DIRECTIONAL PHASE TRANSITIONS IN REWARD GUIDANCE FOR FLOW MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Recent advances in reward-guided generation with flow models have demonstrated remarkable performance in steering outputs toward desired reward functions. In particular, inference-time alignment methods are drawing attention, as they adjust the sampling trajectory without extensive retraining of the pretrained model. Recent analyses of sampling trajectories, however, reveal that they are not dynamically homogeneous but pass through distinct phases. Yet, the impact of these phases on inference-time reward alignment remains under-explored. To bridge this gap, we introduce the directional expansion rate, a metric that quantifies how the velocity field of a flow model amplifies or attenuates perturbations toward higher reward. We observe a consistent *contraction–expansion–contraction* pattern across distinct tasks: data assimilation, PDE solving. To explain this behavior, we derive the phase structure in a closed-form expression for a Gaussian mixture model, establishing a formal link to mode commitment. Based on this analysis, we propose an adjoint-based control strategy that operates exclusively within the expansion window to steer mode selection before the commitment occurs. Across the tasks, we demonstrate that the proposed method generates better-aligned samples with the target reward function.

1 INTRODUCTION

Flow models have become the dominant paradigm of generative models recently, producing high-fidelity samples across images, video, and scientific simulation (Labs et al., 2025; Jin et al., 2025; Chen et al., 2025; Cheng et al., 2025). Building on this, reward-aligned generation methods are widely explored to make generative models satisfy downstream criteria such as human preferences (Xu et al., 2023; Ma et al., 2025) and other task-specific objectives (Utkarsh et al., 2025; Yang et al., 2026). Accordingly, various methods for reward alignment have been proposed, which generally fall into two categories: fine-tuning approaches (Black et al., 2024; Fan et al., 2023; Liu et al., 2025) and test-time alignment approaches (Kim et al., 2026b; Long et al., 2026; Hwang et al., 2026). Driven by the computational cost of training, test-time approaches have gained attention, improving performance without extra training. Within this paradigm, control-theoretic approaches stand out by accounting for how interventions propagate through the sampling dynamics for the terminal reward. By framing generation as trajectory optimization, prior works leverage adjoint states from Pontryagin's Maximum Principle (PMP) to align trajectories with the reward (Wang et al., 2025; Chang et al., 2026; Huang et al., 2026). While these frameworks address how to steer the trajectory, when to intervene remains underexplored. Prior work reports that intervention timing substantially affects sample quality, but treats it merely as a hyperparameter to tune (Kynkäänniemi et al., 2024; Huang et al., 2026). This suggests that finding the optimal intervention timing requires a systematic analysis of its interaction with the underlying generative dynamics.

This perspective connects naturally to a parallel line of research that characterizes the sampling trajectory as temporally inhomogeneous process (Permenter & Yuan, 2024; Karczewski et al., 2026; Sclocchi et al., 2025; Jin et al., 2026). One direction analyzes a symmetry-breaking transition, at which a trajectory initialized from isotropic noise breaks its symmetry and commits to a particular mode (Raya & Ambrogioni, 2023; Biroli et al., 2024; Li & Chen, 2024; Lobashev et al., 2026). For the ODE trajectories of flow-based models, prior work decomposes sampling into distinct dynamic stages (Wan et al., 2025) and uses velocity field divergence to measure local expansion and contraction along the trajectory (Min & Gomes, 2026). Existing analyses, however, measure tem-

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.

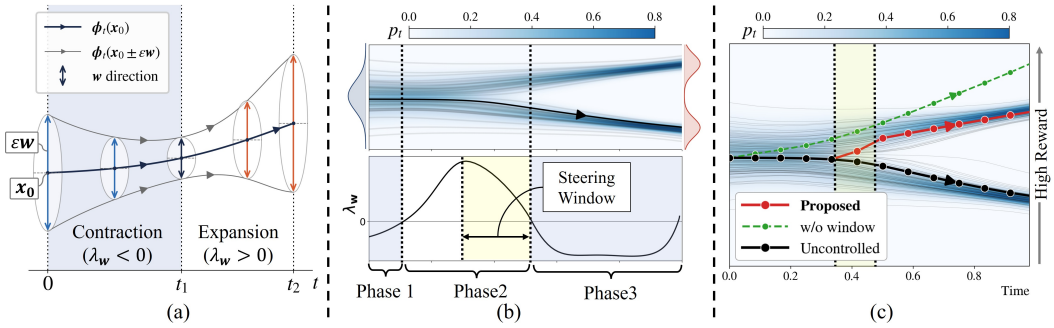

Figure 1: (a) **Directional Expansion Rate**: $\lambda_{\boldsymbol{w}}$ measures perturbation growth along $\boldsymbol{w}$, indicating whether the field ϕ_t contracts or expands along it. (b) **Phase Structure**: Three phases in generation and the proposed steering window (yellow) (c) **Proposed method**: Windowed guidance (red) successfully reroutes the trajectory away from uncontrolled one (black), whereas the guidance without the window (green) drives it off-distribution. Heatmaps show marginal density p_t at time $t \in [0, 1]$

poral inhomogeneity by aggregating over all directions (e.g., via divergence). Trajectory steering, by contrast, perturbs the state along the direction that most steeply increases the terminal reward. The dynamics along this particular direction cannot be inferred from such direction-agnostic quantities. Consequently, how the trajectory responds to directional perturbations across sampling stages remains unexplored. To address this gap, we analyze the phase structure of the sampling dynamics along the axis of intervention.

