WHAT DO FLOW-BASED INVERSE SOLVERS APPROXIMATE? A POSTERIOR-TRANSPORT VIEW

Anonymous authors
Paper under double-blind review

ABSTRACT

A growing family of training-free solvers—FlowDPS, FLOWER, PnP-Flow and their diffusion ancestors (DPS, DAPS)—repurpose a pretrained flow-matching prior to solve imaging inverse problems by adding a measurement-guidance term to the deterministic probability-flow ODE. Despite strong empirical results, what these per-step corrections actually approximate—and how far the resulting samples are from the true posterior $p(\mathbf{x} \mid \mathbf{y})$—has not been characterized. We give a posterior-transport account of flow-based inverse problem solving. Our starting point is a simple but consequential fact: for a *deterministic* flow prior, Bayesian conditioning is realized entirely by a *reweighting of the source distribution*, not by a drift correction; pushing the reweighted source through the *unmodified* velocity field yields exact posterior samples. From this we define a canonical minimum-energy corrector that exactly transports the unconditional source to the posterior. Existing guidance methods can then be placed within the same correction framework as *local surrogate* correctors—differing in their endpoint estimate and mobility—and their discrepancy from the canonical corrector controls the terminal posterior bias in Wasserstein distance. A controlled 2D study with a closed-form posterior supports the theory: source reweighting is comparable to a two-sample Monte-Carlo reference on every metric (within a factor of 2, below it on sliced-W_2), whereas trajectory guidance incurs substantially larger error—6–14$\times$ that reference on sliced-W_2 and 200–800$\times$ on energy distance and MMD—and collapses posterior modes *across all tested guidance strengths*. Guided by the analysis we propose a cheap, principled velocity-correction solver that is competitive across two in-domain priors (AFHQ, CelebA) and two out-of-distribution settings while, unlike point-estimate source-space optimizers, producing diverse conditional samples whose sample-based uncertainty correlates with reconstruction error (moderately so once the inpainting mask is controlled for).

1 INTRODUCTION

Flow matching (Lipman et al., 2023; Liu et al., 2023; Albergo & Vanden-Eijnden, 2023) has become a state-of-the-art framework for generative modeling, learning a velocity field $\boldsymbol{v}_t$ whose probability-flow ODE transports a simple source p_0 to a data distribution p_1. A pretrained flow model is an implicit prior over images, and a large recent literature reuses it—without retraining—to solve linear and nonlinear inverse problems such as super-resolution, deblurring and inpainting (Kim et al., 2025; Pourya et al., 2026; Martin et al., 2025; Ben-Hamu et al., 2024). These solvers all share a template inherited from diffusion posterior sampling (Chung et al., 2023; Zhang et al., 2025): integrate the prior flow ODE and, at each step, nudge the trajectory toward the measurement $\mathbf{y}$ with a likelihood-based correction.

This template works well empirically, but it is built bottom-up by analogy to diffusion guidance, and a basic question has gone unanswered: *what does the per-step correction approximate, and how biased is the resulting sampler relative to the true posterior $p(\mathbf{x} \mid \mathbf{y})$?* Perceptual metrics (LPIPS, FID) reward producing *a* plausible reconstruction, and so they mask the distinction between sampling the posterior and merely landing on its typical set. For reliability-critical restoration—and for any use of these solvers to quantify uncertainty—the distinction matters.

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

We take a top-down, posterior-transport view. The key observation is that the *deterministic* nature of the probability-flow ODE changes the structure of Bayesian conditioning relative to the stochastic (diffusion) case. For an SDE prior, conditioning on $\mathbf{y}$ adds a Doob h-transform drift to the dynamics (Denker et al., 2024). For a deterministic flow, we show there is *no canonical added drift*: the exact posterior is obtained by leaving the velocity field untouched and *reweighting the source distribution* by the endpoint likelihood (Section 4, Proposition 1). Conditioning lives in the source, not in the trajectory.

This reframing has three consequences. (i) It explains what trajectory-guidance solvers are doing: because sampling the reweighted source is intractable, they start from the *unconditional* source and inject a correction field to morph it into the posterior; we identify the canonical (minimum-kinetic-energy) correction and show FlowDPS, FLOWER and PnP-Flow can be read as distinct local approximations within this correction framework (Proposition 3). (ii) It yields a posterior-bias bound in Wasserstein distance controlled by the correction-estimation error (Theorem 1). (iii) It tells us how existing solvers fail—on multimodal posteriors a local correction can systematically mis-weight the modes, with mode collapse as the extreme case—and what a faithful solver should instead do.

We validate the theory on a 2D problem with a closed-form Gaussian-mixture posterior, removing all model error: source reweighting reaches the two-sample Monte-Carlo reference on sliced-Wasserstein, energy distance, MMD and mode-weight recovery, while trajectory guidance is 6–$14\times$ that reference on sliced-W_2 and 200–$800\times$ on energy distance and MMD, and collapses modes for *every* guidance strength (Section 6). On AFHQ, CelebA, and out-of-distribution image restoration we reproduce the standard benchmark and, guided by the analysis, give a cheap velocity-correction solver (no ODE backpropagation) that is competitive on quality while producing diverse *conditional* samples whose sample-based uncertainty correlates with the error—something point-estimate source-space optimizers (e.g. D-Flow) cannot provide (Section 7).

Contributions.

- **Exact posterior transport for flow priors.** We prove that, for a deterministic flow prior, the posterior is the pushforward of a reweighted source under the *unmodified* velocity field—conditioning is source reweighting, not drift (Proposition 1).
- **A unifying view of flow inverse solvers.** We characterize the canonical correction field and place FlowDPS, FLOWER and PnP-Flow within it as distinct local surrogates, differing in endpoint estimate and mobility (Proposition 3), with a Wasserstein posterior-bias bound on any such corrector (Theorem 1).
- **Empirical confirmation and a practical solver.** A closed-form 2D study confirms the theoretical predictions in a controlled setting; on image restoration our analysis-guided solver matches strong baselines in quality while delivering error-correlated sample-based uncertainty.

2 Related work

Diffusion-based inverse solvers. The dominant template is likelihood-gradient guidance: DPS (Chung et al., 2023) adds a Tweedie-endpoint manifold-constrained gradient to the reverse diffusion, MCG (Chung et al., 2022) augments it with a manifold projection, ΠGDM (Song et al., 2023) replaces the point likelihood by an isotropic-Gaussian (pseudoinverse) approximation, and DDNM (Wang et al., 2023) projects onto the measurement null-space. A second line reduces the greedy per-step error our analysis flags: DAPS (Zhang et al., 2025) decouples consecutive steps with noise annealing and an MCMC (Langevin/HMC) update so the time-marginals anneal to the posterior; MGPS (Moufad et al., 2025) performs variational posterior sampling with midpoint guidance; and SITCOM (Alkhouri et al., 2025) is optimization-based, enforcing measurement, forward- and *backward*-diffusion consistency by optimizing the model input at each step. Provably consistent SMC/annealing samplers also exist (Xu & Chi, 2024; Wu et al., 2023; Cardoso et al., 2024). Our viewpoint organizes these: pure diffusion guidance (DPS/MCG/ΠGDM/DDNM) is the diffusion analogue of the greedy local correctors Proposition 3 formalizes for FlowDPS, FLOWER and PnP-Flow; the annealing/MCMC methods (DAPS/MGPS) and the optimization methods (SITCOM, and D-Flow below) are partial moves *toward the source side*—trading per-step locality for exploration or latent optimization— which our source-transport analysis (Section 4) identifies as the exact object.

Flow-matching inverse solvers. FlowDPS (Kim et al., 2025), FLOWER (Pourya et al., 2026) and PnP-Flow (Martin et al., 2025) port the guidance template to deterministic flow priors; Feng et al. (2025) instead unify guidance-by-velocity-modification—modifying $\boldsymbol{v}_t$ where we hold it fixed. We do not propose another such solver; we explain what their corrections approximate (Proposition 3), bound the posterior bias (Theorem 1), and show that for a deterministic flow conditioning is a source reweighting rather than a drift.

Source-space and exact-posterior methods. D-Flow (Ben-Hamu et al., 2024) optimizes the source latent, returning a single MAP-like point. Wang et al. (2025) and Kalaivanan et al. (2025) independently establish source-space exact-transport identities closely related to our Lemma 1; Appendix F details the overlap and our dynamical contribution beyond it. On the diffusion side, provably consistent posterior samplers with bounds exist (Xu & Chi, 2024); they target SDE priors and do not provide the flow velocity-plus-correction decomposition or the unification of flow solvers.

h-transforms and guidance. Doob h-transforms underpin conditional diffusion training and guidance (Denker et al., 2024); the deterministic-flow case we treat is structurally different (Section 4).

3 Background and problem setup

Flow-matching priors. A flow-matching model learns a time-dependent velocity field $\boldsymbol{v}_t : \mathbb{R}^d \to \mathbb{R}^d$, $t \in [0, 1]$, such that the probability-flow ODE

$$\frac{d\mathbf{x}_t}{dt} = \boldsymbol{v}_t(\mathbf{x}_t), \qquad \mathbf{x}_0 \sim p_0 = \mathcal{N}(0, I), \tag{1}$$

transports the source p_0 to the data distribution $p_1 \approx p_{\text{data}}$. Let $\Phi_{s\to t}$ denote the flow map of equation 1 (the solution operator from time s to t); it is a deterministic diffeomorphism, and $p_t = (\Phi_{0\to t})_{\#}p_0$. We write $\Phi := \Phi_{0\to 1}$, so $p_1 = \Phi_{\#}p_0$.

Inverse problems. We observe $\mathbf{y} = \mathcal{A}(\mathbf{x}) + \boldsymbol{\eta}$, $\boldsymbol{\eta} \sim \mathcal{N}(0, \sigma^2 I)$, with a known (possibly non-invertible) forward operator $\mathcal{A}$ and likelihood $p(\mathbf{y} \mid \mathbf{x})$. Using the flow model as prior, the Bayesian posterior is $p(\mathbf{x} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{x})\, p_1(\mathbf{x})$. We want samples from $p(\mathbf{x} \mid \mathbf{y})$, not just a single reconstruction.

Trajectory-guidance solvers. The dominant template integrates equation 1 from $\mathbf{x}_0 \sim p_0$ while adding a measurement correction: $\frac{d\mathbf{x}_t}{dt} = \boldsymbol{v}_t(\mathbf{x}_t) + \mathbf{u}_t(\mathbf{x}_t)$, where $\mathbf{u}_t$ is built from a one-step endpoint estimate $\hat{\mathbf{x}}_1(\mathbf{x}_t) = \mathbf{x}_t + (1-t)\boldsymbol{v}_t(\mathbf{x}_t)$ (a flow Tweedie formula) and the likelihood gradient $\nabla_{\mathbf{x}_t} \log p(\mathbf{y} \mid \hat{\mathbf{x}}_1)$. FlowDPS (Kim et al., 2025) uses this gradient directly; FLOWER (Pourya et al., 2026) replaces it with an isotropic-Gaussian endpoint-posterior projection solved by conjugate gradients; PnP-Flow (Martin et al., 2025) alternates a data-fidelity step with flow-path reprojection and denoising. D-Flow (Ben-Hamu et al., 2024) is the exception: it optimizes the *source* latent $\mathbf{z}$ to fit $\mathbf{y}$, returning a single MAP-like point.

4 Exact posterior transport under a flow prior

We first establish that, for a deterministic flow prior, conditioning is a source operation. Throughout, Φ is the prior flow map and $p_0 = \mathcal{N}(0, I)$.

Lemma 1 (Posterior = reweighted source; classical)**.** *Define the* source posterior

$$p_0^{\mathbf{y}}(\mathbf{z}) \;\propto\; p\big(\mathbf{y} \mid \Phi(\mathbf{z})\big)\, p_0(\mathbf{z}). \tag{2}$$

Then $\mathbf{x} = \Phi(\mathbf{z})$ with $\mathbf{z} \sim p_0^{\mathbf{y}}$ is distributed exactly as $p(\mathbf{x} \mid \mathbf{y})$, i.e. $p(\cdot \mid \mathbf{y}) = \Phi_{\#}p_0^{\mathbf{y}}$.

Lemma 1 is the elementary change-of-variables identity underlying latent-space Bayesian inference and normalizing-flow posterior inference; we claim no novelty for it and state it only to fix notation. Our contribution is what it implies for the *dynamics*, which is specific to flow matching and, to our knowledge, has not been made explicit. Define the time-t reweighting $r_t(\mathbf{x}) := p(\mathbf{y} \mid \Phi_{t\to 1}(\mathbf{x}))/p(\mathbf{y})$.

Proposition 1 (Conditioning is a prior-flow invariant)**.** *The posterior marginal is the reweighted prior marginal, $p_t^{\mathbf{y}} = r_t\, p_t$, and r_t is conserved along prior-flow characteristics, $\partial_t r_t + \boldsymbol{v}_t \cdot \nabla r_t = 0$. Consequently the posterior marginal path is transported by the* unmodified *prior velocity,*

$$\partial_t p_t^{\mathbf{y}} + \nabla \cdot \big(p_t^{\mathbf{y}}\, \boldsymbol{v}_t\big) = 0, \qquad p_1^{\mathbf{y}} = p(\cdot \mid \mathbf{y}), \tag{3}$$

i.e. no drift correction is required (proof in Appendix A).

The conservation $\frac{D}{Dt} r_t = 0$ holds because the endpoint $\Phi_{t\to 1}(\mathbf{x}_t) = \mathbf{x}_1$ is constant along a deterministic characteristic: the "filtering" expectation that defines the diffusion h-function collapses to a point evaluation. This yields our central conceptual point, which sharply separates the flow case from the diffusion case it is usually derived by analogy to.

Proposition 2 (Zero-drift dichotomy / small-noise h-transform limit)**.** *For the family of noisy bridges $d\mathbf{x}_t = [\boldsymbol{v}_t(\mathbf{x}_t) + \epsilon^2 \nabla \log h_t^\epsilon(\mathbf{x}_t)]\, dt + \epsilon\, dW_t$ conditioned on $\mathbf{y}$ at $t = 1$, the Doob drift correction $\epsilon^2 \nabla \log h_t^\epsilon$ vanishes pointwise as $\epsilon \to 0$, while the endpoint constraint is enforced in the limit through the reweighted initial law $p_0^{\mathbf{y}}$. Hence in the deterministic ($\epsilon = 0$) flow limit, conditioning carries* no *drift and is realized entirely by source reweighting (proof in Appendix A).*

Proposition 2 shows that copying the diffusion Doob h-transform drift into the small-noise limit vanishes; it does *not* show that no drift-based corrector can be exact (one does, $\mathbf{u}^\star$ below)—only that existing solvers, built by that analogy rather than by solving equation 4, have no reason to land on it (Appendix A states this precisely).

Why solvers add a drift anyway. Proposition 1 also exposes the catch: it requires *initializing at the reweighted source* $p_0^{\mathbf{y}}$. One can sample $p_0 = \mathcal{N}(0, I)$, but $p_0^{\mathbf{y}}$ is intractable (it requires the full flow map inside the likelihood), and naive importance reweighting of p_0 suffers from vanishing effective sample size in high dimension. Trajectory-guidance solvers sidestep this by starting from the *unconditional* p_0 and adding a correction field $\mathbf{u}_t$ to bend the marginals from p_t to $p_t^{\mathbf{y}}$. The next section makes this precise.

5 Flow inverse solvers as approximate posterior corrections

Suppose we integrate the corrected ODE $\frac{d\mathbf{x}_t}{dt} = \boldsymbol{v}_t(\mathbf{x}_t) + \mathbf{u}_t(\mathbf{x}_t)$ from $\mathbf{x}_0 \sim p_0$ and let $q_t^{\mathbf{u}}$ be the resulting marginals. We seek $\mathbf{u}_t$ that lands on the posterior, $q_1^{\mathbf{u}} = p(\cdot \mid \mathbf{y})$. Among all such fields, the canonical choice is the one of minimum kinetic energy:

$$\mathbf{u}_t^\star = \arg\min_{\mathbf{u}} \int_0^1 \mathbb{E}_{q_t^{\mathbf{u}}} \big\|\mathbf{u}_t(\mathbf{x}_t)\big\|^2\, dt \quad \text{s.t.} \quad q_1^{\mathbf{u}} = p(\cdot \mid \mathbf{y}). \tag{4}$$

This is a Benamou–Brenier / Schrödinger-type control problem: $\mathbf{u}^\star$ is the optimal-transport correction taking the prior path to the posterior path. It is the deterministic analogue of the Doob/Föllmer drift, but defined through a chosen mobility (here the Euclidean metric) rather than a diffusion coefficient.

We stress what is and is not being claimed. *Any* field with $q_1^{\mathbf{u}} = p(\cdot \mid \mathbf{y})$ is a valid corrector; $\mathbf{u}^\star$ is merely the *canonical* (minimum-energy) representative, and Theorem 1 bounds the bias of *any* corrector by its discrepancy from $\mathbf{u}^\star$. Existing solvers do not attempt to compute $\mathbf{u}^\star$; they follow a likelihood gradient and are *greedy, local* correctors that in general do *not* reach the posterior at all. So Proposition 3 should be read not as "they approximate the single object $\mathbf{u}^\star$" but as "they are local likelihood-gradient correctors whose gap to the canonical exact corrector $\mathbf{u}^\star$ is what Theorem 1 controls and what Section 6 measures."

