 0.0077	-4.151 ± 0.032

Table 42: moons: complete target sweep, part 2.

Method	Strength	Energy	W_2	SW	Mean log q
og	0.2	$0.005958 \pm 8.1 \times 10^{-5}$	1.667 ± 0.015	1.084 ± 0.0077	-4.118 ± 0.033
og	0.3	0.01346 ± 0.0002	1.664 ± 0.016	1.087 ± 0.0076	-4.054 ± 0.034
og	0.45	0.03039 ± 0.00047	1.662 ± 0.019	1.093 ± 0.0079	-3.958 ± 0.033
og	0.7	0.07349 ± 0.0012	1.66 ± 0.022	1.103 ± 0.0084	-3.808 ± 0.032
raw	0.005	$0.0002006 \pm 1.1 \times 10^{-5}$	1.672 ± 0.015	1.078 ± 0.0081	-4.226 ± 0.033
raw	0.01	$0.0008089 \pm 4.7 \times 10^{-5}$	1.671 ± 0.015	1.079 ± 0.0079	-4.201 ± 0.033
raw	0.02	0.003293 ± 0.0002	1.668 ± 0.015	1.081 ± 0.0079	-4.152 ± 0.032
raw	0.035	0.01036 ± 0.00067	1.663 ± 0.015	1.084 ± 0.008	-4.08 ± 0.033
raw	0.05	0.02174 ± 0.0015	1.659 ± 0.016	1.087 ± 0.0079	-4.008 ± 0.032
raw	0.08	0.05893 ± 0.0045	1.653 ± 0.019	1.094 ± 0.0081	-3.866 ± 0.03
raw	0.12	0.1433 ± 0.012	1.647 ± 0.022	1.102 ± 0.0087	-3.688 ± 0.025
raw	0.2	0.4545 ± 0.044	1.641 ± 0.028	1.113 ± 0.0083	-3.378 ± 0.023
unguided	0	0 ± 0	1.674 ± 0.015	1.077 ± 0.0082	-4.25 ± 0.033