Contributions of this work. In this work, we introduce the *directional expansion rate* $\lambda_{\boldsymbol{w}}(\boldsymbol{x}, t)$ to characterize the generative dynamics of flow models along the reward-update direction $\boldsymbol{w}$. Inspired by classical deformation analysis in continuum mechanics (Batchelor, 2000; Gurtin et al., 2010; Ottino, 1989), the quantity $\lambda_{\boldsymbol{w}} = \boldsymbol{w}^\top \boldsymbol{J}(\boldsymbol{x}, t)\boldsymbol{w}$ measures local stretching rate in a given direction, where $\boldsymbol{J}(\boldsymbol{x}, t)$ is a jacobian of the field $\boldsymbol{\phi}_t(\boldsymbol{x})$ in Figure 1 (a). Specifically, its sign indicates whether the pretrained velocity field $\boldsymbol{v}_{\boldsymbol{\theta}}$ amplifies or attenuates perturbations along $\boldsymbol{w}$. By setting $\boldsymbol{w}$ as the guidance direction toward higher reward, $\lambda_{\boldsymbol{w}}$ measures how the underlying dynamics respond to the injected steering signal. Next, we present a consistent *contraction–expansion–contraction* pattern along the reward-update direction, as shown in Figure 1 (b). To explain this phenomenon, we analyze a minimal, analytically tractable model (i.e., a two-component Gaussian mixture) and derive closed-form expressions for the phase transition times. Our analysis connects this three-phase structure to the process of mode commitment. Specifically, the initial contraction draws trajectories toward the overall data mean, consistent with Wan et al. (2025). During expansion, trajectories separate and commit to individual modes. The final contraction then draws each trajectory toward its selected mode, refining the output. Building on the observed dynamics, we propose a steering window based on the phase-aware criterion which determines when the guidance should be applied. In this work, we instantiate this criterion using adjoint-based trajectory control, restricting intervention to the interval from the peak of $\lambda_{\boldsymbol{w}}$ to the end of the expansion phase, as shown in Figure 1 (c). For tasks requiring strict feasibility at the output, we apply an extra projection step during the final contraction phase to enforce hard constraints while preserving the selected mode. Experimentally, we show consistent improvements across tasks for scientific application such as data assimilation under observational constraints, and hard-constrained sampling for solving PDEs.

Our main contributions are summarized as follows:

- **Characterization of three-phase sampling dynamics.** We introduce the directional expansion rate $\lambda_{\boldsymbol{w}}(\boldsymbol{x}, t)$ and identify a contraction–expansion–contraction structure in flow sampling. Through a two-component Gaussian mixture model, we connect these phases to mode commitment and derive closed-form expressions for their transition times.
- **Dynamics-informed reward control.** We use the directional expansion rate to set an intervention window for adjoint-based control, complemented by projection during the final contraction phase for tasks with hard output constraints.
- **Consistent empirical improvements across tasks.** We demonstrate the effectiveness of the proposed strategy in data assimilation, and hard-constrained PDE sampling.

2 PRELIMINARIES & RELATED WORK

We consider flow-based generative models, which transport a Gaussian prior $\boldsymbol{x}_0 \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ to the data distribution $\boldsymbol{x}_1 \sim p_1$, following the notation of previous work (Lipman et al., 2023). Specifically, a flow matching model (Lipman et al., 2023) follows the generative trajectory by integrating an instantaneous velocity field $\boldsymbol{v}_\theta(\boldsymbol{x}_t)$. A complementary line of work learns a *flow map* $\Phi_{s\to t} : \mathbb{R}^d \to \mathbb{R}^d$ that moves a state directly from time s to time t, $\boldsymbol{x}_t = \Phi_{s\to t}(\boldsymbol{x}_s)$ (Huang et al., 2026; Geng et al., 2025). Both approaches describe the same underlying generative trajectory and differ only in how they traverse it: over an infinitesimally small interval, $\Phi_{t\to t+dt}$ reduces to a displacement along the instantaneous velocity $\boldsymbol{v}_\theta$.

To align these models with a reward r at inference time, recent work casts the problem as optimal control over the generative trajectory (Huang et al., 2026; Wang et al., 2025). A control signal $\boldsymbol{w}_t$ is added to the pretrained dynamics, which minimizes a cost effort while maximizing terminal reward,

$$\min_{\tilde{\boldsymbol{w}}} \int_0^1 \tfrac{1}{2}\|\tilde{\boldsymbol{w}}_t\|^2 \, dt \; - \; r(\tilde{\boldsymbol{w}}_1) \quad \text{s.t.} \quad \dot{\boldsymbol{x}}_t = \boldsymbol{v}_\theta(\boldsymbol{x}_t, t) + \tilde{\boldsymbol{w}}_t. \tag{1}$$

where $\boldsymbol{x}_1$ is the terminal state obtained by rolling out the controlled trajectory through the flow model. The optimality conditions determine $\tilde{\boldsymbol{w}}_t$ in closed form through the adjoint state, which is propagated backward from $\nabla r(\boldsymbol{x}_1)$. This answers *how* to guide the trajectory, but not *when*: the control acts over the entire interval, and its intervention time is left as a hyperparameter. In this paper, we show that the trajectory is temporally inhomogeneous with respect to $\tilde{\boldsymbol{w}}_t$, and that this phase structure is determined by the pretrained model itself. We consider two different scientific applications with distinct terminal objectives, ranging from enforcing observational consistency and hard physical constraints:

- **Data assimilation.** The goal of this task is to estimate the dynamic states such as turbulent flows or climate patterns, from incomplete observations (Chen et al., 2025). Specifically, we reconstruct the state $\boldsymbol{x}(\tau)$ at physical time τ from observations $\boldsymbol{y}_k = \mathcal{A}_k(\boldsymbol{x}^\star(\tau_k)) + \boldsymbol{\eta}_k$ collected at $\{\tau_k\}_{k=1}^K$, where $\boldsymbol{x}^\star(\tau)$ denote the underlying states, $\mathcal{A}_k$ the observation operator, and $\boldsymbol{\eta}_k$ measurement noise. To align the generation output $\boldsymbol{x}_1(\tau_k)$ with these observations, we maximize the reward $r(\boldsymbol{x}_1) = -\left(\sum_{k=1}^K \left\|\boldsymbol{y}_k - \mathcal{A}_k\big(\boldsymbol{x}_1(\tau_k)\big)\right\|_2^2\right)^{1/2}$.
- **PDE Solving under Hard Physical Constraints.** This task generates PDE solutions that strictly satisfy the given initial conditions and physical conservation laws (Utkarsh et al., 2025; Cheng et al., 2025). These constraints define a feasible set $\mathcal{C} = \{\boldsymbol{x} : h(\boldsymbol{x}) = \mathbf{0}\}$, where h measures the constraint residuals. Since a soft penalty alone cannot guarantee exact feasibility ($\boldsymbol{x}_1 \in \mathcal{C}$), we incorporate the hard constraints via the projection operator $\Pi_{\mathcal{C}}$, after steering the trajectory with the relaxed reward $r(\boldsymbol{x}_1) = -\|h(\boldsymbol{x}_1)\|_2$. See Appendix C.2 for PDE details and the constraints.