Proposition 3 (Existing solvers as local corrections in the framework of equation 4)**.** *Write the generic correction as $\mathbf{u}_t(\mathbf{x}) = M_t(\mathbf{x})\, \nabla_{\mathbf{x}} \log p\big(\mathbf{y} \mid \hat{\mathbf{x}}_1(\mathbf{x}, t)\big)$ for a mobility operator M_t and endpoint estimate $\hat{\mathbf{x}}_1$. Then:*

1. ***FlowDPS*** *uses $\hat{\mathbf{x}}_1 = \mathbf{x} + (1-t)\boldsymbol{v}_t(\mathbf{x})$ (flow Tweedie) and a scalar mobility $M_t = \zeta_t I$, i.e. a zeroth-order (point-mass endpoint) instance of the template;*
2. ***FLOWER*** *replaces the point mass by an isotropic-Gaussian endpoint* surrogate *$\mathcal{N}(\hat{\mathbf{x}}_1, \sigma_{r,t}^2 I)$ and sets M_t to the corresponding Gaussian covariance (a* Π*GDM/CG projection);*

3. ***PnP-Flow*** *replaces the gradient step by a data-fidelity proximal map followed by flow-path reprojection, a splitting approximation of the same correction.*

All three replace the exact endpoint map $\Phi_{t\to 1}$ and the global transport problem in equation 4 by local endpoint and mobility surrogates, and therefore incur a nonzero gap $\mathbf{u}_t - \mathbf{u}_t^\star$ even with exact velocity fields. We avoid calling $\hat{\mathbf{x}}_1$ an approximate "endpoint posterior": under the deterministic prior ODE that law is a Dirac mass (Appendix A). This is a structural *placement, not an approximation result: we claim no rate of the form $\mathbf{u}^{\textit{FlowDPS}} = \mathbf{u}^\star + O(\epsilon)$. The proposition supplies the common template—endpoint estimate plus mobility—within which the three solvers are distinct local choices; Theorem 1 converts a gap in that template into a posterior-bias bound.*

Theorem 1 (Posterior-bias bound)**.** *Let $p_1^{\mathbf{u}}$ be the terminal law of the corrected ODE with an approximate field $\mathbf{u}$, and let $p(\cdot \mid \mathbf{y})$ be the true posterior. Assuming $\boldsymbol{v}_t + \mathbf{u}_t^\star$ and $\boldsymbol{v}_t + \mathbf{u}_t$ are uniformly Lipschitz in $\mathbf{x}$ (Assumption 1, Appendix A),*

$$W_2\big(p_1^{\mathbf{u}},\, p(\cdot \mid \mathbf{y})\big) \;\le\; L\int_0^1 \big(\mathbb{E}_{q_t^{\mathbf{u}}}\|\mathbf{u}_t - \mathbf{u}_t^\star\|^2\big)^{1/2} dt \;+\; \epsilon_{\mathrm{disc}} + \epsilon_{\mathrm{flow}}, \tag{5}$$

where ϵ_{disc} is ODE discretization error and ϵ_{flow} is the flow-model approximation error.

The bound in equation 5 is a one-directional stability guarantee (small gap ⇒ small bias; a large gap does *not* imply large bias, and it supplies no rate at which existing solvers approach $\mathbf{u}^\star$—see the discussion in Appendix A). It identifies the right quantity, not a computed one: we do not solve for $\mathbf{u}^\star$ itself or measure $\|\mathbf{u}_t - \mathbf{u}_t^\star\|$ directly anywhere in this paper. Section 6 instead measures the bound's left-hand side—the resulting terminal-distribution error against the closed-form posterior—which is a consequence of, but not a direct measurement of, the correction gap; it is substantial for likelihood-gradient guidance and non-vanishing across all tested guidance strengths.

A large-deviation reading. In the small-observation-noise limit ($\mathbf{y} = A(\mathbf{x}) + \sqrt{\epsilon}\,\boldsymbol{\eta}$, prior held fixed) the source posterior in equation 2 is a tilted measure that, by Laplace's method, concentrates on a categorical mixture over the measurement-consistent basins with weights $w_k \propto p_0(\mathbf{z}_k^\star)\,|\nabla^2(\ell_{\mathbf{y}} \circ \Phi)(\mathbf{z}_k^\star)|^{-1/2}$, $\ell_{\mathbf{y}}(\mathbf{x}) = \frac{1}{2}\|A(\mathbf{x}) - \mathbf{y}\|^2$ (Appendix B.1). This asymptotic picture explains two things we observe. Exact source reweighting is unbiased, but its effective sample size can decay exponentially (Proposition 6), which is why it cannot be run at image scale. And a greedy likelihood-gradient correction transports mass by local ascent, so the weights that survive are those of its *regions of attraction* rather than the posterior's Laplace weights—a guidance-strength-independent mismatch $\frac{1}{2}\sum_k |\alpha_k - w_k|$ in an idealized caricature, and the milder mis-weighting we measure for real solvers, whose prior velocity still spreads mass across modes (Table 1). Finally, $\mathbf{u}^\star$ is the deterministic-limit minimum-action (Freidlin–Wentzell) control (Kifer, 1988; Dembo & Zeitouni, 2010) steering the *unconditional* source to the posterior, which situates the picture in the Schrödinger-bridge / Föllmer continuum (Appendix F). These statements are asymptotic and explanatory, not finite-σ guarantees.

6 A CONTROLLED STUDY WITH A CLOSED-FORM POSTERIOR

To test the theory free of model error we use a 2D flow prior whose target is a random 8-component Gaussian mixture; the conditional Euler velocity is available in closed form, so the flow map is essentially exact. With a linear-Gaussian measurement the true posterior is again a Gaussian mixture, available in closed form. We compare against this ground truth: (i) **exact source reweighting** (oracle)—integrate the unmodified ODE from p_0 and weight by the endpoint likelihood (Lemma 1); and (ii) **FlowDPS-style guidance**— the Tweedie-endpoint likelihood-gradient correction, swept over guidance strength. Metrics are sliced-Wasserstein (W_2), energy distance, multi-scale MMD^2, and mode-weight ℓ_1 error, against a *two-sample Monte-Carlo reference*: the same statistics between two independent draws of the true posterior—a finite-sample reference scale, not a lower bound, since our weighted estimator can fall below it on sliced-W_2.

Scope of this comparison. All methods here act on the *same* analytic flow prior with a closed-form posterior, so differences reflect the solver and not the prior. This is the only place such a comparison against diffusion-native solvers is unconfounded—a shared imaging table would mix

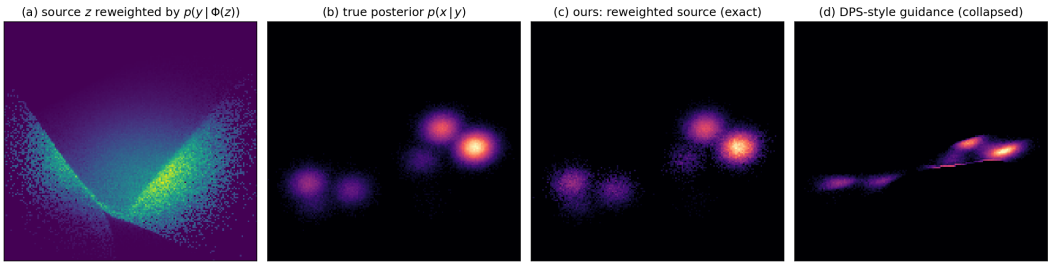


Figure 1: **Conditioning lives in the source** (densities, 8-mode 2D GMM prior, linear measurement). **(a)** The source $\mathcal{N}(0, I)$ reweighted by the endpoint likelihood $p(\mathbf{y} \mid \Phi(\mathbf{z}))$—conditioning is a reweighting on the source. **(b)** Closed-form true posterior. **(c)** Exact source reweighting (oracle) pushed through the *unmodified* velocity field—visually indistinguishable from (b). **(d)** DPS-style trajectory guidance collapses the multimodal posterior onto a single measurement-consistent ridge.

Table 1: 2**D study against the closed-form posterior** (K=8 modes, 12k samples). Exact source reweighting is comparable to the two-sample Monte-Carlo reference on every metric—within a factor of 2, and below it on sliced-W_2 (a single draw of this reference, not a distribution over repeated trials, so this is not a claim of statistical equivalence)—while trajectory guidance is 6–14$\times$ that reference on sliced-W_2 and 200–800$\times$ on energy distance and MMD^2, and badly misestimates mode weights, for *all* guidance strengths g. Our own image solver's update rule lands in the same regime as guidance, not near the exact-reweighting floor.

Method	sliced-W_2	energy dist.	MMD^2	mode-weight ℓ_1
Two-sample MC reference (GT vs GT)	0.060	0.0004	1.5×10^{-4}	0.024
Exact source reweighting (oracle)	**0.032**	**0.0006**	$\mathbf{2.1 \times 10^{-4}}$	**0.028**
FlowDPS-style guidance, g=0.5	0.71	0.122	4.4×10^{-2}	0.53
FlowDPS-style guidance, g=1.0	0.38	0.087	3.0×10^{-2}	0.16
FlowDPS-style guidance, g=2.0	0.53	0.243	7.8×10^{-2}	0.52
FlowDPS-style guidance, g=4.0	0.84	0.307	9.5×10^{-2}	0.56
DAPS (Zhang et al., 2025) (annealing)	0.87	0.076	1.5×10^{-2}	0.29
MGPS (Moufad et al., 2025) (midpoint)	0.45	0.092	3.2×10^{-2}	0.17
Ours (velocity-correction solver, γ retuned)	0.38	0.115	3.7×10^{-2}	0.29

prior and solver—so the imaging tables (Section 7) stay within the flow-prior family. We reproduce the mechanisms of DAPS (Zhang et al., 2025) (decoupled annealing with an exact Gaussian endpoint-posterior inner step, favourable to DAPS) and MGPS (Moufad et al., 2025) (midpoint guidance). Both improve on pure likelihood-gradient guidance, yet remain $\sim$100–230$\times$ the Monte-Carlo reference on energy distance and MMD^2 and still mis-weight modes ($\ell_1 = 0.17$–0.29): even an annealing/MCMC solver that provably targets the posterior in the limit is, at practical budgets, far from exact source reweighting. We do *not* report a toy number for the optimization-based SITCOM (Alkhouri et al., 2025), whose objective is defined for the diffusion setup and whose 2D port was unstable; reporting it would be a strawman (Appendix C).

Table 1 and Figure 1 support the theory. Source reweighting through the unmodified velocity field is comparable to the two-sample Monte-Carlo reference on all four metrics, consistent with Proposition 1: the conditioning is captured entirely in the source. Trajectory guidance, by contrast, is one order of magnitude worse on sliced-W_2 and two-to-three orders of magnitude worse on energy distance and MMD^2, and collapses the multimodal posterior onto a single ridge (Figure 1d); increasing the guidance strength does not fix this—the error plateaus and then worsens—consistent with the correction-gap perspective of Theorem 1 and with the mode-weight analysis of Proposition 8. Our own solver's exact update rule (Algorithm 1, γ retuned for this scale) lands in the *same* regime as guidance here rather than near the floor—an honest negative result, and the expected one, since Section 7.1 already treats it as a heuristic.

Table 2: **AFHQ-Cat restoration** (10 shared images; PSNR ↑ / LPIPS ↓; every method returns one reconstruction, ours a *single* sample). Our velocity-correction solver attains the best perceptual quality (LPIPS) on box inpainting and is competitive throughout, at the lowest cost of any method measured (Table 3). Best / second-best LPIPS per column in **bold** / underline.

Method	Super-resolution	Gaussian deblur	Box inpainting
FLOWER (Pourya et al., 2026)	25.95 / 0.276	27.25 / 0.301	25.65 / 0.077
OT-ODE	25.18 / **0.107**	26.40 / **0.117**	25.23 / 0.083
Flow-Priors	23.36 / 0.276	25.85 / 0.183	26.15 / 0.125
D-Flow (Ben-Hamu et al., 2024)	24.11 / 0.180	26.70 / 0.173	24.32 / 0.085
PnP-Flow (Martin et al., 2025)	–	–	25.44 / 0.116
Ours (velocity correction)	25.07 / 0.122	25.53 / 0.130	25.57 / **0.052**

7 IMAGE RESTORATION

Setup. We use AFHQ-Cat 256×256 with a pretrained optimal-transport flow-matching prior, and the standard inverse problems of the FLOWER benchmark (Pourya et al., 2026): $\times4$ super-resolution, Gaussian deblurring, and box inpainting, all at measurement noise $\sigma=0.05$. We report PSNR and LPIPS. All methods share the same pretrained prior.

FlowDPS is absent from Table 2 for an implementation reason: its released code is built around Stable Diffusion 3, a *text-conditioned latent* flow model, while this benchmark's priors are unconditional and pixel-space (Appendix C). We evaluate its *mechanism* in the 2D study instead, where it can be isolated exactly (Table 1).

An analysis-guided solver. Proposition 3 says a faithful solver should (i) use a well-chosen mobility M_t rather than a fixed scalar step, and (ii) preserve, rather than collapse, posterior multimodality. We adopt a cheap velocity-correction scheme that integrates the prior ODE from a *random* source draw and applies a per-step, mobility-modulated, *normalized* measurement correction (a discrete per-step correction rather than an Euler discretization of $\boldsymbol{v}_t + \mathbf{u}_t$; see Appendix E); it requires one forward/backward through the endpoint per step—*no* backpropagation through the ODE, unlike source-space optimizers such as D-Flow. Different source draws give different conditional samples, from which we form a sample mean and a per-pixel sample-based uncertainty map. We say *conditional* rather than *posterior* samples deliberately: only the exact source reweighting of Section 4 is a posterior sampler, and it does not scale (Section 7.1).

Table 2 shows our solver is competitive on distortion (PSNR) and strong on perceptual quality: it attains the *best* LPIPS on box inpainting (0.052, vs. 0.077 for FLOWER) and the second-best on super-resolution and deblurring, behind only OT-ODE. We do not claim uniform state-of-the-art on single-reconstruction metrics—on these constraining problems several methods cluster closely—because that is precisely the regime where such metrics are least discriminative (Section 6). The case for our solver is made on the axes those metrics ignore, which we isolate next.

Statistical significance and sample size. Paired Wilcoxon signed-rank tests on per-image LPIPS ($n=10$) give **ours** $<$ **FLOWER** on super-resolution and deblurring ($p=0.002$ each) and marginally on inpainting ($p=0.064$); against OT-ODE the picture is mixed (inpainting $p=0.004$ for us, super-resolution $p=0.027$ against, deblurring tied). A larger-scale re-run ($n=50$ for the scalable baselines; $n=15$ for D-Flow, given its cost) preserves all qualitative comparisons involving our method (Appendix C.1)—the inpainting gain over FLOWER, only marginal here, becomes highly significant ($p=2\times10^{-8}$)—and raising D-Flow from $n=5$ to $n=15$ turns its favorable trend into significance on all three operators. Our reconstruction quality is thus competitive—ahead of FLOWER perceptually, on par with OT-ODE—not uniformly state-of-the-art, consistent with our thesis that single-reconstruction metrics are not where the contribution lies.

Can the baselines be ensembled too? Yes—as the theory predicts. Conditional samples are not a property of our solver: Proposition 1 places the conditional variability in the *source draw*, so *any* solver integrating from a random source inherits it. Rerunning FLOWER and OT-ODE with 8 independent initial-noise draws per image (measurement noise fixed) and ensembling them as

Table 3: **Ensembling the baselines, and cost.** Every solver that starts from a random source can be rerun with independent initial noise; we did so (8 draws per image, measurement noise fixed, 10 AFHQ inpainting images) and measured the ensemble spread and the uncertainty–error correlation the same way for all three. FLOWER and OT-ODE give error-correlated uncertainty comparable to ours, so this is *not* a property unique to our solver; D-Flow returns a single optimized reconstruction rather than a conditional sampler, so it does not natively provide sample-based uncertainty. What separates our solver is cost. Runtime is wall-clock seconds per 256×256 image (AFHQ super-resolution, same GPU, each method at its reported step budget, ours with a single sample).

Method	cond. samples	spread	uncert.–error ρ	ODE-backprop-free	runtime (s/img)
FLOWER	✓ (8 draws)	0.009	0.74±0.04	✓	124.7
OT-ODE	✓ (8 draws)	0.029	0.76±0.03	✓	20.2
D-Flow	✗ (single MAP)	n/a	n/a	✗ (needs backprop)	503.7
Ours	✓ (8 draws)	0.028	0.75±0.04	✓	**6.6**

Table 4: **Sample-based uncertainty** (8 samples). Pearson ρ (95% CI, n images per row) between the per-pixel sample std and the per-pixel error of the sample mean, averaged over images. The first two rows share the 10 images of Table 2, so the D-Flow comparison is matched; D-Flow is a point estimate and has no uncertainty to report. Appendix C.4 gives the mask-controlled value, 0.44 ± 0.05 within the missing region.

Method	operator (n images)	sample-mean PSNR	diversity	uncertainty–error ρ
D-Flow (single MAP)	AFHQ inpainting (10)	24.32	0 (point est.)	n/a
Ours (8 samples)	AFHQ inpainting (10)	**27.64**	0.028	**0.75±0.04**
Ours (8 samples)	AFHQ inpainting (100)	28.15	0.028	**0.77±0.01**
Ours (8 samples)	CelebA inpainting (100)	32.65	0.019	**0.65±0.02**
Ours (8 samples)	AFHQ super-res (20)	27.59	0.057	**0.53±0.02**

we do ours gives error-correlated uncertainty indistinguishable from ours—$\rho = 0.74 \pm 0.04$ and 0.76 ± 0.03 against our 0.75 ± 0.04 on the same 10 images (Table 3)—differing only in spread, FLOWER's being $3\times$ tighter. What remains specific to our solver is cost: an 8-sample ensemble costs less than a *single* FLOWER reconstruction. D-Flow is the exception: it optimizes the source to a single point rather than sampling, so it does not natively provide an ensemble.

Sample-based uncertainty. The across-sample standard deviation gives a per-pixel uncertainty map, which Figure 2 shows concentrates where the error does. Its Pearson correlation with the per-pixel error is $\rho = 0.77 \pm 0.01$ (95% CI, 100 AFHQ inpainting images), 0.65 ± 0.02 on CelebA inpainting (100), and 0.53 ± 0.02 on super-resolution (20, where the error is more diffuse), with all reported as the mean per-image correlation with a 95% CI over images (Table 4). We avoid the word "calibrated": ρ says the uncertainty tracks error in aggregate, not that it is a credible interval, which would need a coverage-based evaluation. Nor is it ours alone—ensembled baselines reach the same ρ (Table 3)—though the point-estimate optimizer D-Flow furnishes no uncertainty at all.

Is the correlation just the mask? For inpainting both error and uncertainty concentrate in the missing region, so part of the global ρ only reflects that both know where the mask is. Controlling for this costs about half the effect but not the effect: over 40 images the correlation *restricted to the missing pixels* is 0.44 ± 0.05 and the partial correlation removing the mask indicator is 0.43 ± 0.05, against 0.76 ± 0.02 globally on those images—positive on *every* image tested (Appendix C.4). Super-resolution, with no mask to give the support away, is the complementary control and yields 0.53 ± 0.02 outright. Ensembling also helps quality: the sample mean over $1/4/8$ draws raises box-inpainting PSNR $26.87 \to 28.16 \to 29.24$ (Table 7), and on Table 2's images the 8-sample mean reaches 27.64 against D-Flow's 24.32, at far lower cost. More examples: Appendix D; Appendix C explains why the three box-inpainting numbers differ.

Ablation. The quality does not come merely from gradient normalization or step-size tuning: removing the Tweedie endpoint (guiding on $\mathbf{x}_t$) collapses LPIPS from 0.045 to 0.248, and removing per-image gradient normalization is catastrophic (0.239). Mobility and guidance strength are sec-

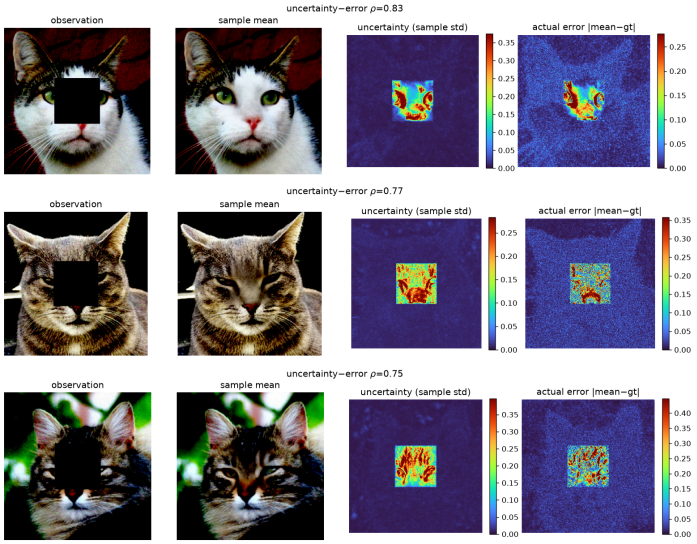


Figure 2: **The uncertainty map tracks the error** (box inpainting, 3 examples). Per row: degraded observation; sample mean; per-pixel uncertainty (across-sample std); actual error $|\text{mean} - \text{gt}|$ (both in the `turbo` colormap, observation/mean gamma-brightened for display). Uncertainty and error co-localize on the inpainted region (per-image ρ in titles).

ondary knobs, and the sample mean over $1/4/8$ draws improves PSNR $26.9 \to 28.2 \to 29.2$—the ensembling benefit single-shot solvers forgo. Full ablation: Table 7, Appendix C.

A second prior, and out-of-distribution priors. A second pretrained prior (CelebA, 128^2 faces) shows the same pattern—best or second-best LPIPS on super-resolution and random inpainting, $\rho = 0.65 \pm 0.02$ over 100 images—and applying the AFHQ-*cat* prior to AFHQ-*dog*/*wild* images (the only public flow priors) shows our solver *generalizes markedly better perceptually* than guidance (n=50 SR LPIPS: dog 0.183 vs. FLOWER's 0.260, wild 0.195 vs. 0.309), because guidance over-commits to the mismatched prior's modes. Tables and maps: Appendix C.

7.1 Scope, and the theory–practice gap

The concept, which the theory establishes. For deterministic flow priors, conditioning is exact source reweighting and carries no drift (Lemma 1, Propositions 1–2), and trajectory guidance is a greedy local corrector whose posterior gap is bounded by its distance to the canonical exact corrector $\mathbf{u}^\star$ (Theorem 1); the 2D study supports this against a closed-form posterior (Table 1).

The image solver, which is a separate heuristic. The exact estimator is intractable beyond low dimension (Table 5); our velocity-correction solver is an amortized heuristic *motivated by*, not derived from, the analysis, and its gradient-normalization ingredient is a numerical stabilizer the analysis does not predict. We do *not* claim it samples the posterior: its faithfulness is assessed only indirectly ($\rho = 0.53$–0.77 globally, ≈ 0.44 mask-controlled; ensembled baselines reach the same ρ; and Table 1 shows it inherits none of exact reweighting's mode-weight accuracy either). Closing that gap—a provably faithful *and* scalable source-side sampler—is the open problem this viewpoint makes precise.

8 Conclusion

For a deterministic flow prior, conditioning reweights the source, and trajectory-guidance solvers are local corrections within the minimum-kinetic-energy framework needed from the unconditional source. A closed-form study confirms source reweighting is exact while trajectory guidance is biased and mis-weights modes, a basis for evaluating inverse solvers beyond single-reconstruction metrics.

Reproducibility statement

All theoretical claims are stated with their assumptions in the main text or the proof appendices, and are proved in full in Appendix A (Lemma 1, Propositions 1 and 2, Theorem 1) and Appendix B (the large-deviation statements). The controlled 2D study is fully specified in Appendix E.2: the prior, the measurement operator, the closed-form posterior and the Monte-Carlo reference are all analytic, so it can be reproduced exactly from the description. Our velocity-correction solver is given as pseudocode in Algorithm 1 with every hyperparameter listed in Appendix E; it is built on the publicly released FLOWER benchmark code and the publicly released AFHQ-Cat and CelebA flow-matching priors, and all baselines are run through that same harness on the same images, prior and noise level. Dataset construction, including the out-of-distribution reuse of the AFHQ-cat prior, is described in Appendix C. Source code will be released.

References

Michael S Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. In *International Conference on Learning Representations*, 2023.

Ismail Alkhouri, Shijun Liang, Cheng-Han Huang, Jimmy Dai, Qing Qu, Saiprasad Ravishankar, and Rongrong Wang. Sitcom: Step-wise triple-consistent diffusion sampling for inverse problems. In *International Conference on Machine Learning*, pp. 1128–1158. PMLR, 2025.

Heli Ben-Hamu, Omri Puny, Itai Gat, Brian Karrer, Uriel Singer, and Yaron Lipman. D-flow: Differentiating through flows for controlled generation. In *International Conference on Machine Learning*, pp. 3462–3483. PMLR, 2024.

Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

Gabriel Cardoso, Yazid Janati el Idrissi, Sylvain Le Corff, and Eric Moulines. Monte carlo guided denoising diffusion models for bayesian linear inverse problems. In *International Conference on Learning Representations*, 2024.

Hyungjin Chung, Byeongsu Sim, Dohoon Ryu, and Jong Chul Ye. Improving diffusion models for inverse problems using manifold constraints. *Advances in Neural Information Processing Systems*, 35:25683–25696, 2022.

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.

Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion schrödinger bridge with applications to score-based generative modeling. *Advances in neural information processing systems*, 34:17695–17709, 2021.

Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer, 2nd edition, 2010.

Alexander Denker, Francisco Vargas, Shreyas Padhy, Kieran Didi, Simon Mathis, Vincent Dutordoir, Riccardo Barbano, Emile Mathieu, Urszula J Komorowska, and Pietro Lio. Deft: Efficient fine-tuning of diffusion models by learning the generalised h-transform. *Advances in Neural Information Processing Systems*, 37:19636–19682, 2024.

Paul Dupuis and Richard S Ellis. *A weak convergence approach to the theory of large deviations*. John Wiley & Sons, 1997.

Berthy T Feng, Jamie Smith, Michael Rubinstein, Huiwen Chang, Katherine L Bouman, and William T Freeman. Score-based diffusion models as principled priors for inverse imaging. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 10520–10531, 2023.

Ruiqi Feng, Chenglin Yu, Wenkang Deng, Peiyan Hu, and Tailin Wu. On the guidance of flow matching. In *International Conference on Machine Learning*, 2025.

Adhithyan Kalaivanan, Zheng Zhao, Jens Sjölund, and Fredrik Lindsten. ESS-flow: Training-free guidance of flow-based models as inference in source space. *arXiv preprint arXiv:2510.05849*, 2025.

Yuri Kifer. *Random Perturbations of Dynamical Systems*, volume 16 of *Progress in Probability and Statistics*. Birkhäuser Boston, 1988.

Jeongsol Kim, Bryan Sangwoo Kim, and Jong Chul Ye. Flowdps: Flow-driven posterior sampling for inverse problems. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 12328–12337, 2025.

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

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

S'egolène Martin, Anne Gagneux, Paul Hagemann, and Gabriele Steidl. Pnp-flow: Plug-and-play image restoration with flow matching. In *International Conference on Learning Representations*, 2025.

Badr Moufad, Yazid Janati el Idrissi, Lisa Bedin, Alain Oliviero Durmus, Randal Douc, Eric Moulines, and Jimmy Olsson. Variational diffusion posterior sampling with midpoint guidance. In *International Conference on Learning Representations*, 2025.

George Papamakarios, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *Journal of Machine Learning Research*, 22(57):1–64, 2021.

Mehrsa Pourya, Bassam El Rawas, and Michael Unser. Flower: A flow-matching solver for inverse problems. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=QGd34p02mI.

Sanjeev Raja, Martin Sipka, Michael Psenka, Tobias Kreiman, Michal Pavelka, and Aditi S Krishnapriyan. Action-minimization meets generative modeling: Efficient transition path sampling with the onsager-machlup functional. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 50972–51008, 2025.

Danilo Rezende and Shakir Mohamed. Variational inference with normalizing flows. In *International conference on machine learning*, pp. 1530–1538. PMLR, 2015.

Jiaming Song, Arash Vahdat, Morteza Mardani, and Jan Kautz. Pseudoinverse-guided diffusion models for inverse problems. In *International conference on learning representations*, 2023.

Belinda Tzen and Maxim Raginsky. Theoretical guarantees for sampling and inference in generative models with latent diffusions. In *Conference on Learning Theory*, pp. 3084–3114. PMLR, 2019.

Yinhuai Wang, Jiwen Yu, and Jian Zhang. Zero-shot image restoration using denoising diffusion null-space model. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=mRieQgMtNTQ.

Zifan Wang, Alice Harting, Matthieu Barreau, Michael M Zavlanos, and Karl H Johansson. Source-guided flow matching. *arXiv preprint arXiv:2508.14807*, 2025.

Luhuan Wu, Brian Trippe, Christian Naesseth, David Blei, and John P Cunningham. Practical and asymptotically exact conditional sampling in diffusion models. *Advances in Neural Information Processing Systems*, 36:31372–31403, 2023.

Jian Xu, Delu Zeng, John Paisley, and Qibin Zhao. Onsager-machlup posterior transport for deep gaussian processes. *arXiv preprint arXiv:2605.23434*, 2026.

Xingyu Xu and Yuejie Chi. Provably robust score-based diffusion posterior sampling for plug-and-play image reconstruction. *Advances in Neural Information Processing Systems*, 37:36148–36184, 2024.

Bingliang Zhang, Wenda Chu, Julius Berner, Chenlin Meng, Anima Anandkumar, and Yang Song. Improving diffusion inverse problem solving with decoupled noise annealing. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 20895–20905, 2025.

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

A PROOFS

Throughout, $\Phi_{s\to t}$ is the flow map of the probability-flow ODE $\dot{\mathbf{x}}_t = \boldsymbol{v}_t(\mathbf{x}_t)$, $\Phi := \Phi_{0\to 1}$, and $p_0 = \mathcal{N}(0, I)$. We make the standard regularity assumption.

Assumption 1 (Regularity)**.** *The velocity field $\boldsymbol{v}_t$, the canonical corrector $\mathbf{u}_t^\star$ and the approximate corrections $\mathbf{u}_t$ are continuous in t and L-Lipschitz in $\mathbf{x}$, uniformly in $t \in [0,1]$ (the Grönwall step of Theorem 1 is applied to the exact field $\boldsymbol{v}_t + \mathbf{u}_t^\star$); the flow map Φ is a C^1-diffeomorphism with $\det \nabla\Phi \neq 0$; and the likelihood $\mathbf{x} \mapsto p(\mathbf{y} \mid \mathbf{x})$ is bounded and measurable. These hold for the smooth velocity fields produced by trained flow models on bounded domains.*

A.1 PROOF OF LEMMA 1 AND PROPOSITION 1

Reweighting identity. Let $p_0^{\mathbf{y}}(\mathbf{z}) = p(\mathbf{y} \mid \Phi(\mathbf{z}))p_0(\mathbf{z})/Z$ with $Z = \int p(\mathbf{y} \mid \Phi(\mathbf{z}))p_0(\mathbf{z})\, d\mathbf{z}$. For any bounded test function g,

$$\begin{aligned}
\mathbb{E}_{\mathbf{z}\sim p_0^{\mathbf{y}}}\big[g(\Phi(\mathbf{z}))\big] &= \frac{1}{Z}\int g(\Phi(\mathbf{z}))\, p(\mathbf{y} \mid \Phi(\mathbf{z}))\, p_0(\mathbf{z})\, d\mathbf{z} \\
&= \frac{1}{Z}\int g(\mathbf{x})\, p(\mathbf{y} \mid \mathbf{x})\, p_0(\Phi^{-1}(\mathbf{x}))\, \big|\det \nabla\Phi^{-1}(\mathbf{x})\big|\, d\mathbf{x} \qquad (\mathbf{x} = \Phi(\mathbf{z})) \\
&= \frac{1}{Z}\int g(\mathbf{x})\, p(\mathbf{y} \mid \mathbf{x})\, p_1(\mathbf{x})\, d\mathbf{x} = \int g(\mathbf{x})\, p(\mathbf{x} \mid \mathbf{y})\, d\mathbf{x}, \qquad (6)
\end{aligned}$$

where the third equality uses the change-of-variables identity for the pushforward density, $p_1(\mathbf{x}) = p_0(\Phi^{-1}(\mathbf{x}))\,|\det \nabla\Phi^{-1}(\mathbf{x})|$, and the last uses $Z = \int p(\mathbf{y} \mid \mathbf{x})p_1(\mathbf{x})\, d\mathbf{x} = p(\mathbf{y})$ together with Bayes' rule $p(\mathbf{x} \mid \mathbf{y}) = p(\mathbf{y} \mid \mathbf{x})p_1(\mathbf{x})/p(\mathbf{y})$. As this holds for all bounded g, $\Phi_{\#}p_0^{\mathbf{y}} = p(\cdot \mid \mathbf{y})$.

Posterior path. Let μ_0 be any initial law and $\mu_t := (\Phi_{0\to t})_{\#}\mu_0$. Because $\{\Phi_{0\to t}\}_t$ is the flow of $\dot{\mathbf{x}} = \boldsymbol{v}_t(\mathbf{x})$, the curve $t \mapsto \mu_t$ satisfies the continuity (Liouville) equation $\partial_t\mu_t + \nabla\cdot(\mu_t \boldsymbol{v}_t) = 0$ in the weak sense: for $g \in C_c^\infty$, $\frac{d}{dt}\int g\, d\mu_t = \int \nabla g \cdot \boldsymbol{v}_t\, d\mu_t$ by differentiating $g(\Phi_{0\to t}(\mathbf{x}))$ and integrating against μ_0. Taking $\mu_0 = p_0^{\mathbf{y}}$ gives $p_t^{\mathbf{y}} = (\Phi_{0\to t})_{\#}p_0^{\mathbf{y}}$ with $\partial_t p_t^{\mathbf{y}} + \nabla\cdot(p_t^{\mathbf{y}}\boldsymbol{v}_t) = 0$ and terminal value $p_1^{\mathbf{y}} = \Phi_{\#}p_0^{\mathbf{y}} = p(\cdot \mid \mathbf{y})$. The velocity field is the prior $\boldsymbol{v}_t$, independent of $\mathbf{y}$; conditioning enters only through the initial law $p_0^{\mathbf{y}}$. □

A.2 PROOF OF PROPOSITION 2

Additional regularity. Beyond Assumption 1 this proof uses that the terminal likelihood $\mathbf{x} \mapsto p(\mathbf{y} \mid \mathbf{x})$ is C^1 and bounded away from zero on the relevant compact set, and that the reference SDE admits a differentiable stochastic flow (we use its first variation $J_{t\to 1}^\epsilon$ below). Assumption 1 alone requires only boundedness and measurability of the likelihood, so we state these extra conditions here rather than strengthening the global assumption.

Fix a bounded, C^1 terminal weight $g(\mathbf{x}_1) = p(\mathbf{y} \mid \mathbf{x}_1)$ with $g \geq c > 0$ on the relevant compact set (so $\log g$ is well defined and Lipschitz), and let Φ be a C^1 flow (Assumption 1). Consider the reference SDE with small noise $d\mathbf{x}_t = \boldsymbol{v}_t(\mathbf{x}_t)\, dt + \epsilon\, dW_t$ and its endpoint-conditioned (Doob) bridge, whose drift is $\boldsymbol{v}_t + \epsilon^2\nabla \log h_t^\epsilon$ with the harmonic function

$$h_t^\epsilon(\mathbf{x}) = \mathbb{E}\big[g(\mathbf{x}_1^\epsilon)\,\big|\,\mathbf{x}_t^\epsilon = \mathbf{x}\big], \qquad (7)$$

the expectation taken under the reference SDE. We show the correction vanishes pointwise.

Step 1 (the harmonic function converges to a point evaluation). Let $\mathbf{x}_1^\epsilon$ solve the reference SDE from $\mathbf{x}_t^\epsilon = \mathbf{x}$. By standard small-noise stability of SDEs (Grönwall on $\mathbb{E}\|\mathbf{x}_1^\epsilon - \Phi_{t\to 1}(\mathbf{x})\|^2 \le C\epsilon^2$ under Assumption 1), $\mathbf{x}_1^\epsilon \to \Phi_{t\to 1}(\mathbf{x})$ in L^2, hence in probability, as $\epsilon \to 0$. As g is bounded and continuous, dominated convergence gives $h_t^\epsilon(\mathbf{x}) \to g(\Phi_{t\to 1}(\mathbf{x})) =: h_t^0(\mathbf{x})$.