2.1 RELATED WORKS

Dynamic Phases of Generative Trajectory The generative trajectory of a diffusion or flow-based model shows distinct behavior along the sampling path, with early steps governed by mean-reverting dynamics and later steps refining high-frequency details. A recent line of work has explored this evolution through the lens of dynamical phase transitions, focusing on both the timing of the transition and the behavior within each phase. Specifically, prior work has analyzed the symmetry-breaking transition at which a trajectory commits to a mode at a critical time (Raya & Ambrogioni, 2023; Biroli et al., 2024). A related transition has also been identified in hierarchical data (Lobashev et al., 2026; Sclocchi et al., 2025), with non-asymptotic bounds on the resulting critical windows (Li & Chen, 2024). In flow-based models, Wan et al. (2025) decomposed the ODE into an early mean-attraction stage and a late absorbing stage, while Min & Gomes (2026) used the divergence of the velocity field to identify the trajectory dynamics throughout generation. These studies characterize how the generative dynamics unfold on its own, but not how it responds to an external reward signal injected along the way. In our work, we address this gap by characterizing the phase-dependent effects of reward guidance on the generative trajectory.

Reward-Aligned Generation Inference-time alignment steers a pretrained model toward outputs favored by an external criterion without additional training. The central difficulty is a mismatch of

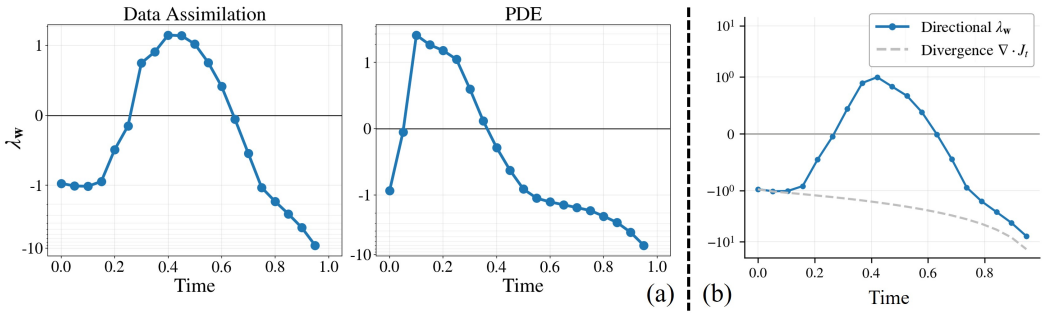

Figure 2: (a) Log-scaled directional expansion rate $\lambda_w(t)$ from $t = 0$ (noise) to $t = 1$ (data) across tasks, which exhibits a shared phase structure: Contraction–Expansion–Contraction pattern. (b) Comparison between the proposed $\lambda_{\boldsymbol{w}}$ and divergence $\nabla \cdot \boldsymbol{J}_t$ of the pretrained flow velocity

evaluation points: the reward scores terminal outputs, whereas steering must act on noisy intermediate states. This has motivated an optimization perspective in which generation is treated not as a fixed forward pass but as a trajectory to be optimized for terminal reward. One line of research optimizes the initial noise alone, searching for a seed whose trajectory attains high terminal reward (Ben-Hamu et al., 2024; Hwang et al., 2026; Eyring et al., 2024; Kim et al., 2026a). A second direction steers the trajectory throughout sampling by perturbing the per-step drift, using the reward gradient obtained from an adjoint state derived via Pontryagin's Maximum Principle (Wang et al., 2025; Chang et al., 2026). However, the timing of this steering remains underexplored, even though applying it uniformly over the whole trajectory is known to degrade sample quality (Huang et al., 2026). In our work, we analyze how the steering interacts with the generation phases, and use this analysis to determine the window over which it should be applied.

3 Directional Phase Transition

3.1 Key Observation

Directional Expansion Rate. Inspired by the classical deformation analysis in continuum mechanics (Batchelor, 2000; Gurtin et al., 2010), we examine the instantaneous directional expansion rate to analyze how a perturbation along a designated direction evolves locally under the velocity field $\dot{\boldsymbol{x}}_t = \boldsymbol{v}_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t)$ of a flow model with learned parameter $\boldsymbol{\theta}$. For an arbitrary state $\boldsymbol{x}$ and flow time t, let $\boldsymbol{J}(\boldsymbol{x}, t) = \partial_{\boldsymbol{x}} \boldsymbol{v}_{\boldsymbol{\theta}}(\boldsymbol{x}, t)$ denote the Jacobian of the velocity field, and $\boldsymbol{S}(\boldsymbol{x}, t) = 1/2 \left(\boldsymbol{J}(\boldsymbol{x}, t) + \boldsymbol{J}(\boldsymbol{x}, t)^\top \right)$ define its symmetric rate-of-strain tensor. For a unit direction $\boldsymbol{w}$, we define the instantaneous *directional expansion rate* as

$$\boxed{\lambda_{\boldsymbol{w}}(\boldsymbol{x}, t) := \boldsymbol{w}^\top \boldsymbol{S}(\boldsymbol{x}, t)\, \boldsymbol{w} = \boldsymbol{w}^\top \boldsymbol{J}(\boldsymbol{x}, t)\, \boldsymbol{w}} \tag{2}$$

where the equality follows because the antisymmetric part of $\boldsymbol{J}(\boldsymbol{x}, t)$ does not contribute to changes in perturbation magnitude. Along a sampling trajectory $\boldsymbol{x}_t$, we use the shorthand notations $\boldsymbol{J}_t := \boldsymbol{J}(\boldsymbol{x}_t, t)$. The sign of $\lambda_{\boldsymbol{w}}(\boldsymbol{x}_t, t) = \boldsymbol{w}^\top \boldsymbol{J}_t \boldsymbol{w}$ then characterizes the local behavior of perturbations along $\boldsymbol{w}$ on the trajectory: $\lambda_{\boldsymbol{w}}(\boldsymbol{x}_t, t) > 0$ indicates amplification along $\boldsymbol{w}$, while $\lambda_{\boldsymbol{w}}(\boldsymbol{x}_t, t) < 0$ indicates damping of the perturbation by the flow.

In reward-guided generation, the intervention to the generation trajectory is applied along the *reward direction* $\boldsymbol{w}_t := \tilde{\boldsymbol{w}}_t / \|\tilde{\boldsymbol{w}}_t\| = \nabla_{\boldsymbol{x}_t} r(\boldsymbol{x}_1) / \|\nabla_{\boldsymbol{x}_t} r(\boldsymbol{x}_1)\|$, obtained by transporting the terminal reward gradient backward via the adjoint ODE (Wang et al., 2025; Chang et al., 2026). We therefore evaluate the directional expansion rate along the sampling trajectory in the reward direction and, for brevity, write $\lambda_{\boldsymbol{w}}(t) := \lambda_{\boldsymbol{w}_t}(\boldsymbol{x}_t, t)$. While prior adjoint-based control methods treat guidance timing merely as a hand-tuned hyperparameter or leave it unaddressed, our direct analysis of $\lambda_{\boldsymbol{w}}(t)$ uncovers the generator's intrinsic phase structure. This transforms guidance timing from a manual choice into a model-derived property.

Observation of Phase Transition. Using the reward function defined in Section 2, we present the evolution of $\lambda_{\boldsymbol{w}}(t)$ across multiple flow models, including a reaction–diffusion PDE, a

Navier–Stokes data assimilation (DA) model. As shown in Figure 2 (a), the sign of $\lambda_{\boldsymbol{w}}(t)$ resolves sampling into three phases, following a *contraction–expansion–contraction* pattern. However, this is not captured by the derivative $\nabla \cdot \boldsymbol{J}_t$ which is direction-agnostic, remaining negative and fails to capture direction-dependent behavior(Figure 2 (b)). This gap directly motivates to explore directional behavior of generative process using the introduced $\lambda_{\boldsymbol{w}}(t)$. Intuitively, $\lambda_{\boldsymbol{w}}(t)$ indicates whether a reward-aligned perturbation injected at time t grows or decays along the remaining trajectory, and hence how much a control applied at t can change the terminal sample. Under this lens, we present the three phases with distinct roles:

- **Phase 1 (contraction).** The trajectory converges toward the mean of the data distribution, damping reward-aligned perturbations and leaving the terminal state insensitive to intervention.
- **Phase 2 (expansion).** The perturbations are amplified, so a control applied here redirects the terminal state at low cost; this is the window where reward guidance is effective.
- **Phase 3 (re-contraction).** The trajectory commits to a mode: inter-mode transitions become increasingly costly, while intra-mode refinement remains effective.

3.2 Theoretical Justification

Inspired by recent analyses of inhomogeneous generation characteristics in two-component Gaussians (Biroli et al., 2024; Kong et al., 2026), we leverage this setup as a minimal analytic model to explore the phase transitions driven by mode commitment. Crucially, despite its simplicity, this framework captures the core dynamics we observe in complex generative models. Specifically, we consider

$$p_1(\boldsymbol{x}) = \tfrac{1}{2}\mathcal{N}(\boldsymbol{\mu}, \sigma^2 \boldsymbol{I}) + \tfrac{1}{2}\mathcal{N}(-\boldsymbol{\mu}, \sigma^2 \boldsymbol{I}), \tag{3}$$

with symmetric means $\pm\boldsymbol{\mu} \in \mathbb{R}^d$ and a standard deviation σ. Next, we define the terminal reward. In practice, reward models reflect preferences between semantically distinct outputs. Under the setting of Equation (3), the semantically distinct outputs are the two modes, so such a preference reduces to a preference between modes. We model this via a reward function φ that increases along the mode axis $\boldsymbol{e}_{\boldsymbol{\mu}} := \boldsymbol{\mu}/\|\boldsymbol{\mu}\|$, which favors the $+\boldsymbol{\mu}$ mode:

$$r(\boldsymbol{x}) = \varphi(\boldsymbol{e}_{\boldsymbol{\mu}}^\top \boldsymbol{x}), \tag{4}$$

where φ is an arbitrary differentiable function with $\varphi' > 0$. From p_1 and r and noise level $\rho_t^2 := t^2\sigma^2 + (1-t)^2$, we derive $\lambda_{\boldsymbol{w}}$ in closed form, as stated in Proposition 3.1.

Proposition 3.1 (Directional expansion rate of the two-component Gaussian mixture)**.** *For the mixture in Equation* (3)*, the Jacobian of the velocity field is isotropic apart from a rank-one term along* $\boldsymbol{e}_{\boldsymbol{\mu}}$ *at every* $(\boldsymbol{x}_t, t)$*. Consequently:*

(i) the reward direction is fixed, $\boldsymbol{w}_t = \boldsymbol{e}_{\boldsymbol{\mu}}$ *for all* $t \in [0, 1)$*;*

(ii) the directional expansion rate is the corresponding eigenvalue, for all $t \in [0, 1)$ *we have*

$$\lambda_{\boldsymbol{w}}(t) = \frac{t\sigma^2 - (1-t)}{\rho_t^2} + \frac{\|\boldsymbol{\mu}\|^2\, t(1-t)}{\rho_t^4}\, \operatorname{sech}^2\Big(\frac{t\,\boldsymbol{\mu}^\top \boldsymbol{x}_t}{\rho_t^2}\Big). \tag{5}$$