Step 2 (C^1 convergence and a uniform gradient bound). Differentiating equation 7 in $\mathbf{x}$ and using the first-variation (derivative) flow $J_{t\to 1}^\epsilon = \partial \mathbf{x}_1^\epsilon / \partial \mathbf{x}$, which converges to the deterministic Jacobian $\nabla \Phi_{t\to 1}(\mathbf{x})$ in L^2,

$$\nabla h_t^\epsilon(\mathbf{x}) = \mathbb{E}\big[(J_{t\to 1}^\epsilon)^\top \nabla g(\mathbf{x}_1^\epsilon)\big] \xrightarrow[\epsilon\to 0]{} \big(\nabla \Phi_{t\to 1}(\mathbf{x})\big)^\top \nabla g(\Phi_{t\to 1}(\mathbf{x})) = \nabla h_t^0(\mathbf{x}), \tag{8}$$

which is finite and bounded on compacts (product of bounded C^1 quantities). Since also $h_t^\epsilon \to h_t^0 \ge c > 0$, the logarithmic gradient is uniformly bounded for small ϵ: $\sup_{\epsilon\le\epsilon_0} \|\nabla \log h_t^\epsilon(\mathbf{x})\| = \sup_\epsilon \|\nabla h_t^\epsilon / h_t^\epsilon\| \le M(\mathbf{x}) < \infty$.

Step 3 (the drift vanishes; the source carries the conditioning). Therefore the Doob correction satisfies $\|\epsilon^2 \nabla \log h_t^\epsilon(\mathbf{x})\| \le \epsilon^2 M(\mathbf{x}) \to 0$ pointwise. At the same time the bridge's initial law is the reference initial law reweighted by h_0^ϵ, $p_0^{\mathbf{y},\epsilon} \propto h_0^\epsilon\, p_0 \to g(\Phi(\cdot))\, p_0 = p(\mathbf{y} \mid \Phi(\cdot))\, p_0 \propto p_0^{\mathbf{y}}$. Passing to the limit, the conditioned process is the deterministic flow Φ started from $p_0^{\mathbf{y}}$ with *zero* added drift, matching Proposition 1. □

Remark (what this does and does not rule out). The proposition shows that *one specific* way of building a corrector—copying the diffusion Doob h-transform drift (as in classifier/measurement guidance and Denker et al. (2024)) into the small-noise limit—vanishes and carries no information at $\epsilon = 0$. It does *not* show that no drift-based corrector can be exact: $\mathbf{u}^\star$ of equation 4 is one, by construction, and it is not obtained by this small-noise analogy (it solves the transport problem directly, at $\epsilon = 0$, rather than arising as a limit of Doob drifts). What the proposition explains, rather than merely observes, is why existing solvers—which build their correction by analogy to the diffusion h-transform rather than by solving equation 4—have no principled reason to land on $\mathbf{u}^\star$ (Section 5); it is not a structural impossibility result for drift-based correction in general.

Remark (the tempered-likelihood regime). Step 2's $O(1)$ gradient bound uses that g is fixed in ϵ. If instead the likelihood is tempered at the noise scale, $g^\epsilon = \exp(-\ell_{\mathbf{y}} \circ \Phi_{t\to 1}/\epsilon^2)$ (the large-deviation scaling of Appendix B), then $\log h_t^\epsilon \asymp -S_t/\epsilon^2$ for an action S_t and $\epsilon^2 \nabla \log h_t^\epsilon \to -\nabla S_t \neq 0$: the correction converges to the Freidlin–Wentzell minimum-action drift rather than to zero. The two regimes are consistent—both reach $p(\cdot \mid \mathbf{y})$— and correspond exactly to the Section 4 (source reweighting, no drift) vs. Section 5 (start from p_0, pay an action) dichotomy.

A.3 Exact reweighting of the marginal path and proof of Proposition 3

We first record the deterministic collapse of the h-function that distinguishes the flow case from the diffusion case.

Proposition 4 (Deterministic reweighting)**.** *For every t, the posterior marginal equals the prior marginal reweighted by a* pointwise *likelihood,*

$$p_t^{\mathbf{y}}(\mathbf{x}) = p_t(\mathbf{x})\, \frac{p\big(\mathbf{y} \mid \Phi_{t\to 1}(\mathbf{x})\big)}{p(\mathbf{y})}. \tag{9}$$

Proof. Under the prior flow, the terminal state is the deterministic function $\mathbf{x}_1 = \Phi_{t\to 1}(\mathbf{x}_t)$ of $\mathbf{x}_t$. Hence conditioning the endpoint on $\mathbf{y}$ reweights the law of $\mathbf{x}_t$ by $p(\mathbf{y} \mid \mathbf{x}_1 = \Phi_{t\to 1}(\mathbf{x}_t))$: $p_t^{\mathbf{y}}(\mathbf{x}) \propto p_t(\mathbf{x})\, p(\mathbf{y} \mid \Phi_{t\to 1}(\mathbf{x}))$, and normalizing by $p(\mathbf{y})$ gives equation 9. The conditional expectation defining the diffusion h-function, $h_t(\mathbf{x}) = \mathbb{E}[p(\mathbf{y} \mid \mathbf{x}_1) \mid \mathbf{x}_t = \mathbf{x}]$, collapses to a point evaluation because $\mathbf{x}_1 \mid \mathbf{x}_t$ is a Dirac mass. □

The identity in equation 9 defines the *instantaneous likelihood score*

$$\mathbf{s}_t(\mathbf{x}) := \nabla_{\mathbf{x}} \log p\big(\mathbf{y} \mid \Phi_{t\to 1}(\mathbf{x})\big) = \big(\nabla_{\mathbf{x}} \Phi_{t\to 1}(\mathbf{x})\big)^\top \nabla_{\mathbf{x}_1} \log p(\mathbf{y} \mid \mathbf{x}_1)\big|_{\mathbf{x}_1 = \Phi_{t\to 1}(\mathbf{x})}. \tag{10}$$

The canonical correction. To reach $p(\cdot \mid \mathbf{y})$ when starting from the *unconditional* p_0 (rather than $p_0^{\mathbf{y}}$), one integrates $\dot{\mathbf{x}} = \boldsymbol{v}_t + \mathbf{u}_t$ with marginals $q_t^{\mathbf{u}}$, $q_0 = p_0$, and requires $q_1^{\mathbf{u}} = p(\cdot \mid \mathbf{y})$. Among all admissible $\mathbf{u}$, the minimum-kinetic-energy field in equation 4 exists and is a gradient field $\mathbf{u}_t^\star = \nabla\phi_t$, where $(q_t^\star, \phi_t)$ solve the Benamou–Brenier optimality system relative to the reference drift $\boldsymbol{v}_t$,

$$\partial_t q_t^\star + \nabla\cdot\big(q_t^\star(\boldsymbol{v}_t + \nabla\phi_t)\big) = 0, \qquad \partial_t\phi_t + \boldsymbol{v}_t\cdot\nabla\phi_t + \tfrac{1}{2}\|\nabla\phi_t\|^2 = 0, \tag{11}$$

with boundary data $q_0^\star = p_0$, $q_1^\star = p(\cdot \mid \mathbf{y})$ (Benamou & Brenier, 2000). This is the deterministic, mobility-Euclidean analogue of the Doob/Föllmer drift; the existence and gradient form follow from the dynamic-OT formulation under Assumption 1.

Existing solvers as approximations. Trajectory-guidance solvers use the *greedy* correction aligned with the instantaneous likelihood score equation 10, with two approximations:

1. **Endpoint map.** They replace the exact endpoint map $\Phi_{t\to 1}(\mathbf{x})$ by its one-step (Tweedie) linearization $\hat{\mathbf{x}}_1(\mathbf{x}, t) = \mathbf{x} + (1-t)\boldsymbol{v}_t(\mathbf{x})$, i.e. $\hat{\mathbf{x}}_1 = \mathbf{x} + (1-t)\boldsymbol{v}_t(\mathbf{x}) = \Phi_{t\to 1}(\mathbf{x}) + O\big((1-t)^2\big)$, so that $\mathbf{s}_t(\mathbf{x}) \approx \nabla_{\mathbf{x}} \log p(\mathbf{y} \mid \hat{\mathbf{x}}_1(\mathbf{x}, t))$.
2. **Mobility and greediness.** They scale this score by a mobility M_t and follow it greedily rather than solving equation 11: *FlowDPS* takes $M_t = \zeta_t I$ scalar (point-mass endpoint); *FLOWER* models $\mathbf{x}_1 \mid \mathbf{x}_t \approx \mathcal{N}(\hat{\mathbf{x}}_1, \sigma_{r,t}^2 I)$ and sets M_t to the induced Gaussian posterior covariance, solved by conjugate gradients (ΠGDM-type); *PnP-Flow* replaces the gradient step by a data-fidelity proximal map and a flow-path reprojection (operator splitting).

Both approximations are exact only in degenerate cases (linear $\Phi_{t\to 1}$; Gaussian interpolation conditional; vanishing transport cost), so in general $\mathbf{u}_t \neq \mathbf{u}_t^\star$ and a posterior bias remains, quantified next. □

A.4 Proof of Theorem 1

Let $w_t = \boldsymbol{v}_t + \mathbf{u}_t^\star$ and $w_t' = \boldsymbol{v}_t + \mathbf{u}_t$ be the exact and approximate total velocities, generating flows $\Psi_{0\to t}$ and $\Psi'_{0\to t}$ from the shared initial law p_0. Couple trajectories by the same initial point $\mathbf{x}_0 \sim p_0$ and write $\mathbf{x}_t = \Psi_{0\to t}(\mathbf{x}_0)$, $\mathbf{x}_t' = \Psi'_{0\to t}(\mathbf{x}_0)$. Then

$$\tfrac{d}{dt}\|\mathbf{x}_t - \mathbf{x}_t'\| \le \|w_t(\mathbf{x}_t) - w_t(\mathbf{x}_t')\| + \|w_t(\mathbf{x}_t') - w_t'(\mathbf{x}_t')\| \le L\|\mathbf{x}_t - \mathbf{x}_t'\| + \|\mathbf{u}_t^\star(\mathbf{x}_t') - \mathbf{u}_t(\mathbf{x}_t')\|, \tag{12}$$

using Assumption 1. Grönwall's inequality with $\mathbf{x}_0 = \mathbf{x}_0'$ gives $\|\mathbf{x}_1 - \mathbf{x}_1'\| \le \int_0^1 e^{L(1-s)}\|\mathbf{u}_s^\star(\mathbf{x}_s') - \mathbf{u}_s(\mathbf{x}_s')\|\, ds$. Taking the $L^2(p_0)$ norm, using the coupling as an admissible plan for W_2 and Minkowski's integral inequality,

$$W_2\big(\Psi_{0\to 1\,\#}p_0,\ \Psi'_{0\to 1\,\#}p_0\big) \le e^L \int_0^1 \big(\mathbb{E}_{q_s^{\mathbf{u}}}\|\mathbf{u}_s - \mathbf{u}_s^\star\|^2\big)^{1/2}\, ds. \tag{13}$$

Since $\Psi_{0\to 1\,\#}p_0 = p(\cdot \mid \mathbf{y})$ exactly and $\Psi'_{0\to 1\,\#}p_0 = p_1^{\mathbf{u}}$ in continuous time with exact $\boldsymbol{v}_t$, adding the ODE-discretization error ϵ_{disc} (from finite step size) and the flow-model error ϵ_{flow} (from $\hat{\boldsymbol{v}}_t \neq \boldsymbol{v}_t$) by the triangle inequality yields equation 5 with $L \leftarrow e^L$. □

Remark (a one-directional bound, not an algorithmic error analysis). Equation equation 5 only runs in one direction: a small correction–$\mathbf{u}^\star$ gap implies small posterior bias, but the converse is false. First, since $\mathbf{u}^\star$ is merely the minimum-energy member of a generally larger family of exact correctors (any $\mathbf{u}$ with $q_1^{\mathbf{u}} = p(\cdot \mid \mathbf{y})$ qualifies), a solver could be far from $\mathbf{u}^\star$ specifically while landing near a different exact corrector with small true bias; a large gap to $\mathbf{u}^\star$ does not certify a large bias. Second, the bound supplies no rate at which FlowDPS, FLOWER, or PnP-Flow approach $\mathbf{u}^\star$ as a function of, e.g., problem dimension or measurement operator—Proposition 3 places them as local approximations within the same template, but proves no convergence statement—and $\mathbf{u}^\star$ itself is generally intractable to evaluate, so the gap cannot be measured directly for these solvers—indeed we never solve for $\mathbf{u}^\star$ or evaluate $\|\mathbf{u}_t - \mathbf{u}_t^\star\|$ anywhere in this paper, including the toy study. We therefore treat equation 5 as identifying the correction–$\mathbf{u}^\star$ gap as the relevant quantity, not as a predictive error analysis of any existing algorithm; the toy study (Section 6) instead measures the bound's consequence—the resulting terminal-distribution error against the closed-form posterior, which *is* available exactly there—as an empirical proxy for that gap's cost, not the gap itself.

Remark (non-vanishing floor). If the correction has a mis-specified direction or mobility—e.g. a scalar M_t where $\mathbf{u}^\star$ requires an anisotropic one, or a greedy field where transport requires a global rerouting—then $\inf_{c>0} \mathbb{E}\|c\,\mathbf{u}_t - \mathbf{u}_t^\star\|^2 > 0$, so no scalar rescaling of the guidance drives the *bound* to zero. We stress that equation 5 is an upper bound: this says the guarantee cannot be recovered by tuning c, not that the true error is bounded below. It is consistent with, but does not by itself predict, the plateau observed in Table 1. The corresponding mode-weight characterization is Proposition 8, which gives an *exact* asymptotic TV mismatch for the idealized caricature—itself zero in the exceptional case $\alpha = w$—and so is a characterization rather than a universal positive lower bound.

A.5 Proof of Proposition 5

The proposition is a Regime-A statement (Appendix B.2): the prior $p_0 = \mathcal{N}(0, I)$ is held *fixed* and only the observation noise vanishes, so it is proved by Laplace's method with a smooth amplitude, not by a large-deviation principle for a tempered Gibbs measure. Write $g := \ell_{\mathbf{y}} \circ \Phi$, so that by equation 17

$$p_0^{\mathbf{y},\epsilon}(\mathbf{z}) = \frac{1}{Z_\epsilon}\, e^{-g(\mathbf{z})/\epsilon}\, p_0(\mathbf{z}), \qquad Z_\epsilon = \int e^{-g(\mathbf{z})/\epsilon} p_0(\mathbf{z})\, d\mathbf{z}, \tag{14}$$

a Laplace-type integral in which the exponent is g alone and the fixed prior p_0 appears as a smooth, strictly positive *amplitude*.

Assume as stated that $g \in C^2$ has finitely many minimizers $\{\mathbf{z}_k^\star\}_{k=1}^K$, each isolated with $\nabla^2 g(\mathbf{z}_k^\star) =: H_k \succ 0$, and normalize $\min g = 0$. Assume in addition the standard tail condition that for every $\delta > 0$ the minimizers are globally separated from the tails, $\eta_\delta := \inf\{g(\mathbf{z}) : \mathbf{z} \notin \bigcup_k B_\delta(\mathbf{z}_k^\star)\} > 0$, and that $\int e^{-g/\epsilon_0} p_0 < \infty$ for some $\epsilon_0 > 0$. These are two separate assumptions: integrability of p_0 gives the second but *not* the first, since $g \geq 0$ with isolated minimizers is still compatible with $g(\mathbf{z}_n) \to 0$ along a sequence escaping to infinity. We assume the separation directly; it holds, for instance, whenever g is coercive or the measurement-consistent set is bounded.

Fix $\delta > 0$. Outside $\bigcup_k B_\delta(\mathbf{z}_k^\star)$ the integrand of equation 14 is bounded by $e^{-\eta_\delta(1/\epsilon - 1/\epsilon_0)}\, e^{-g/\epsilon_0} p_0$, so that part of the integral is $O(e^{-\eta_\delta/2\epsilon})$ and exponentially negligible. On $B_\delta(\mathbf{z}_k^\star)$ substitute $\mathbf{z} = \mathbf{z}_k^\star + \sqrt{\epsilon}\,\mathbf{u}$ and use $g(\mathbf{z}) = \frac{1}{2}(\mathbf{z} - \mathbf{z}_k^\star)^\top H_k (\mathbf{z} - \mathbf{z}_k^\star) + o(\|\mathbf{z} - \mathbf{z}_k^\star\|^2)$ together with continuity of p_0 at $\mathbf{z}_k^\star$; dominated convergence gives the classical Laplace expansion (Dembo & Zeitouni, 2010, Ch. 3)

$$\int_{B_\delta(\mathbf{z}_k^\star)} e^{-g/\epsilon} p_0 = (2\pi\epsilon)^{d/2}\, |H_k|^{-1/2}\, p_0(\mathbf{z}_k^\star)\, \big(1 + o(1)\big), \qquad \epsilon \to 0. \tag{15}$$

Summing over k gives Z_ϵ to leading order, and hence

$$p_0^{\mathbf{y},\epsilon}\big(B_\delta(\mathbf{z}_k^\star)\big) \;\longrightarrow\; w_k = \frac{p_0(\mathbf{z}_k^\star)\, |H_k|^{-1/2}}{\sum_j p_0(\mathbf{z}_j^\star)\, |H_j|^{-1/2}}, \tag{16}$$

which is the stated weight: the prior enters through the amplitude prefactor $p_0(\mathbf{z}_k^\star)$ and *not* through the exponent. Since $\delta > 0$ was arbitrary and the complement carries vanishing mass, $p_0^{\mathbf{y},\epsilon}$ converges weakly to $\sum_k w_k \delta_{\mathbf{z}_k^\star}$. Finally Φ is continuous, so the continuous mapping theorem gives $\Phi_{\#} p_0^{\mathbf{y},\epsilon} \Rightarrow \sum_k w_k \delta_{\Phi(\mathbf{z}_k^\star)}$ with the same weights. □

Remark (what $\hat{\mathbf{x}}_1$ approximates). It is tempting to describe the Gaussian step of FLOWER and the proximal step of PnP-Flow as approximating an "endpoint posterior" $p(\mathbf{x}_1 \mid \mathbf{x}_t)$, and we deliberately do not. Under the deterministic prior ODE the endpoint is a *deterministic* function of the current state, $\mathbf{x}_1 = \Phi_{t\to 1}(\mathbf{x}_t)$, so that conditional law is a Dirac mass (Proposition 1) and has no higher-order structure to neglect. The object those solvers actually smooth is the endpoint conditional of the *flow-matching interpolation* used to train the velocity field—the law of $\mathbf{x}_1$ given a point on the stochastic interpolant $\mathbf{x}_t = (1-t)\mathbf{x}_0 + t\mathbf{x}_1$ with independently sampled $\mathbf{x}_0 \sim p_0, \mathbf{x}_1 \sim p_1$, which is genuinely non-degenerate (at t=0, e.g., it is simply p_1 itself, since $\mathbf{x}_0$ carries no information about the independently drawn $\mathbf{x}_1$). This is a *different* object from the ODE endpoint, not an approximation of it that improves with training or integration accuracy: the marginal velocity field $\boldsymbol{v}_t(\mathbf{x}) = \mathbb{E}[\mathbf{x}_1 - \mathbf{x}_0 \mid \mathbf{x}_t = \mathbf{x}]$ used to define the deterministic flow is already a conditional *expectation* over this same coupling, so no amount of learning or integration precision makes the

(degenerate, Dirac) ODE endpoint and the (non-degenerate) interpolation conditional coincide—they answer different questions at different points in the construction. FLOWER and PnP-Flow smooth the latter as a computationally convenient heuristic; we do not claim this is justified by, or convergent to, the former. In Proposition 3 we therefore phrase the common template as a *surrogate for the exact endpoint map* $\Phi_{t\to 1}$ and the global transport problem in equation 4, which is what the bound of Theorem 1 controls.