Derivations and proofs are provided in Appendix A. The two terms of Equation (5) have distinct origins and dependencies. The first term is isotropic and independent of the state $\boldsymbol{x}_t$, contracting for most of the interval and turning positive only near the end. In contrast, the second term depends on the state and contributes to mode selection. It is non-negative everywhere and separates neighboring trajectories along the mode axis. However, it decays once a trajectory has moved away from the decision boundary, so the expansion ends as soon as the mode is settled. Their interaction drives the three sign changes of $\lambda_{\boldsymbol{w}}$ that form the phase transitions. The early stage is dominated by contraction; the mode-selection term then surpasses it, making $\lambda_{\boldsymbol{w}} > 0$ and opening an expanding phase; after mode commitment, the term decays, restoring contraction until the isotropic term reverses sign. We denote the three transitions t_e, t_c and t_σ in order of occurrence and derive them below. Derivations for each time are given in Appendix B.

Expansion onset t_e**.** The expansion region first opens on the hyperplane $\{\boldsymbol{x}_t : \boldsymbol{x}_t^\top \boldsymbol{\mu} = 0\}$, where the mode is undetermined. In this hyperplane, the transition time t_e is derived as follows:

$$t_e = \frac{2}{2 + \|\boldsymbol{\mu}\|^2 + \|\boldsymbol{\mu}\|\sqrt{\|\boldsymbol{\mu}\|^2 + 4}}, \tag{6}$$

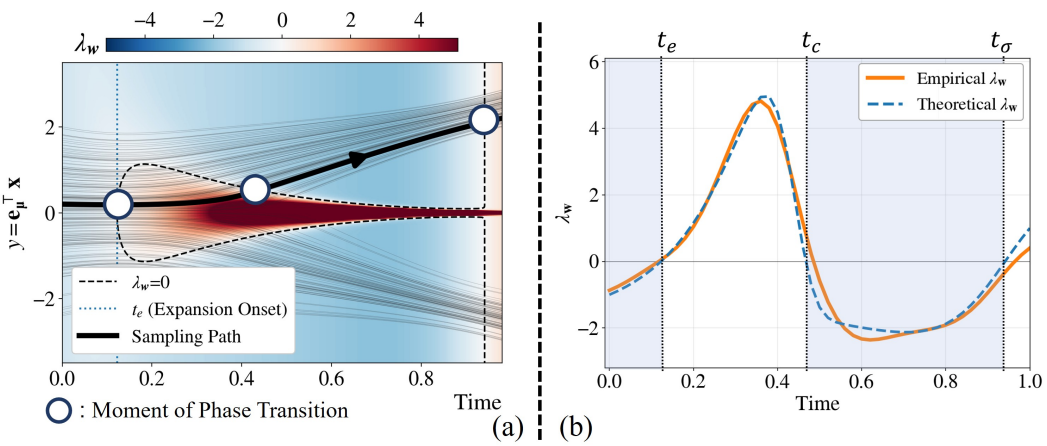

Figure 3: Phase structure measured by the directional expansion rate $\lambda_{\boldsymbol{w}}$. (a) Heatmap of theoretical $\lambda_{\boldsymbol{w}}$ over (y, t) plane. The dashed curve is the zero-crossing of $\lambda_{\boldsymbol{w}}$, the transition boundary between phases. (b) $\lambda_{\boldsymbol{w}}$ along the selected sampling path in (a), giving the phase transitions at t_e, t_c and t_σ.

which lower-bounds the beginning of the expansion phase for every trajectory. Additionally, in the well-separated regime where $\|\boldsymbol{\mu}\| \gg 1$, $t_e \approx 1/\|\boldsymbol{\mu}\|^2$. This provides a clear intuition: greater mode separation induces an earlier expansion stage. This transition is the flow matching counterpart of the symmetry-breaking time studied in diffusion models (Raya & Ambrogioni, 2023), where the symmetric fixed point of the reverse dynamics loses stability.

Contraction time t_c. The second transition varies across trajectories. Once the trajectory enters the expansion region after t_e, it begins to commit to a mode, and the mode-selection term decays as $|\boldsymbol{\mu}^\top \boldsymbol{x}_1|$ grows. This transition boundary $\mathcal{B}$ is the curve

$$\mathcal{B} := \{(\boldsymbol{x}, t) : f(\boldsymbol{x}, t) = 0\} \quad \text{where} \quad f(\boldsymbol{x}, t) = |\boldsymbol{e}_{\boldsymbol{\mu}}^\top \boldsymbol{x}| - \frac{\rho_t^2}{t\|\boldsymbol{\mu}\|} \cosh^{-1}\sqrt{\frac{\|\boldsymbol{\mu}\|^2\, t(1-t)}{\rho_t^2\,(1-t-t\sigma^2)}}, \quad (7)$$

and we define the contraction time t_c as the moment the trajectory crosses $\mathcal{B}$,

$$t_c(\boldsymbol{x}_0) := \min\{\, t > t_e :\ f(\boldsymbol{x}_t, t) = 0,\ \tfrac{d}{dt} f(\boldsymbol{x}_t, t) > 0 \,\}. \quad (8)$$

The transition changes the sign of $\lambda_{\boldsymbol{w}}$ from positive to negative, as visualized in Figure 3. This crossing is the decisive point in generative process. Before it, expansion amplifies small differences along $\boldsymbol{e}_{\boldsymbol{\mu}}$, so the mode assignment is still being decided. After the crossing, contraction pulls neighboring trajectories back together, refining detail within a mode that has already been chosen.

The analysis also predicts a final re-expansion phase at the end , as $\lambda_{\boldsymbol{w}}$ reduces to the isotropic term after the mode-selection term is exponentially suppressed. Specifically, $\lambda_{\boldsymbol{w}}$ becomes positive at

$$t_\sigma = \frac{1}{1+\sigma^2}. \quad (9)$$

Since the final phase interval $\sigma^2/(1+\sigma^2)$ lies below the discrete time-step resolution ($t_\sigma \approx 1$), we treat t_σ as a purely theoretical transition time.