Remark. The tempered-prior (Regime B) statements of Appendix B are proved differently—there the prior does enter the exponent, and Varadhan's lemma and the contraction principle apply to the resulting Gibbs family. We keep the two arguments separate because they establish different limits.

B Extended large-deviation analysis

B.1 A large-deviation view of source reweighting and guidance collapse

The source representation in equation 2 has a clean small-noise asymptotic geometry that sharpens *why* source reweighting is exact yet hard, and why guidance collapses. We take the *small-observation-noise* regime, in which only the measurement noise shrinks, $\mathbf{y} = A(\mathbf{x}) + \sqrt{\epsilon}\,\boldsymbol{\eta}$, $\boldsymbol{\eta} \sim \mathcal{N}(0, I)$, and the *source prior is kept fixed* ($p_0 = \mathcal{N}(0, I)$, matching the actual model—we do *not* temper the prior). Writing $\ell_{\mathbf{y}}(\mathbf{x}) = \frac{1}{2}\|A(\mathbf{x}) - \mathbf{y}\|^2$, the source posterior is the tilted measure

$$p_0^{\mathbf{y},\epsilon}(\mathbf{z}) \;\propto\; \exp\big(-\ell_{\mathbf{y}}(\Phi(\mathbf{z}))/\epsilon\big)\, p_0(\mathbf{z}). \tag{17}$$

Proposition 5 (Small-noise concentration and Laplace weights)**.** *Let $\ell_{\mathbf{y}} \circ \Phi$ be C^2 with isolated nondegenerate minimizers $\{\mathbf{z}_k^\star\}$ on $\operatorname{supp} p_0$ (the measurement-consistent set). Then as $\epsilon \to 0$, $p_0^{\mathbf{y},\epsilon}$ concentrates on $\{\mathbf{z}_k^\star\}$ and, by Laplace's method, converges weakly to the categorical mixture $\sum_k w_k\, \delta_{\mathbf{z}_k^\star}$ with basin weights*

$$w_k \;\propto\; p_0(\mathbf{z}_k^\star)\, \big|\nabla^2(\ell_{\mathbf{y}} \circ \Phi)(\mathbf{z}_k^\star)\big|^{-1/2}, \tag{18}$$

the prior entering through the prefactor $p_0(\mathbf{z}_k^\star)$. Since Φ is a diffeomorphism, the data posterior $p_1^{\mathbf{y},\epsilon} = \Phi_{\#} p_0^{\mathbf{y},\epsilon}$ concentrates on $\{\Phi(\mathbf{z}_k^\star)\}$ with the same weights.

Regime and caveat. Proposition 5 is an *asymptotic, explanatory* statement about the small-observation-noise limit with a fixed prior; it is not a claim about the fixed-finite-σ posterior, and the basin picture is exact only as $\epsilon \to 0$. We use it to explain—not to certify—the finite-noise phenomena observed empirically (ESS collapse, mode mis-weighting). The path-space/Freidlin–Wentzell statements below additionally use a small *dynamics* noise (a reference SDE); these are likewise asymptotic.

The isolated-nondegenerate-minimizer hypothesis on $\ell_{\mathbf{y}} \circ \Phi$ deserves a further caveat, and we state it without the loophole we previously (incorrectly) offered. For a linear operator A with $\operatorname{rank}(A) < d$—the *under*determined case that includes our super-resolution and inpainting experiments, *and the toy study's own measurement* ($A \in \mathbb{R}^{1\times 2}$, $\operatorname{rank}(A) = 1 < d = 2$, Appendix E.2)—the zero level set $\{\mathbf{z} : A\Phi(\mathbf{z}) = \mathbf{y}\}$ is a $(d - \operatorname{rank}(A))$-dimensional submanifold, not a set of isolated points, and $\nabla^2(\ell_{\mathbf{y}} \circ \Phi) = J_\Phi^\top A^\top A J_\Phi$ has rank at most $\operatorname{rank}(A) < d$ at *every* point of it, for *any* prior: a multimodal prior changes how mass is distributed along this submanifold, but cannot turn points of it into isolated nondegenerate critical points of $\ell_{\mathbf{y}} \circ \Phi$ in the ambient space. Consequently Proposition 5's hypothesis is not satisfied by the rank-deficient toy operator used to illustrate it either, and we do not have a rigorous small-noise account of that toy study's basin structure; a correct treatment would analyze the fixed-prior limit fiber-wise on the measurement-consistent submanifold (a coarea-formula argument), which we have not carried out. What is unaffected by this gap is the toy study's *empirical* evidence (Table 1): the mode-weight and distributional metrics there are computed directly from the closed-form posterior and the simulated solvers, not derived from Proposition 5. The proposition's role is to explain, heuristically, why mode mis-weighting of the kind observed should be expected in the well- or over-determined case ($\operatorname{rank}(A) = d$, e.g. denoising); for the underdetermined case, including our own toy and image experiments, its explanatory content is an open problem rather than an established one.

Proposition 5 (proof in Appendix A) makes three points precise. **(i) Why source reweighting is exact but hard.** The exact estimator weights p_0 by $e^{-\ell_{\mathbf{y}}(\Phi(\mathbf{z}))/\epsilon}$, but the rate function localizes

mass in $O(\sqrt{\epsilon})$ neighborhoods of $\{\mathbf{z}_k^\star\}$; unconditional draws $\mathbf{z} \sim p_0$ rarely hit them, so the effective sample size collapses—a large-deviation account of the high-dimensional ESS failure noted above. **(ii) Why guidance mis-weights modes.** The true posterior is a Laplace *mixture* over basins with weights w_k. Source reweighting preserves these inter-basin weights, whereas a likelihood-gradient correction reweights them incorrectly because it transports mass by local ascent rather than by the global optimal coupling. We make this precise only for an *idealized caricature*—an autonomous source-space likelihood descent, whose terminal weights are those of the regions of attraction, giving a guidance-strength-independent $\mathrm{TV} = \frac{1}{2}\sum_k |\alpha_k - w_k|$ (Proposition 8 below). Real solvers integrate $\boldsymbol{v}_t + \mathbf{u}_t$, and the prior velocity $\boldsymbol{v}_t$ itself spreads mass across modes, so their failure is *softer*: mode mis-weighting rather than exact single-basin collapse. The idealized bound is therefore an upper caricature of the effect; what we verify empirically is the milder but real mis-weighting (Table 1: large mode-weight error and a non-vanishing distributional gap that no guidance strength removes) and, complementarily, the exponential ESS decay of exact source IS (Proposition 6 and Table 5 below). **(iii) What $\mathbf{u}^\star$ is.** The minimum-kinetic-energy correction in equation 4 is the deterministic-limit minimum-action (Freidlin–Wentzell) control (Kifer, 1988; Dembo & Zeitouni, 2010) steering the *unconditional* source to the posterior; from the reweighted source $p_0^{\mathbf{y}}$ no action is needed (Proposition 1). This situates our transport picture within the Schrödinger-bridge / Föllmer-drift continuum (Appendix F).

B.2 Two asymptotic regimes, kept separate

The statements below develop the picture of Appendix B.1 in more detail. Two different small-parameter scalings appear in this appendix and we keep them strictly separate; no statement is transferred between them.

Regime A (fixed prior, vanishing observation noise). Only the measurement noise shrinks, $\mathbf{y} = A(\mathbf{x}) + \sqrt{\epsilon}\,\boldsymbol{\eta}$, and the source prior is held fixed at $p_0 = \mathcal{N}(0, I)$, matching the actual model. The source posterior is equation 17 and the prior enters the Laplace weights of Proposition 5 through the *prefactor* $p_0(\mathbf{z}_k^\star)$. This is the regime used in the main text and in Appendix B.1.

Regime B (jointly tempered prior and likelihood). This regime is used *only* for the exponential-rate ESS calculation of Proposition 6; every other statement in this appendix, including the basin statements (Propositions 7 and 8), is a Regime-A statement with the fixed prior $p_0 = \mathcal{N}(0, I)$. For the rate computation it is convenient to temper the prior at the same rate, $p_0^\epsilon \propto e^{-I_0/\epsilon}$ with $I_0 = -\log p_0$, so that the source posterior is a single Gibbs measure $p_0^{\mathbf{y},\epsilon} \propto e^{-I_{\mathbf{y}}/\epsilon}$ with $I_{\mathbf{y}}(\mathbf{z}) = I_0(\mathbf{z}) + \ell_{\mathbf{y}}(\Phi(\mathbf{z})) - \min I_{\mathbf{y}}$, $\ell_{\mathbf{y}}(\mathbf{x}) = \frac{1}{2}\|A(\mathbf{x}) - \mathbf{y}\|^2$, and minimizers $\{\mathbf{z}_k^\star\}_{k=1}^K$ with Hessians $H_k = \nabla^2 I_{\mathbf{y}}(\mathbf{z}_k^\star)$. Putting prior and likelihood on the same exponential scale is what makes that rate function well defined; the price is that the prior sits in the *exponent* rather than the prefactor, so Regime B is genuinely a different limit from Regime A—which is precisely why we confine it to Proposition 6. Both are asymptotic devices used to *explain* the finite-σ behaviour we measure, not finite-σ guarantees.

B.3 Why importance sampling collapses: exponential ESS decay

We work in Regime B (Appendix B.2), so the proposal is the *tempered* source $p_0^\epsilon \propto e^{-I_0/\epsilon}$, not the fixed $\mathcal{N}(0, I)$ of Regime A. Consider self-normalized importance sampling of $p_0^{\mathbf{y},\epsilon}$ with proposal p_0^ϵ and weights $w(\mathbf{z}) = p(\mathbf{y} \mid \Phi(\mathbf{z})) = e^{-\ell_{\mathbf{y}}(\Phi(\mathbf{z}))/\epsilon}$. The standard normalized effective sample size is $\mathrm{ESS}/N = \big(\mathbb{E}_{p_0^\epsilon} w\big)^2 / \mathbb{E}_{p_0^\epsilon} w^2 = 1/\big(1 + \chi^2(p_0^{\mathbf{y},\epsilon} \,\|\, p_0^\epsilon)\big)$.

Proposition 6 (Exponential ESS decay)**.** *In Regime B (jointly tempered prior and likelihood, Appendix B.2), $\mathbb{E}_{p_0^\epsilon}[w^m] \asymp e^{-\frac{1}{\epsilon}\inf_{\mathbf{z}}[\,I_0(\mathbf{z}) + m\,\ell_{\mathbf{y}}(\Phi(\mathbf{z}))\,]}$ by Laplace's method—the I_0/ϵ in the exponent is the tempered proposal itself—so*

$$\epsilon \log \frac{\mathrm{ESS}}{N} \;\xrightarrow[\epsilon\to 0]{}\; -\Big(\,2\,m_1 - m_2 - m_0\,\Big) \;=:\; -\Delta \le 0, \tag{19}$$

where $m_j = \inf_{\mathbf{z}}[I_0(\mathbf{z}) + j\,\ell_{\mathbf{y}}(\Phi(\mathbf{z}))]$ for $j \in \{0, 1, 2\}$. For each fixed $\mathbf{z}$ the map $j \mapsto I_0(\mathbf{z}) + j\,\ell_{\mathbf{y}}(\Phi(\mathbf{z}))$ is affine, so m_j, being a pointwise infimum of affine functions, is concave *in j; hence $m_1 \ge \frac{1}{2}(m_0 + m_2)$ and*

$$\Delta = 2m_1 - m_0 - m_2 \;\ge\; 0 \quad \textit{always.} \tag{20}$$

Strict decay, $\Delta > 0$, holds exactly when this concavity is strict *at j=1—equivalently, when the minimizing configuration moves with j, so that no single $\mathbf{z}$ attains the infimum simultaneously at $j = 0, 1, 2$. This is an assumption on the joint geometry of I_0 and $\ell_{\mathbf{y}} \circ \Phi$ and does* not *follow merely from $\ell_{\mathbf{y}} \circ \Phi$ being non-constant: if one $\mathbf{z}$ happened to be optimal at all three exponents, m_j would be locally affine and $\Delta = 0$. When $\Delta > 0$, ESS/N decays exponentially in $1/\epsilon$—and, in product or high-dimensional settings where log-weight contributions accumulate across coordinates, exponentially in the dimension. This is the large-deviation form of the statement that naive source reweighting is exact but intractable in high dimension.*

We illustrate the same ESS-collapse phenomenon in the actual fixed-prior setting (Table 5): for a Gaussian flow prior with a random linear measurement observing $d/4$ coordinates at fixed noise σ=0.3, the importance-sampling ESS/N collapses from 0.97 at d=2 to the floor $\approx 1/N$ by d=128, and the IS posterior-mean estimate degrades from 3% relative error to *worse than the prior mean*. Exact source reweighting is thus usable only at very low dimension—hence the need for an amortized solver at image scale.

Table 5: **Exact source importance sampling collapses with dimension** (Gaussian flow, observe $d/4$ coords, σ=0.3, N=4×10^4 proposals). $\mathrm{ESS}/N \to 1/N$ and the IS posterior-mean error exceeds 1 (worse than the prior).

dimension d	2	8	32	128	512	2048
ESS/N	0.97	3.2×10^{-3}	3.2×10^{-4}	2.5×10^{-5}	2.5×10^{-5}	2.5×10^{-5}
rel. error of IS mean	0.03	0.06	0.83	1.89	1.97	2.23

B.4 Source reweighting preserves Laplace weights; greedy guidance does not

Proposition 7 (Basin-weight preservation)**.** *In Regime A (fixed prior p_0, vanishing observation noise), as $\epsilon \to 0$ both the true posterior and exact source reweighting (importance resampling) converge to the same categorical law over basins, $\sum_k w_k \, \delta_{\mathbf{x}_k^\star}$ with $\mathbf{x}_k^\star = \Phi(\mathbf{z}_k^\star)$ and the Laplace weights of Proposition 5,*

$$w_k \;\propto\; p_0(\mathbf{z}_k^\star)\,\big|\nabla^2(\ell_{\mathbf{y}} \circ \Phi)(\mathbf{z}_k^\star)\big|^{-1/2}. \tag{21}$$

Source reweighting therefore preserves the relative basin (mode) weights by construction.

Proposition 8 (Greedy guidance replaces Laplace weights by attraction-region weights)**.** *Consider the idealized caricature of pure likelihood ascent: the* autonomous *source-space gradient flow*

$$\frac{d\mathbf{z}_s}{ds} = -\,c\,\nabla_{\mathbf{z}}\big[\ell_{\mathbf{y}}\big(\Phi(\mathbf{z}_s)\big)\big], \qquad \mathbf{z}_0 \sim p_0, \quad c > 0, \tag{22}$$

run to convergence as $s \to \infty$. This is an autonomous source-space analogue *of pure likelihood ascent—a caricature motivated by the same endpoint likelihood objective, not a change of variables of $\nabla_{\mathbf{x}} \log r_t$. (Under the nonlinear map $\mathbf{x} = \Phi_{0\to t}(\mathbf{z})$ a Euclidean gradient does not transform as a vector field: $\nabla_{\mathbf{z}}(\ell_{\mathbf{y}} \circ \Phi) = J_\Phi^\top \nabla_{\mathbf{x}} \ell_{\mathbf{y}}$, whereas pulling back an $\mathbf{x}$-space field would involve J_Φ^{-1}, and the two agree only under an isometry. We therefore define the caricature directly in source space, where it is autonomous and the basin argument below is clean.) Each trajectory converges to the minimizer $\mathbf{z}^\star_{k(\mathbf{z}_0)}$ of $\ell_{\mathbf{y}} \circ \Phi$ whose region of attraction $\mathcal{R}_k \subset \mathbb{R}^d$ contains $\mathbf{z}_0$. Its terminal law is therefore the categorical measure $\sum_k \alpha_k \, \delta_{\mathbf{x}_k^\star}$ with* attraction-region *weights $\alpha_k = p_0(\mathcal{R}_k)$, and for any guidance strength $c > 0$,*

$$\mathrm{TV}\big(\textit{guidance},\, p(\cdot \mid \mathbf{y})\big) \;=\; \tfrac{1}{2}\sum_k \big|\alpha_k - w_k\big|, \tag{23}$$

independently of c. *The α_k are fixed by the geometry of the regions of attraction of $\ell_{\mathbf{y}} \circ \Phi$ under p_0, whereas the w_k are the Laplace weights of Proposition 7; the two are computed from different objects and there is no mechanism forcing them to agree. In the special case where the initializations that carry p_0-mass all fall in one region of attraction ($\alpha_{k^\star} = 1$) this reduces to* $\mathrm{TV} = 1 - w_{k^\star} \geq 1 - \max_k w_k$.