Pre-committed trajectories. If $|\boldsymbol{e}_{\boldsymbol{\mu}}^\top \boldsymbol{x}_t|$ is already large at initialization ($t = 0$), $\lambda_{\boldsymbol{w}}$ stays negative before t_σ, as the mode-selection term never overcomes the isotropic contraction. Such trajectories start from a tail of p_0, and the resulting samples lie in the corresponding tail of p_1. Accordingly, the intervention near t_e carries a risk of drifting the sample away from the typical set, which motivates opening the steering window at $t_{\text{peak}} = \arg\max_t \lambda_{\boldsymbol{w}}$ rather than at t_e (Section 3.3).

Toy experiment. We analyze the phase structure along $y := \boldsymbol{e}_{\boldsymbol{\mu}}^\top \boldsymbol{x}$ in 8-dimensional gaussian mixture, as shown in Figure 3. By Proposition 3.1, $\lambda_{\boldsymbol{w}}$ depends solely on the state through y, allowing the phase transition to be captured in the (y, t) plane. Directions orthogonal to $\boldsymbol{e}_{\boldsymbol{\mu}}$ are governed only the isotropic term and undergo no transition induced by the mode separation, which is out of our scope. Further experimental details are provided in Appendix C.1. Fig. 3(a) shows

Algorithm 1 Reward Aligned Generation using Phase-aware Window

Require: Model $\boldsymbol{v}_\theta$, Reward $r(\cdot)$, Time grid $\{t_l\}_{l=1}^{N_t}$, Initial parameters $\{t_c, t_{peak}, \boldsymbol{x}_{peak}\}$, Update interval n_c, (Optional) Hard Constraint $\mathcal{C}$

for $n = 1, \ldots, N_{\text{iter}}$ **do**
  $\{\lambda_{\boldsymbol{w}}(t), \tilde{\boldsymbol{w}}_t\}_{t=t_{peak}}^{1} \leftarrow \text{Adjoint}(\{\boldsymbol{x}_t\}_{t=t_{peak}}^{1}, r(\cdot))$
  if $n \bmod n_c = 0$: $t_c \leftarrow \min\{\, t > t_{\text{peak}} : \lambda_{\boldsymbol{w}}(t) \le 0 \,\}$ ▷ Update t_c every n_c steps
  for $t = t_{peak}, \ldots, t_{N_t - 1}$ **do**
    $\boldsymbol{x}_{t+\Delta t} \leftarrow \boldsymbol{x}_t + \Delta t\, \boldsymbol{v}_\theta(\boldsymbol{x}_t, t)$
    if $t_{peak} \le t < t_c$ **then**
      $\boldsymbol{x}_{t+\Delta t} \leftarrow \boldsymbol{x}_{t+\Delta t} + \tilde{\boldsymbol{w}}_t$ ▷ Intervention in mode-selection
    else if $t \ge t_c$ **and** $\mathcal{C}$ given **then**
      $\boldsymbol{x}_{t+\Delta t} \leftarrow \Pi_{\mathcal{C}}(\boldsymbol{x}_{t+\Delta t})$ ▷ Projection for hard constraint
return $\boldsymbol{x}_{t_{N_t}} = \boldsymbol{x}_1$

the zero-crossing of $\lambda_{\boldsymbol{w}}$ by the dashed curve, the transition boundary $\mathcal{B}$ in the (y, t) plane. We explicitly mark t_e where $\mathcal{B}$ meets the axis $y = 0$, which refers to the onset time for the expansion region. We then select a single trajectory to examine how $\lambda_{\boldsymbol{w}}$ evolves along an actual generation path, as shown in Fig. 3(b). Three transitions are visible: t_e, where expansion begins; t_c, where mode commitment restores contraction; and t_σ, where expansion returns at the terminal. The estimated transition times agree closely with the closed-form predictions; specifically, the estimated values $(\hat{t}_e, \hat{t}_c, \hat{t}_\sigma) = (0.11,\, 0.53,\, 0.97)$ well aligned with theoretical values $(t_e, t_c, t_\sigma) = (0.12,\, 0.51,\, 0.94)$.

3.3 METHODS

We now use this phase structure to determine *when* the intervention should be applied. We estimate $\lambda_{\boldsymbol{w}}$ along trajectories of the pretrained model, detect its transitions, and take the interval they bound as a guidance window. In the experiments below, we instantiate it using adjoint-based trajectory control (Wang et al., 2025). The overall method is shown in Algorithm 1.

Step 1. Measuring $\lambda_{\boldsymbol{w}}$ and detecting phases. We measure $\lambda_{\boldsymbol{w}}(t) = \boldsymbol{w}_t^\top J_t \boldsymbol{w}_t$ along an uncontrolled trajectory integrated on a coarse grid of N_t timesteps. From the measured $\lambda_{\boldsymbol{w}}$ we obtain the peak expansion time and the contraction time,

$$t_{\text{peak}} := \arg\max_t \lambda_{\boldsymbol{w}}(t), \quad t_c = \min\{\, t > t_{\text{peak}} \;:\; \lambda_{\boldsymbol{w}}(t) \le 0 \,\}. \tag{10}$$

We utilize $\boldsymbol{x}_{peak}$, unguided state at t_{peak}, as a starting point for reward-aligned trajectory.

Step 2. Control with Phase-aware Window. The detected times define the interval over which reward guidance acts: the trajectory control $\tilde{\boldsymbol{w}}_t$ at time step t is applied only on $[t_{\text{peak}}, t_c)$. Guidance thus acts only while the trajectory is expanding along $\boldsymbol{w}_t$, where a perturbation can still route it to a different mode. Since t_c is trajectory-dependent, we re-evaluate $\lambda_{\boldsymbol{w}}$ every k steps on the trajectory being controlled and close the window at the first non-positive value, so that each trajectory is cut at its own commitment time.