Proof. Resampling from an exact reweighting reproduces the limiting mixture, giving Proposition 7. For Proposition 8, the deterministic descent map $\mathbf{z} \mapsto \mathbf{x}^\star_{k(\mathbf{z})}$ sends each initialization to one basin

minimizer, and the regions of attraction $\{\mathcal{R}_k\}$ partition $\operatorname{supp} p_0$ up to the p_0-null union of the stable manifolds of the saddle points; hence the pushforward of p_0 is the categorical law with weights $\alpha_k = p_0(\mathcal{R}_k)$. The total variation between two categorical laws on the same atoms $\{\mathbf{x}_k^\star\}$ is $\frac{1}{2}\sum_k |\alpha_k - w_k|$, which is equation 23; the degenerate case $\alpha_{k^\star} = 1$ gives $1 - w_{k^\star}$. Because equation 22 is *autonomous*, replacing c by $c' > 0$ merely reparameterizes time—$\mathbf{z}_s^{(c')} = \mathbf{z}^{(c)}_{(c'/c)s}$, by uniqueness of solutions for the locally Lipschitz field—so the trajectories trace the same orbits, every region of attraction is unchanged, and so is every α_k; the value is therefore c-independent. (This is exactly why we state the caricature in autonomous source coordinates: for the non-autonomous field $\nabla_{\mathbf{x}} \log r_t$ along the ODE, scaling by c does *not* in general reduce to a time reparameterization.) □

The mismatch in equation 23 is the asymptotic counterpart of the empirical plateau in Table 1: the guidance error is a *mode-weight mismatch* fixed by the posterior's basin geometry rather than by a tunable step size: tuning c cannot alter the mismatch at all, so whenever $\alpha \neq w$ no choice of c removes it, and it vanishes only in the exceptional case $\alpha = w$. We stress what this does *not* say. It does not say that greedy guidance must collapse onto one mode: that is only the degenerate case $\alpha_{k^\star} = 1$, which requires the extra assumption that essentially all p_0-mass lies in a single region of attraction, and we do not claim it holds in general. The generic statement is the weaker and more robust one—the surviving weights are attraction-region weights rather than posterior weights—and it is the mis-weighting, not exact single-mode collapse, that we actually observe (Table 1, mode-weight ℓ_1).

B.5 $\mathbf{u}^\star$ AS A FREIDLIN–WENTZELL MINIMUM-ACTION CORRECTION

Finally we make explicit the action interpretation of the canonical corrector. For the small-noise controlled dynamics $d\mathbf{x}_t = (\boldsymbol{v}_t + \mathbf{u}_t)\,dt + \sqrt{\epsilon}\,dW_t$, the Freidlin–Wentzell rate of a path γ is $S[\gamma] = \frac{1}{2}\int_0^1 \|\dot\gamma_t - \boldsymbol{v}_t(\gamma_t)\|^2\,dt$ (Kifer, 1988; Dupuis & Ellis, 1997). The most likely way to transport p_0 to $p(\cdot \mid \mathbf{y})$ minimizes $\mathbb{E}\,S$ over controls with the prescribed endpoint law, which is exactly the kinetic-energy objective equation 4 with $\mathbf{u}_t = \dot\gamma_t - \boldsymbol{v}_t(\gamma_t)$; its minimizer is the Benamou–Brenier geodesic relative to $\boldsymbol{v}_t$ (Appendix A). Thus $\mathbf{u}^\star$ is simultaneously (i) the minimum-kinetic-energy corrector, (ii) the deterministic ($\epsilon \to 0$) Schrödinger-bridge/Föllmer drift, and (iii) the Freidlin–Wentzell minimum-action control from the unconditional source to the posterior—three views of one object. From the reweighted source $p_0^{\mathbf{y}}$, the required action is zero (Proposition 1); the action is the price of starting from p_0.

C ADDITIONAL QUANTITATIVE RESULTS

C.1 ROBUSTNESS TO SAMPLE SIZE: A FULL n=50 COMPARISON

Table 2's comparison uses 10 images per operator, shared across methods; here we re-run *every* method on *all three* AFHQ operators at n=50 (D-Flow at n=15, given its cost of $\sim$ 500s/image), rather than spot-checking individual numbers.

Table 6: **Full AFHQ-Cat comparison at** n=50 (D-Flow: n=15; PSNR ↑ / LPIPS ↓). The qualitative conclusions involving our method are preserved; D-Flow and OT-ODE swap order on inpainting, and ours/OT-ODE remain a near-tie on deblurring.

Method	Super-resolution	Gaussian deblur	Box inpainting
FLOWER	25.89 / 0.282	27.85 / 0.285	25.98 / 0.070
OT-ODE	25.78 / <u>0.115</u>	27.01 / 0.130	24.25 / 0.096
Flow-Priors	23.54 / 0.264	26.45 / 0.190	26.13 / 0.120
D-Flow	24.52 / 0.170	27.47 / 0.158	25.76 / 0.078
Ours	25.42 / **0.144**	26.37 / **0.130**	26.19 / **0.050**

Every comparison reported in the main text survives at this scale, unchanged in direction and mostly tighter in magnitude (paired Wilcoxon on per-image LPIPS): ours < FLOWER on all three operators, now with $p < 10^{-7}$ throughout (super-resolution $p{=}4\times10^{-14}$, deblur $p{=}2\times10^{-15}$, inpainting $p{=}2\times10^{-8}$)— the inpainting comparison was only marginal at n=10 (p=0.064) and is now highly

significant. Against OT-ODE the same pattern holds as at n=10: significantly worse on super-resolution (p=7 $\times$ 10^{-10}), significantly better on inpainting (p=4 $\times$ 10^{-15}), and genuinely tied on deblurring (p=0.73; the means are now within 0.001 LPIPS of each other, 0.1297 vs. 0.1303). The larger D-Flow sample additionally resolves the earlier power problem: at n=15, ours is significantly better on all three operators (p=0.004, 0.015, 0.0004), where n=5 gave only a favorable trend. CelebA robustness (n=100 super-resolution: ours 0.029 vs. OT-ODE 0.026) and the OOD advantage at n=50 (Table 9) are unchanged from the earlier check. In short, the n=10 main table is representative, not favorably selected.

C.2 Ablation

Table 7 isolates the ingredients of our solver on box inpainting, addressing whether the quality comes merely from gradient normalization or step-size tuning. It does not: removing the Tweedie endpoint estimate (guiding on $\mathbf{x}_t$ rather than $\hat{\mathbf{x}}_1$) collapses LPIPS from 0.045 to 0.248, and removing per-image gradient normalization is catastrophic (the unnormalized update diverges for all but the smallest γ, peaking at LPIPS 0.24). Both ingredients are necessary, not incidental. The mobility schedule is a minor knob (constant slightly beats $\sigma^2_{r,t}$, which vanishes near $t = 1$); guidance strength behaves monotonically without a sharp optimum; and the sample mean over draws steadily improves PSNR (26.9 $\rightarrow$ 28.2 $\rightarrow$ 29.2 for $1/4/8$ samples), the ensembling benefit that single-shot solvers forgo.

Table 7: **Ablation** (box inpainting, 20 images, PSNR ↑ / LPIPS ↓). The Tweedie endpoint and gradient normalization are both essential; mobility and guidance strength are secondary; more conditional samples monotonically improve the sample-mean PSNR.

Variant	PSNR	LPIPS
Full (γ=2, const mobility, normalized, Tweedie)	**26.87**	**0.045**
− Tweedie endpoint (guide on $\mathbf{x}_t$)	18.80	0.248
− gradient normalization (best γ=0.01)	17.50	0.239
mobility $\sigma^2_{r,t}$ (vs. constant)	25.26	0.120
guidance $\gamma = 0.5 \,/\, 1 \,/\, 4$	16.12 / 24.07 / 26.95	0.427 / 0.134 / 0.059
splitting steps N=50 / 100 (fixed γ)	23.77 / 26.87	0.154 / 0.045
conditional samples $1 \,/\, 4 \,/\, 8$	26.87 / 28.16 / 29.24	0.045 / 0.042 / 0.042

C.3 A second in-domain prior and out-of-distribution priors

We repeat the study with a second pretrained flow prior, CelebA (128×128 faces), and additionally probe robustness with two *out-of-distribution* (OOD) settings in which the AFHQ-*cat* prior is applied to AFHQ-*dog* and AFHQ-*wild* images (prior–data mismatch). (Only AFHQ-cat and CelebA flow priors are publicly released; we use both, plus the OOD reuse of the cat prior.)

Second in-domain dataset (CelebA). Table 8 shows the same pattern as AFHQ: our solver is competitive on quality—best or second-best LPIPS on super-resolution and random inpainting—without being uniformly top on PSNR. Its uncertainty remains strongly error-correlated on faces: $\rho = 0.65 \pm 0.02$ (inpainting, 100 images), again far exceeding the zero uncertainty of the point-estimate baselines. The sample mean over 8 draws lifts super-resolution to $32.9/0.021$ (PSNR/LPIPS), competitive with the best baseline.

Out-of-distribution priors (AFHQ-dog / -wild). Under prior–data mismatch our solver *generalizes markedly better perceptually* than guidance (super-resolution, n=50): on dog images ours attains LPIPS 0.183 vs. FLOWER's 0.260, and on wild images 0.195 vs. 0.309 (Table 9); the inpainting LPIPS are also lower on both (dog 0.061 vs. 0.071, wild 0.058 vs. 0.078). Guidance overcommits to the (mismatched) prior's modes, whereas our correction stays closer to the measurement. Uncertainty still localizes correctly on the inpainted region even OOD (per-image $\rho \approx 0.7$; Figure 3).

Table 8: **CelebA** (128^2, 20 imgs; PSNR ↑/LPIPS ↓). Best/2nd LPIPS per row in **bold**/underline. Uncertainty–error $\rho = 0.65 \pm 0.02$ (inpainting, 100 images).

Operator	Ours	FLOWER	OT-ODE	D-Flow
Super-res.	30.9 / 0.028	32.3 / 0.032	31.6 / **0.018**	29.5 / 0.036
Inpainting	31.0 / 0.028	31.2 / **0.018**	29.8 / 0.033	31.0 / 0.021
Deblur	31.5 / 0.064	34.9 / **0.025**	–	–
Rand.-inp.	33.1 / **0.011**	33.8 / 0.020	–	–

Table 9: **OOD** (cat prior on dog/wild, n=50; LPIPS, ours wins all).

	Ours	FLOWER
dog SR	**.183**	.260
dog inp	**.061**	.071
wild SR	**.195**	.309
wild inp	**.058**	.078

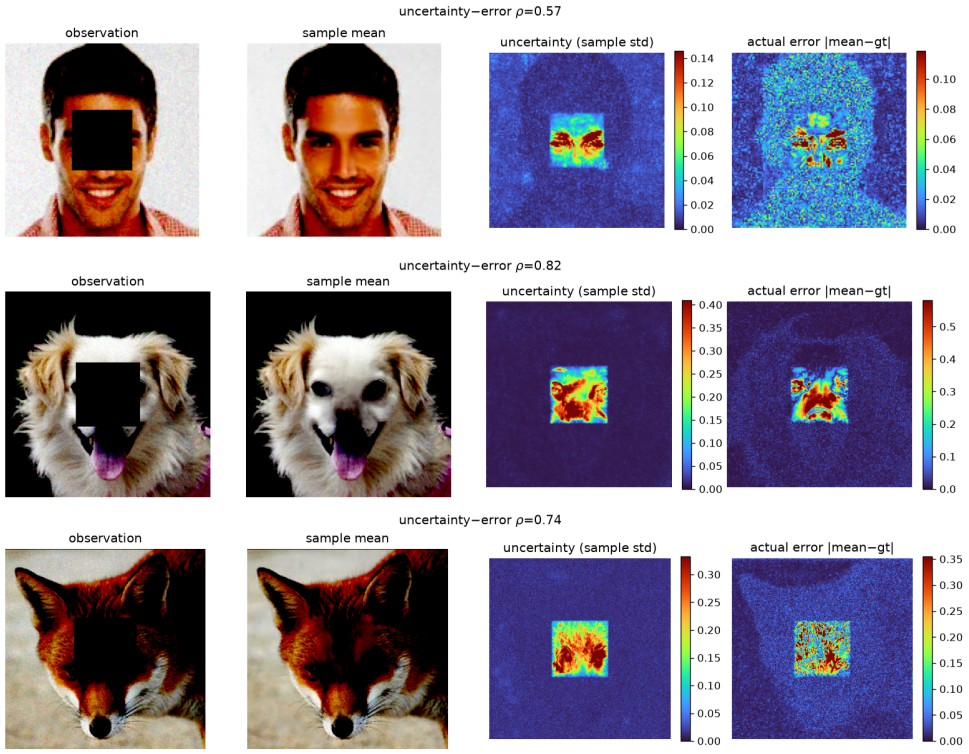


Figure 3: **Uncertainty localizes across datasets and OOD** (box inpainting). Rows: CelebA face, AFHQ-dog (OOD), AFHQ-wild (OOD). Columns: observation, sample mean, uncertainty (sample std), error. The uncertainty tracks the error on the inpainted region even when the prior is mismatched (per-image ρ in titles).

C.4 Is the uncertainty–error correlation an artefact of the mask?

The global correlation ρ of Table 4 is computed over all pixels. For box inpainting this is open to an obvious objection: the reconstruction error is concentrated in the missing region and so is the sample spread, so a positive ρ could reflect nothing more than the fact that both quantities know where the mask is. We therefore recompute the correlation under two controls, on the same run and the same images.

Within-mask correlation. Restrict both the per-pixel sample std and the per-pixel error to the missing pixels (9.8% of the image) and correlate there. This removes the in/out contrast entirely and asks only whether, *inside* the region the solver has to invent, higher uncertainty goes with larger error.

Partial correlation. Remove the group mean of both quantities inside and outside the mask, then correlate the residuals over the whole image. This is the partial correlation of uncertainty and error controlling for the mask indicator.

Over 40 AFHQ images with 8 samples (95% CIs), the within-mask correlation is 0.44 ± 0.05, the partial correlation is 0.43 ± 0.05, and the global correlation on the same images is 0.76 ± 0.02. The controls therefore cost roughly half of the global value, and we do not claim otherwise in the main text: the mask *does* inflate the headline number. What survives is a moderate positive within-region correlation that holds on every image tested (no image gives a negative value under either control), so the uncertainty map carries information beyond the mask geometry, but much less than the global ρ alone would suggest. Super-resolution provides the complementary check: its degradation has no mask, no spatial support to give away, and it still yields $\rho = 0.53 \pm 0.02$ over 20 images. We consider the honest summary to be that the uncertainty is *error-correlated* in a useful but modest sense, which is also why Section 7.1 treats this as indirect rather than decisive evidence of posterior faithfulness.

C.5 Reading the three box-inpainting numbers

Three tables report box inpainting for our solver and the values differ, so we state the axes explicitly. Table 2 (25.57 / 0.052) is a *single* sample on the 10 shared benchmark images, matching how every baseline in that table is run. Table 7 (26.87 / 0.045) is also a single sample but on a 20-image set, since the ablation sweeps many configurations and benefits from the extra images. Table 4 (28.15) is the 8-*sample mean* over 100 images, the setting in which an uncertainty map is defined at all. The two axes—number of samples and image set—are stated in each caption, and their separate effects are visible within a single controlled sweep: at a fixed 20-image set the sample count alone moves PSNR $26.87 \to 28.16 \to 29.24$ for $1/4/8$ samples (Table 7), while at a fixed 8 samples the image set moves it from 27.64 (10 images) to 28.15 (100 images). No number is computed from a different configuration than its caption states.

C.6 Ensembling the baselines: protocol

The comparison in Table 3 answers a question the main text should not leave open: can the trajectory-guidance baselines simply be rerun with independent initial noise to obtain an ensemble and an uncertainty map? They can, and the posterior-transport view predicts as much—Proposition 1 locates the conditional variability in the source draw, not in the corrector—so we measured it rather than asserting otherwise.

Protocol. For FLOWER and OT-ODE we ran the released benchmark implementations 8 times per image on the 10 AFHQ box-inpainting images of Table 2. The measurement noise is drawn from a seed fixed by the image index, so all 8 runs see the *same* noisy observation; only the random number stream feeding each solver's own initialization is advanced, making the 8 outputs independent conditional draws of one problem instance. We then treat the 8 reconstructions exactly as we treat ours: the ensemble spread is the mean across-draw per-pixel standard deviation, and ρ is the Pearson correlation between that per-pixel standard deviation and the per-pixel error of the ensemble mean, computed per image and averaged (95% CI over images). No solver hyperparameter was changed.

Result and reading. FLOWER reaches $\rho = 0.74 \pm 0.04$ and OT-ODE 0.76 ± 0.03, against 0.75 ± 0.04 for our solver on the same images, so error-correlated uncertainty is a property of the *family*, not of our corrector. The three differ in spread (0.009, 0.029, 0.028): FLOWER's ensemble is the tightest by a factor of three, which is consistent with its conjugate-gradient projection pulling every draw toward the same measurement-consistent point. D-Flow is the exception in kind rather than degree: it optimizes the source latent to a single MAP-like point rather than sampling, so it does not natively provide an ensemble (we did not run it from multiple initializations to check empirically how tight any such ensemble would be). We therefore removed the uniqueness claim from the main text and kept only what the measurement supports—that our solver obtains the same uncertainty signal at 3–19$\times$ lower cost.

C.7 Why FlowDPS is not in the image benchmark

FlowDPS (Kim et al., 2025) is one of the solvers our analysis targets, so its absence from Table 2 deserves a full statement. The released implementation is written for Stable Diffusion 3: it encodes a text prompt, runs in the SD3 latent space, and its update rule is tied to that architecture (the entry point is an SD3 sampler that calls `encode_prompt` and a `StableDiffusion3Pipeline`).