Step 3. Late-phase projection for hard constraints. When the objective includes a hard constraint $\boldsymbol{x}_1 \in \mathcal{C}$, stopping the intervention at t_c is insufficient, since the constraint must be satisfied exactly. We therefore apply projection $\Pi_{\mathcal{C}}$ during the contraction phase $[t_c, 1]$. Specifically, after t_c, the trajectory commits to a mode, shifting the goal from inter-mode routing to within-mode refinement. Unlike soft penalties—which weigh the reward together with the prior and risk reward hacking when over-optimized—hard constraints must hold exactly, leaving the prior to select only among feasible points. Accordingly, projection onto the nearest feasible point yields a correction bounded by the current violation while guaranteeing feasibility.

Table 1: Prediction accuracy and hard-constraint satisfaction across three PDE benchmarks. Metrics are scaled by 10^{-2} except FPD. Constraint violations below 10^{-5} are reported as 0.0. Best accuracy in **bold**

Dataset	Method	Prediction Accuracy			Hard Constraint Satisfaction	
		MMSE	SMSE	FPD↓	IC↓	CL↓
Heat Equation	D-Flow	3.60±0.51	**0.20**±0.02	0.34±0.10	6.86±0.35	1.12±0.26
	OC-Flow	3.00±0.13	0.40±0.04	0.36±0.04	15.76±0.29	0.05±0.00
	ECI	3.00±0.13	0.68±0.05	1.61±0.14	0.0	2.47±0.04
	PCFM	3.09±0.23	0.69±0.15	1.43±0.25	0.0	0.0
	Proposed	**0.98**±0.05	**0.20**±0.04	**0.26**±0.06	0.0	0.0
Burgers IC	D-Flow	6.63±0.64	24.78±19.98	2.25±0.46	17.03±0.87	1.15±0.25
	OC-Flow	2.68±0.13	3.72±0.23	3.77±0.35	9.03±0.38	0.88±0.03
	ECI	5.06±0.30	5.33±0.33	4.87±0.25	0.0	11.55±0.85
	PCFM	1.12±0.03	0.82±0.01	2.92±0.29	0.0	0.0
	Proposed	**0.67**±0.07	**0.59**±0.09	**1.31**±0.25	0.0	0.0
Reaction-Diffusion	D-Flow	2.00±0.34	12.15±11.50	32.00±4.30	21.37±1.67	1.04±0.24
	OC-Flow	1.31±0.07	0.28±0.02	14.85±0.92	1.72±0.01	0.21±0.01
	ECI	1.54±0.01	**0.20**±0.03	29.44±0.47	0.0	2.24±0.03
	PCFM	0.53±0.04	0.50±0.03	11.77±1.63	0.0	0.0
	Proposed	**0.39**±0.03	0.28±0.01	**11.02**±1.93	0.0	0.0

4 EXPERIMENTS

4.1 SOLVING PDES WITH PHYSICAL HARD CONSTRAINTS

Settings. We evaluate on the three PDE systems governed by hard constraints, the initial condition (IC) and physical conservation law (CL). Following the evaluation protocols of the prior works (Cheng et al., 2025; Utkarsh et al., 2025), we measure mean squared errors of point-wise predictions (MMSE), standard deviation (SMSE), constraint error (CE) and Fréchet Poseidon distance (FPD) (Cheng et al., 2025) using 512 generated samples. We compare our method with training-free baselines, including trajectory optimization methods, OC-Flow (Wang et al., 2025) and D-Flow (Ben-Hamu et al., 2024), as well as projection-based hard constraint solvers including ECI-Sampling(Cheng et al., 2025) and PCFM(Utkarsh et al., 2025). For this task, our method follows the Alg.1 with hard constraint $\mathcal{C}$ as IC and CL. Further experimental details are provided in Appendix C.2.

Results. Table 1 presents constraint violations and solution errors on the three PDE benchmarks. Soft-penalty solvers (D-Flow, OC-Flow) produce plausible outputs with reduced prediction errors and constraint residuals, but do not exactly satisfy the constraints showing nonzero errors on both sides. ECI satisfies the IC via projection yet leaves the CL unconstrained, so its samples violate the physical conservation law. Both PCFM and our method achieve exact satisfaction via projecting samples on both constraints. However, the main difference comes from how they form the generation path: PCFM projects the sample at every time step, where the strict projection distorts the prior (Li et al., 2026; Yang et al., 2026). In contrast, our method routes the trajectory toward reward-consistent modes prior to projection, which yields better prediction accuracy.

4.2 DATA ASSIMILATION

Settings. We consider two dynamic systems for the experiment. The first is the 2D incompressible Navier–Stokes (NS) flow with random forcing (Chen et al., 2024), whose state is the 128×128 vorticity field, conditioned on 10 previous states. We evaluate super-resolution (SR), which observes an $8\times$ downsampled 16×16 field, and sparse observation (SO), which observes 5% of the pixels, both with Gaussian noise of standard deviation 0.05. The second is weather forecasting on the vertically integrated liquid (VIL) product of SEVIR (Veillette et al., 2020), a real-world radar dataset with 128×128 frames recorded every 10 minutes, conditioned on 6 previous states following (Gao et al., 2023), where 10% of the grid cells are observed. We compare against FlowDAS, which applies posterior-sampling guidance at every sampling step, and the trajectory optimization methods OC-