Every prior in this paper's image benchmark—the AFHQ-Cat and CelebA optimal-transport flow-matching models released with FLOWER, and the only publicly released flow priors for these datasets—is unconditional and acts directly in pixel space. There is no text conditioning to supply and no latent space to guide in, so FlowDPS cannot be pointed at these priors; running it would require reimplementing the method for a different class of prior, and any resulting row would report the quality of our reimplementation rather than of FlowDPS. Because a mechanism-level reimplementation is exactly what a strawman looks like, we instead evaluate the FlowDPS *mechanism*—Tweedie-endpoint likelihood-gradient guidance with a scalar mobility—in the one setting where it can be isolated without a prior confound: the controlled 2D study, where every method acts on the same analytic flow prior against a closed-form posterior (Table 1). OT-ODE, which appears throughout the image tables, is the closest runnable relative of that mechanism for an unconditional pixel-space flow prior.

C.8 Why SITCOM is not reported on the 2D prior

SITCOM (Alkhouri et al., 2025) enforces measurement, forward- and backward-diffusion consistency by optimizing the network input at every step. Its triple-consistency objective is defined for the diffusion parameterization, and our attempts to port it to the deterministic 2D flow were numerically unstable across the step-size and inner-iteration grids we tried. Rather than report a number that would understate a method we could not tune faithfully, we omit it from Table 1 and discuss it in Section 2.

D Additional qualitative results

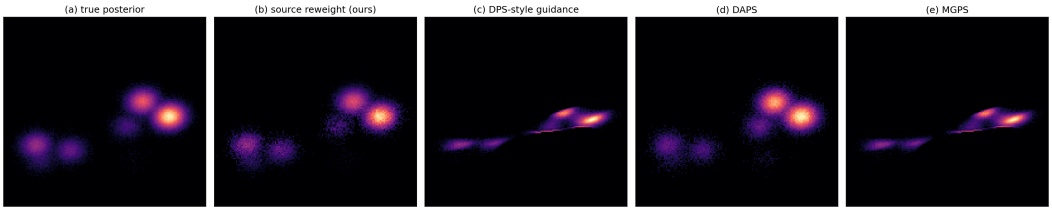


Figure 4: **Full mechanism comparison on the controlled 2D prior** (densities; companion to Table 1). **(a)** closed-form true posterior; **(b)** exact source reweighting (oracle)—matches (a); **(c)** DPS-style guidance collapses onto a measurement-consistent ridge; **(d)** DAPS partially recovers the mode structure but mis-weights it; **(e)** MGPS, like guidance, collapses toward a ridge. Only source-side reweighting reproduces the posterior; all tested trajectory/guidance methods remain biased at the practical budgets used here.

Across all four restoration tasks the uncertainty pattern adapts to the operator and tracks the error, with 3 examples each below (columns: observation, sample mean, per-pixel uncertainty (sample std), actual error; per-image ρ in titles).

E Method and experimental details

E.1 Velocity-correction solver

Guided by Proposition 3, our solver integrates the prior ODE from a random source draw and applies, at each step, a mobility-modulated and *per-image normalized* measurement correction aligned with the instantaneous likelihood score equation 10 (using the Tweedie endpoint $\hat{\mathbf{x}}_1$). Normalization is essential for stability: the raw gradient of $\|H\hat{\mathbf{x}}_1 - \mathbf{y}\|^2/(2\sigma^2)$ has magnitude $O(1/\sigma^2)$ and, applied without normalization, diverges. Algorithm 1 summarizes one conditional sample; S independent draws give a sample mean (reconstruction) and a per-pixel standard deviation (uncertainty map).

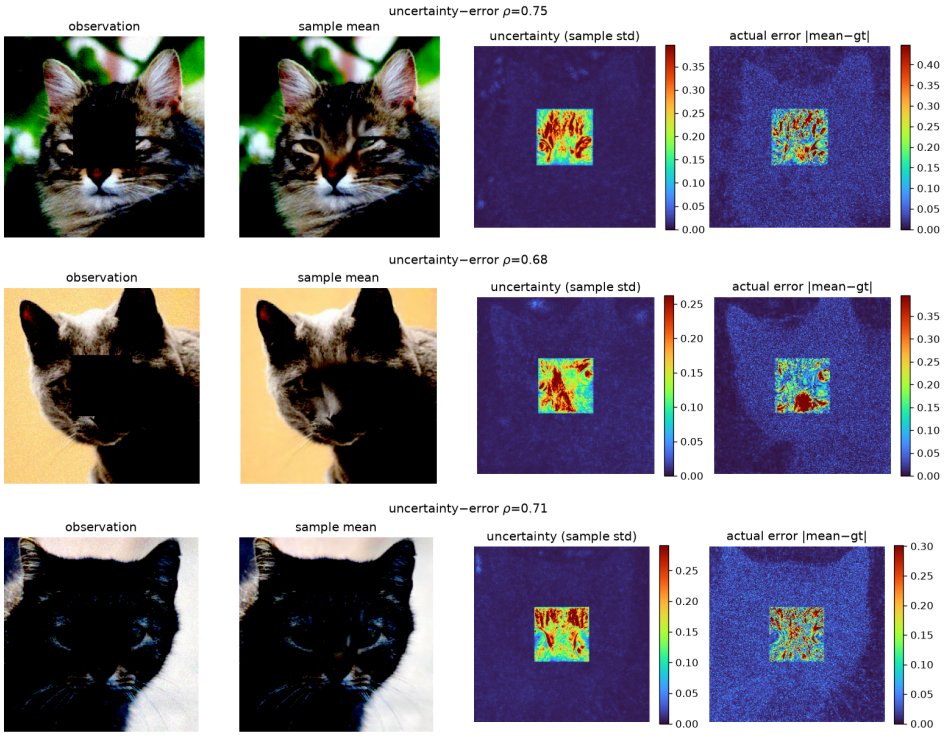


Figure 5: **Additional uncertainty maps** (box inpainting), including darker / harder cases. Per row: observation, sample mean, per-pixel uncertainty (across-sample std), actual error. The uncertainty localizes on the inpainted region across a range of image appearances.

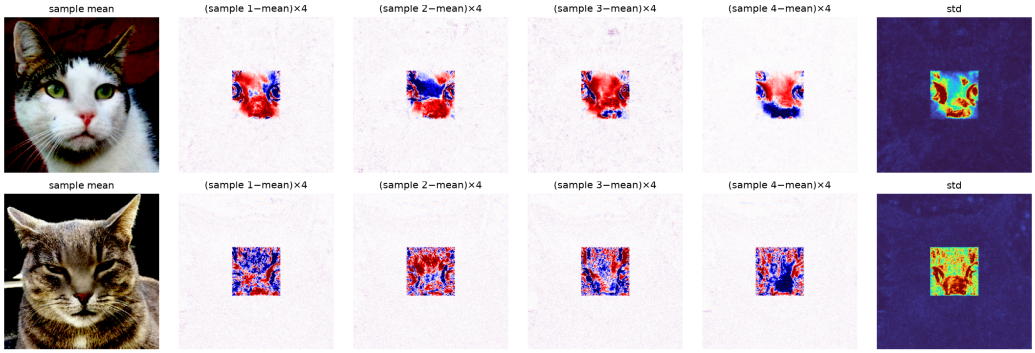


Figure 6: **Conditional-sample diversity, amplified** (box inpainting). Per row: sample mean; four per-sample deviations $(\text{sample}_i - \text{mean}) \times 4$ (red/blue), and the across-sample std. The deviations are concentrated in—and differ across samples within—the inpainted region: the modest but genuine inter-sample variation that single-shot solvers do not capture. (Raw 8-sample montages are visually near-identical precisely because the inter-sample variation is small in magnitude; we therefore visualize it amplified.)

The implemented step is a discrete data correction, not an Euler step of $\dot{\mathbf{x}} = \boldsymbol{v}_t + \mathbf{u}_t$. We flag a discrepancy between the continuous-time analysis and the practical solver, because it changes how Algorithm 1 should be read. Line 6 advances the prior velocity by $\delta = 1/N$ but applies the measurement term at *full* magnitude,

$$\mathbf{x} \leftarrow \mathbf{x} + \delta\, \boldsymbol{v} - \gamma\, w_t\, \hat{\mathbf{g}} \qquad \text{rather than} \qquad \mathbf{x} \leftarrow \mathbf{x} + \delta\big(\boldsymbol{v} - \gamma\, w_t\, \hat{\mathbf{g}}\big). \tag{24}$$

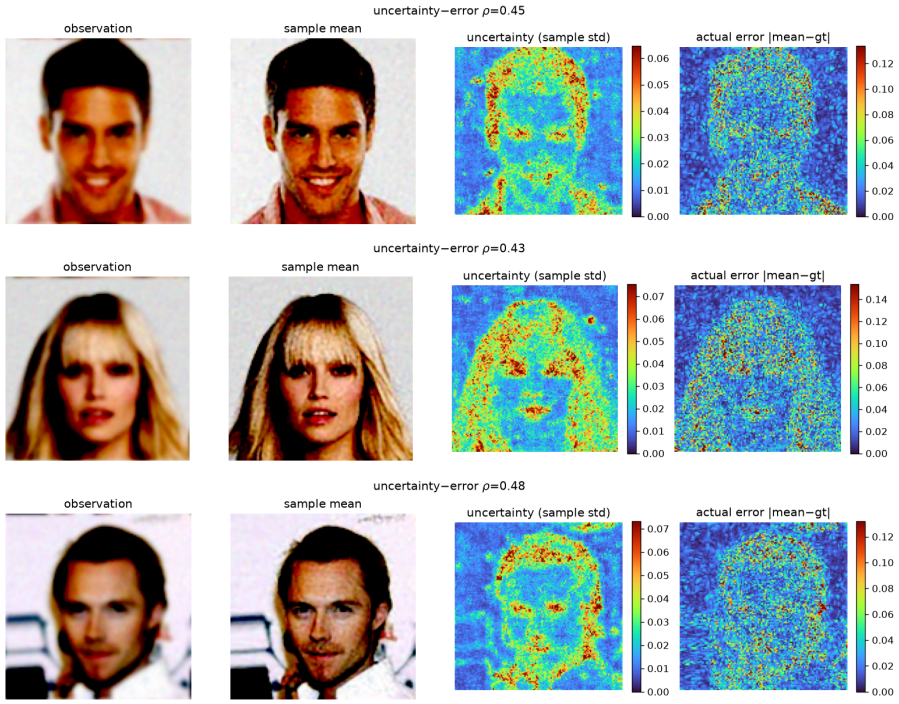


Figure 7: **Gaussian deblurring** (CelebA, 3 examples). Uncertainty concentrates on edges and fine facial structure—the high-frequency content the deblurring must restore.

Because $\hat{\mathbf{g}}$ is normalized to unit norm, each iteration displaces the state by exactly γw_t along the measurement direction: a discrete per-step data-correction scheme, interpretable as an N-dependent effective correction field $\mathbf{u}_t = -(\gamma/\delta)\, w_t \hat{\mathbf{g}} = -\gamma N\, w_t \hat{\mathbf{g}}$ rather than a fixed continuous one, and not the Euler discretization of the corrected ODE. (We do not call this "operator splitting": canonical splitting applies the second operator at the point the first produces, whereas here the correction and $\boldsymbol{v}$ are both evaluated at the same $\mathbf{x}_k$.) Its magnitude scales with the step count, so γ absorbs N and the cumulative correction is controlled by the product $N\gamma$, not by γ alone— the sense in which our solver is "motivated by, not derived from" the analysis (Section 7.1). Theorem 1 can be applied to the corresponding N-dependent effective field under its assumptions, although this does not provide a step-count-independent guarantee for the implemented scheme.

Verification. We confirmed this prediction directly (box inpainting, 20 images, mobility const): holding $N\gamma = 200$ fixed while varying $N \in \{50, 100, 200\}$ (and γ inversely) gives PSNR 26.79, 26.87, 26.82—statistically flat— whereas halving N at *fixed* γ (so $N\gamma = 100$) drops PSNR to 23.77, and applying δ to the correction as in the naive Euler discretization (so $N\gamma = 2$) collapses it to 9.41. The step count N is therefore not doing what an ODE discretization error analysis would suggest; the quantity that matters is $N\gamma$, exactly as equation 24 predicts.

Algorithm 1 Velocity-correction solver (one conditional sample)

input: measurement $\mathbf{y}$, operator H, noise σ, steps N, guidance γ, mobility $\{w_t\}$
$\mathbf{x} \sim \mathcal{N}(0, I)$; $\delta \leftarrow 1/N$
for $i = 0, \ldots, N-1$ **do**
  $t \leftarrow i\delta$; $\boldsymbol{v} \leftarrow \boldsymbol{v}_t(\mathbf{x})$; $\hat{\mathbf{x}}_1 \leftarrow \mathbf{x} + (1-t)\boldsymbol{v}$ {flow Tweedie endpoint}
  $\mathbf{g} \leftarrow \nabla_{\mathbf{x}} \left[\frac{1}{2}\|H\hat{\mathbf{x}}_1 - \mathbf{y}\|^2\right]$; $\hat{\mathbf{g}} \leftarrow \mathbf{g}/(\|\mathbf{g}\| + \epsilon)$ {per-image normalized}
  $\mathbf{x} \leftarrow \mathbf{x} + \delta\, \boldsymbol{v} - \gamma\, w_t\, \hat{\mathbf{g}}$
end for
return $\mathbf{x}$

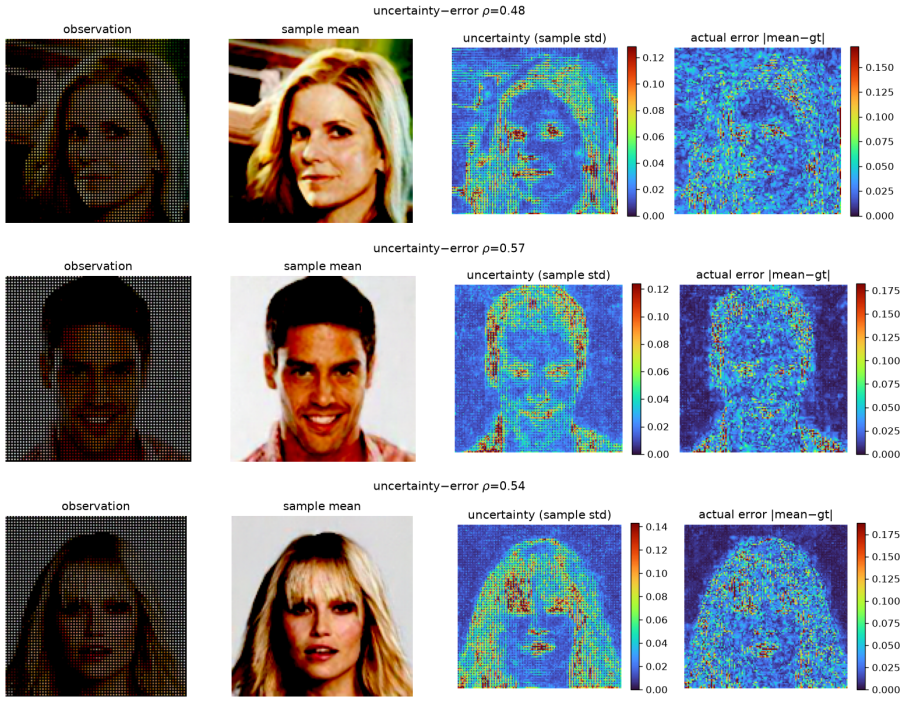


Figure 8: $\times 4$ **super-resolution** (CelebA, 3 examples). Uncertainty spreads over the high-frequency detail that super-resolution hallucinates.

Mobility schedules. We use either a constant mobility $w_t \equiv 1$ (CONST) or the flow noise level $w_t = \sigma_{r,t}^2$ with $\sigma_{r,t} = (1-t)/\sqrt{t^2 + (1-t)^2}$ (SIGMA_R). The latter vanishes as $t \to 1$, removing late-stage data correction; CONST retains correction throughout and is our default. The correction requires one forward and one backward through the endpoint per step—*no* backpropagation through the ODE, in contrast to source-space optimizers such as D-Flow.

Hyperparameters. Unless stated, $N = 100$ splitting steps, $\gamma = 1$, mobility CONST; uncertainty experiments use $S = 8$ samples. All AFHQ runs use the public pretrained OT flow-matching prior and the FLOWER benchmark degradations at $\sigma = 0.05$; we report PSNR and LPIPS (AlexNet). Baselines use their released, per-problem-tuned hyperparameters.

E.2 CLOSED-FORM 2D STUDY

Prior and exact velocity. The target is a Gaussian mixture $p_1 = \sum_{k=1}^{K} w_k \mathcal{N}(\mu_k, \sigma_p^2 I)$ ($K = 8$, random μ_k in $[-5, 5]^2$, random w_k, $\sigma_p = 0.35$) and the source is $p_0 = \mathcal{N}(0, I)$. For the straight (rectified-flow) conditional path $\mathbf{x}_t = (1-t)\mathbf{x}_0 + t\mathbf{x}_1$ with independent $\mathbf{x}_0 \sim p_0, \mathbf{x}_1 \sim p_1$, the marginal velocity $\boldsymbol{v}_t(\mathbf{x}) = \mathbb{E}[\mathbf{x}_1 - \mathbf{x}_0 \mid \mathbf{x}_t = \mathbf{x}]$ is available in closed form. Conditioned on component k, $\mathbf{x}_t \mid k \sim \mathcal{N}(t\mu_k, s_t^2 I)$ with $s_t^2 = (1-t)^2 + t^2\sigma_p^2$, and the joint Gaussianity gives

$$\mathbb{E}[\mathbf{x}_1 \mid \mathbf{x}_t, k] = \mu_k + \tfrac{t\sigma_p^2}{s_t^2}(\mathbf{x}_t - t\mu_k), \qquad \mathbb{E}[\mathbf{x}_0 \mid \mathbf{x}_t, k] = \tfrac{1-t}{s_t^2}(\mathbf{x}_t - t\mu_k), \tag{25}$$

so that $\boldsymbol{v}_t(\mathbf{x}) = \sum_k r_k(\mathbf{x}) \big(\mathbb{E}[\mathbf{x}_1 \mid \mathbf{x}, k] - \mathbb{E}[\mathbf{x}_0 \mid \mathbf{x}, k]\big)$ with responsibilities $r_k(\mathbf{x}) \propto w_k \mathcal{N}(\mathbf{x}; t\mu_k, s_t^2 I)$. This makes the flow map essentially exact (no learned model), isolating the posterior-sampling question from approximation error.

Measurement and exact posterior. With a linear measurement $\mathbf{y} = A\mathbf{x} + \boldsymbol{\eta}$, $\boldsymbol{\eta} \sim \mathcal{N}(0, \sigma^2 I)$, the posterior is again a Gaussian mixture: each component k has $\Sigma_{\text{post}} = (\sigma_p^{-2} I + A^\top \sigma^{-2} A)^{-1}$, mean $\Sigma_{\text{post}}(\sigma_p^{-2}\mu_k + A^\top \sigma^{-2}\mathbf{y})$, and updated weight $\propto w_k \mathcal{N}(\mathbf{y}; A\mu_k, A\sigma_p^2 A^\top + \sigma^2 I)$. We draw from this closed-form posterior for the ground-truth comparison.

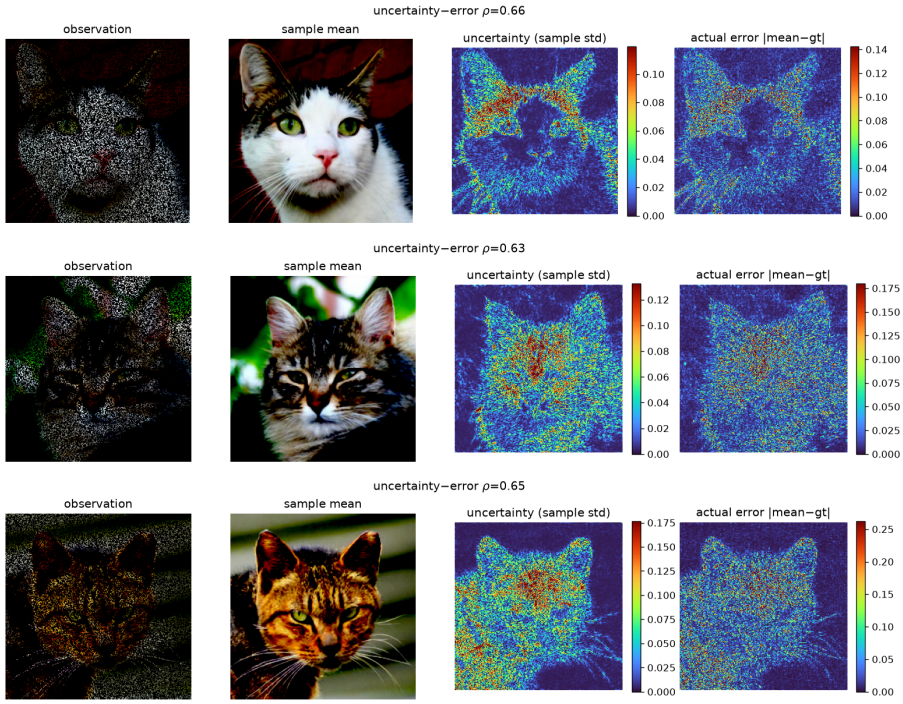


Figure 9: **Random inpainting** (AFHQ-cat, 3 examples). Uncertainty localizes on the missing pixels.

Reproducibility details. We re-ran the released scripts to compile this subsection; two values in earlier drafts of this paper were transcription errors that we correct here against the scripts' actual configuration, not new experiments at a different scale: the sample count was written as "20k" but the scripts use N=12k throughout, and the sliced-W_2 projection count was written as "1000" but the scripts use 400. The metric *values* reported in Table 1 are unchanged from earlier drafts and were re-verified against the corrected N=12k, 400-projection configuration when we prepared this subsection. All numbers use 200-step RK4 integration of the (exact, closed-form) velocity for source reweighting and FlowDPS-style guidance, and 100 Euler steps for FLOWER/DAPS/MGPS. The prior means/weights are drawn with a fixed generator seed (seed=1 for $\{\mu_k, w_k\}$); the global seed is $\mathrm{torch.manual_seed}(0)$ for all sampling calls (source draws, resampling, DAPS/MGPS re-noising). The measurement is a single linear functional $A \in \mathbb{R}^{1\times 2}$, drawn once with seed=7 and normalized to $\|A\|$=1 (so $\mathrm{rank}(A)$=1 $<$ d=2: an underdetermined, sub-Nyquist measurement, chosen deliberately over a random data point so that several modes remain measurement-consistent); the observation is $\mathbf{y} = \mathrm{median}_k(A\mu_k)$, the median of the components' projections, and the observation noise is σ=0.6 (denoted γ in the released code). The exact posterior, used both for the ground-truth draws and the two-sample reference, is the closed-form Gaussian-mixture posterior of the paragraph above.

Methods and metrics. *Source reweighting* integrates the unmodified ODE from 12k draws of p_0 and weights by $p(\mathbf{y} \mid \Phi(\mathbf{z}))$ (Lemma 1); we report effective sample size and use weighted metrics to remove resampling variance. *FlowDPS-style guidance* adds $g \cdot \nabla_{\mathbf{x}_t} \log p(\mathbf{y} \mid \hat{\mathbf{x}}_1)$ to the velocity, swept over $g \in \{0.5, 1, 2, 4\}$. *DAPS* (Table 1) runs 100 decoupled-annealing steps: at each step it forms the *exact* Gaussian endpoint posterior $p(\mathbf{x}_1 \mid \mathbf{x}_t, \mathbf{y})$ under the local Gaussian/linear approximation (closed-form mean and covariance, using the same $A, \mathbf{y}, \sigma$), draws $\mathbf{x}_1$ from it, and re-noises with a *fresh* (decoupled) source draw interpolated toward that sample—the favorable, exact-inner-sampling version of DAPS possible only because the toy posterior is tractable. *MGPS* runs 100 Euler steps with guidance strength g=1, computing the likelihood gradient at the trajectory *midpoint* $\mathbf{x}_t + \frac{1}{2}\delta\, \boldsymbol{v}_t(\mathbf{x}_t)$ (via its own Tweedie endpoint) and applying it to the full step. *Ours* runs Algorithm 1's exact update rule (100 steps, normalized gradient, constant mobility) with the image-domain guidance strength ($\gamma \in \{1, 2\}$) replaced by a value retuned for this scale, since the image

value diverges in 2D; we sweep $\gamma \in \{0.005, 0.01, 0.02, 0.03, 0.05, 0.08, 0.1\}$ and report the value selected by energy distance (γ=0.02), the same selection rule used for the guidance sweep. Metrics are sliced-W_2 (400 random projections), energy distance, multi-scale RBF MMD^2, and the ℓ_1 error of mode weights (nearest-component assignment), each against a two-sample reference given by two independent posterior draws.

F Relation to existing frameworks

Our results sit at the intersection of several literatures. Below we treat each in turn, in every case separating what we *borrow* (established machinery) from what is *new* here (its specialization to deterministic flow-matching posteriors and the consequences for guidance-based solvers). A one-line summary: the source-reweighting identity and the optimal-transport/Schrödinger machinery are classical; their synthesis into a *deterministic-flow posterior-transport account*—with the zero-drift dichotomy, the guidance-as-approximate-corrector reading, and the large-deviation explanation of mode collapse—is, to our knowledge, new.

Source-Guided Flow Matching, ESS-Flow, and guidance-as-drift. Wang et al. (2025) prove, independently of us, the same identity we state as Lemma 1: reweighting the source distribution while leaving a pretrained flow's velocity field unmodified recovers the target conditional exactly. Their Theorem 2 gives a Wasserstein bound for approximate source samplers, separating source-sampling error from velocity-model error—the same object our Theorem 1 bounds, though for the different error source of an approximate *trajectory* correction rather than an approximate *source sampler*. They discuss importance sampling, MALA/ULA/HMC, and D-Flow as instances of source-space inference—these differ sharply in whether they need to differentiate through the flow map: importance sampling does not, MALA/ULA/HMC need $\nabla_{\mathbf{z}} \log p(\mathbf{y} \mid \Phi(\mathbf{z}))$ (one differentiation through Φ, as we discuss below in the *Latent MCMC* paragraph), and D-Flow's LBFGS inner loop backpropagates through the *entire* ODE integration at every iteration, which is what makes it the most expensive method in Table 3—and validate on 2D synthetic data, a Darcy-flow PDE inverse problem, and, on CelebA, denoising, deblurring, super-resolution, and random and box inpainting (their Section 5.3, Table 2): image-restoration tasks essentially the same as ours, just on a different dataset. Kalaivanan et al. (2025) likewise perform training-free guidance as source-space inference, requiring "only forward passes through the generative model and observation process, no gradient or Jacobian computations"—an a gradient-free mechanism distinct from ours (we avoid backpropagating through the ODE but still take one gradient through the endpoint map per step); we do not compare total sampling cost, only the presence or absence of any differentiation through Φ. Their stated limitation is degrading when the prior is uninformative about the target, e.g. in noiseless inpainting. They do report a multi-sample standard deviation on their materials- and protein-generation tasks (not images), but neither a per-pixel uncertainty map, an uncertainty–error correlation, nor an out-of-distribution experiment; Wang et al. (2025) report none of these either. Section 7 reports all three, on images. Feng et al. (2025) instead take the complementary route—modifying the velocity field rather than the source—and give a unified guidance framework with asymptotically exact and local-approximation variants; their route and ours are opposite answers to the same question about where conditioning information should live, rather than overlapping claims. What none of these three papers develops is what we take to be this paper's central dynamical contribution: the connection between source reweighting and the small-noise limit of diffusion Doob h-transform guidance (Proposition 2, the zero-drift dichotomy), the resulting Benamou–Brenier/Schrödinger/Freidlin–Wentzell characterization of the canonical corrector $\mathbf{u}^\star$ (Section B.1), or the placement of existing *trajectory-guidance* solvers (FlowDPS, FLOWER, PnP-Flow) as local approximations within that correction framework (Proposition 3). In short: the source-space exact-transport identity itself is not unique to this paper, and image-restoration experiments on a face dataset already exist for it; our contribution is the dynamical/correction-field account built on top of the identity, its consequences for existing trajectory-guidance solvers, and the AFHQ/OOD/uncertainty results that follow from applying our own solver at that scale.

Normalizing-flow and latent-space Bayesian inference. An invertible generator $\mathbf{x} = \Phi(\mathbf{z})$ turns a data posterior into a reweighted source posterior $p_0^{\mathbf{y}} \propto p(\mathbf{y} \mid \Phi(\mathbf{z}))p_0(\mathbf{z})$ (Lemma 1). This is the change-of-variables identity at the heart of variational inference with normalizing flows (Rezende & Mohamed, 2015; Papamakarios et al., 2021) and of latent-variable Bayesian inference more broadly;

encoder/decoder posterior inference, flow-based VI, and "inference amortization" all exploit it. *Borrowed:* the identity itself. *New:* its *dynamical* consequence for a flow-matching prior (Propositions 1–2). The reweighting $r_t(\mathbf{x}) = p(\mathbf{y} \mid \Phi_{t\to 1}(\mathbf{x}))/p(\mathbf{y})$ is a first integral of the prior flow ($\frac{D}{Dt} r_t = 0$), so the posterior marginal path is carried by the *unmodified* velocity field and conditioning requires no drift. Normalizing-flow posterior inference operates on the static map Φ; it does not analyze the time-marginal dynamics that flow-matching inverse solvers actually run, which is exactly where guidance methods inject their (drift) corrections.

Source-side posterior sampling: importance sampling, MCMC, SMC. Because conditioning lives in the source, drawing $p_0^{\mathbf{y}}$ is a latent-space inference problem, and every generic posterior sampler applies. *(i) Importance sampling:* reweighting unconditional draws $\mathbf{z} \sim p_0$ by $p(\mathbf{y} \mid \Phi(\mathbf{z}))$ is unbiased and is exactly our toy estimator; Proposition 6 explains, via the rate function $I_{\mathbf{y}}$, why its effective sample size can decay exponentially in dimension and inverse-noise. *(ii) Latent MCMC:* Langevin / HMC on $\log p_0^{\mathbf{y}}(\mathbf{z}) = \log p_0(\mathbf{z}) + \log p(\mathbf{y} \mid \Phi(\mathbf{z}))$ needs $\nabla_{\mathbf{z}} \log p(\mathbf{y} \mid \Phi(\mathbf{z})) = (\nabla_{\mathbf{z}} \Phi)^\top \nabla \log p(\mathbf{y} \mid \cdot)$, i.e. differentiation through the flow—the same backpropagation-through-the-ODE cost that makes D-Flow expensive (Table 3); D-Flow itself is the MAP/optimization limit of this chain and recent work adds latent Langevin ("D-Flow-SGLD"). *(iii) SMC:* particle filters that anneal the likelihood along the trajectory give asymptotically exact diffusion posterior samples (Wu et al., 2023; Cardoso et al., 2024) at the price of many particles and resampling. *New / our position:* we do not propose another source-side sampler; our velocity-correction solver is a cheap *amortized forward* surrogate (one ODE pass, no backprop) whose deviation from the exact posterior is controlled by the corrector gap in Proposition 3/Theorem 1. It is complementary to MCMC/SMC: it can warm-start them, and they can debias it.

Conditional and dynamic optimal transport. When one insists on starting from the unconditional p_0, the minimum-energy way to reach $p(\cdot \mid \mathbf{y})$ is the canonical corrector $\mathbf{u}^\star$ of equation 4, a *conditional* dynamic-OT problem relative to the reference drift $\boldsymbol{v}_t$, characterized by the Benamou–Brenier system (Benamou & Brenier, 2000). This connects measurement guidance to a substantial body of work on transport-based generation and conditioning—rectified/OT flow matching, and OT couplings between arbitrary endpoints. *Borrowed:* the Benamou–Brenier characterization and the gradient-field form of $\mathbf{u}^\star$. *New:* the reading of FlowDPS/FLOWER/PnP-Flow as *greedy, locally-linearized* approximations of this conditional-OT corrector (Tweedie endpoint + mobility), and the resulting Wasserstein bias bound. The lens matters: "likelihood gradient" suggests the only error is a step size, whereas "approximate conditional transport" makes clear the error is a mis-specified *field*, which Section 6 confirms cannot be tuned away.

Schrödinger bridges, the Föllmer drift, and the small-noise limit. The stochastic analogue of $\mathbf{u}^\star$ is the entropic Schrödinger bridge / Föllmer drift that steers a reference diffusion to a prescribed terminal law (De Bortoli et al., 2021; Zhang & Chen, 2022; Tzen & Raginsky, 2019); conditional/guided diffusion is the special case where the terminal constraint is a likelihood and the correction is the Doob h-transform drift $a_t \nabla \log h_t$ (Denker et al., 2024). Proposition 2 places the deterministic flow at the *boundary* of this family: as the bridge noise $\epsilon \to 0$, the drift $\epsilon^2 \nabla \log h_t^\epsilon \to 0$, so the small-noise limit is consistent with deterministic transport from the reweighted source. Hence diffusion guidance (drift-based) and flow conditioning (source-based) are two regimes of one continuum, and—this is the practical upshot—an h-transform drift designed for $\epsilon > 0$ has no nonzero limit to "port" onto a flow ($\epsilon = 0$); the information has migrated into the initial law. This is, to our knowledge, the first explicit account of where deterministic flow conditioning sits relative to the Schrödinger/Föllmer picture.

Large deviations and the geometry of mode collapse. Our use of the Laplace principle (Proposition 5) draws on standard large-deviation theory (Dembo & Zeitouni, 2010; Dupuis & Ellis, 1997; Kifer, 1988). *Borrowed:* Varadhan's lemma, the contraction principle, and Laplace's method. *New:* their application to the flow source posterior, in the two regimes we keep separate. Proposition 6 uses a jointly tempered Regime B to quantify the exponential ESS decay through its large-deviation rate $I_{\mathbf{y}}$. Separately, Propositions 5, 7 and 8 work in the fixed-prior Regime A, where the posterior basin weights are $w_k \propto p_0(\mathbf{z}_k^\star)\, |\nabla^2 (\ell_{\mathbf{y}} \circ \Phi)(\mathbf{z}_k^\star)|^{-1/2}$ and greedy source-space ascent generally replaces them by attraction-region weights—a structural, not a magnitude, error. The identification of $\mathbf{u}^\star$ with the Freidlin–Wentzell minimum-action correction from p_0 to $p(\cdot \mid \mathbf{y})$ ties the transport,

bridge, and large-deviation pictures together. A closely related small-noise action principle underlies the Onsager–Machlup posterior transport recently developed for deep Gaussian processes (Xu et al., 2026) and the Onsager–Machlup action-minimization view of generative transition-path sampling (Raja et al., 2025), where the target is characterized as the minimizer of an Onsager–Machlup functional along transport paths; our flow-source action $I_{\mathbf{y}}$ is the inverse-problem analogue of that construction, specialized to a deterministic probability-flow prior.

Diffusion posterior samplers with guarantees, and flow inverse solvers. On the diffusion side, provably consistent posterior samplers exist (Xu & Chi, 2024; Feng et al., 2023), targeting SDE priors via SMC/annealing; they share our goal (true posterior sampling) but not our object (a deterministic flow's velocity-plus-correction decomposition and source-transport viewpoint). On the flow side, FlowDPS (Kim et al., 2025), FLOWER (Pourya et al., 2026), PnP-Flow (Martin et al., 2025) and D-Flow (Ben-Hamu et al., 2024) are the methods we analyze; our contribution is not to outrank them on reconstruction metrics but to explain what their corrections approximate, how they mis-weight posterior modes, and what a faithful—and cheap—alternative looks like.

AI USE STATEMENT

Parts of this paper's text, code and analysis were produced with the assistance of generative AI. The authors take full responsibility for the final content of this paper, including any text produced with the assistance of generative AI.