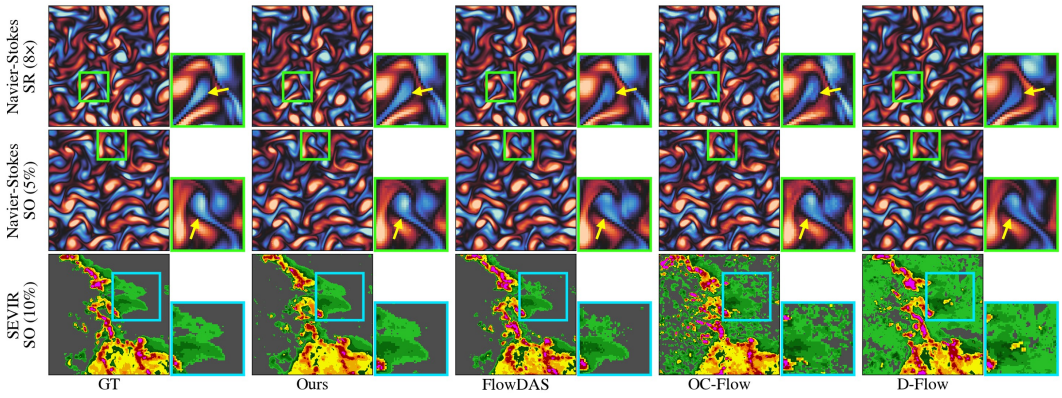

Figure 4: Qualitative comparison of data assimilation tasks. We present outputs at the 3rd rollout prediction step, with zoomed patches highlights the deviation of baselines. For each task, we show the trajectory with median error reduction over FlowDAS.

Table 2: Next-frame data assimilation results . Best results are in **bold**.

	Navier–Stokes (RMSE↓)		SEVIR, SO 10%		
Method	SR 8×	SO 5%	RMSE↓	$CSI_{0.3}$ ↑	$CSI_{0.5}$ ↑
FlowDAS	0.188±0.025	0.160±0.020	0.063±0.012	0.819±0.055	0.715±0.047
OC-Flow	0.247±0.032	0.179±0.018	0.074±0.016	0.778±0.061	0.661±0.053
D-Flow	0.229±0.018	0.216±0.030	0.070±0.013	0.787±0.063	0.656±0.053
Ours	**0.169**±0.028	**0.141**±0.012	**0.060**±0.010	**0.823**±0.054	**0.717**±0.050

Flow (Wang et al., 2025) and D-Flow (Ben-Hamu et al., 2024). Following FlowDAS (Chen et al., 2025), we report the root mean squared error (RMSE) of the next-frame estimate, together with the critical success index (CSI) at the VIL thresholds 0.3 and 0.5 for SEVIR, over 8 held-out test trajectories per task. Further experimental details are provided in Appendix C.3.

Results. Table 2 summarizes the next-frame results. Our method achieves the lowest RMSE on all three tasks and attains the highest CSI in SEVIR. Specifically, compared to OC-Flow and D-Flow, which steer the entire trajectory and only the initial noise respectively, our approach achieves superior results, demonstrating that restricting intervention to the measured expansion window is more effective. Moreover, whereas FlowDAS adjusts the generation path over a fine-grained time grid, our method applies control only within the expansion window on coarse-grained grid, which further shows the efficacy of our targeted intervention. Qualitative comparison in Fig. 4 supports the claim. We visualize 3rd rolled-out predictions which shows the least distortion of our method with the ground truth. We also present the quantitative results for 3rd roll-out Appendix C.3.2.

5 Conclusion

In this work, we introduce the directional expansion rate $\lambda_{\boldsymbol{w}}$ to characterize the phase structure of flow-based sampling along the reward direction. By measuring $\lambda_{\boldsymbol{w}}$ on large-scale pretrained models, we empirically observed a three-phase structure, which we then theoretically grounded through closed-form analysis on a two-component Gaussian mixture. Our theoretical derivations yielded exact transition times (t_e, t_c, t_σ) bounding a mode-selecting expansion window where interventions can effectively alter mode assignment, followed by a contraction phase where mode selection is already settled. Building on these insights, we proposed a phase-aware control strategy that confines reward interventions strictly to the expansion window, complemented by late-stage projections for hard-constrained tasks. Evaluations across data assimilation, and PDE-constrained sampling demonstrate that restricting guidance to the mode-selection phase maximizes terminal reward while preventing quality degradation from full-trajectory interventions, highlighting the key role of the underlying phase structure in informing the timing of intervention.

AI USE STATEMENT

(This section is **required** and does not count toward the page limit.)

In this work, we used generative AI tools for editing and polishing the text. We have not used generative AI tools for primary dataset creation, and data collection are not applicable to this work. Additionally, we used generative AI tools for debugging and optimizing baseline code, refining mathematical formulations. We have reviewed all AI-assisted work. All LLM-suggested text revisions were manually reviewed and rewritten by the authors to preserve the original logical flow, and AI-generated code was thoroughly verified and tested for correctness. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REFERENCES

G. K. Batchelor. *An Introduction to Fluid Dynamics*. Cambridge Mathematical Library. Cambridge University Press, 2000.

Heli Ben-Hamu, Omri Puny, Itai Gat, Brian Karrer, Uriel Singer, and Yaron Lipman. D-flow: Differentiating through flows for controlled generation. In *Forty-first International Conference on Machine Learning*, 2024. URL https://openreview.net/forum?id=SE20BFqj6J.

Giulio Biroli, Tony Bonnaire, Valentin De Bortoli, and Marc Mézard. Dynamical regimes of diffusion models. *Nature Communications*, 15(1):9957, 2024.

Kevin Black, Michael Janner, Yilun Du, Ilya Kostrikov, and Sergey Levine. Training diffusion models with reinforcement learning. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=YCWjhGrJFD.

Jinho Chang, Jaemin Kim, and Jong Chul Ye. Training-free reward-guided image editing via trajectory optimal control. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=YL2NgqN3Vh.

Siyi Chen, Yixuan Jia, Qing Qu, He Sun, and Jeffrey A Fessler. FlowDAS: A stochastic interpolant-based framework for data assimilation. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=1nWqhiulqD.

Yifan Chen, Mark Goldstein, Mengjian Hua, Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Probabilistic forecasting with stochastic interpolants and Föllmer processes. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, 2024.

Chaoran Cheng, Boran Han, Danielle C. Maddix, Abdul Fatir Ansari, Andrew Stuart, Michael W. Mahoney, and Bernie Wang. Gradient-free generation for hard-constrained systems. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=teE4pl9ftK.

Luca Eyring, Shyamgopal Karthik, Karsten Roth, Alexey Dosovitskiy, and Zeynep Akata. ReNO: Enhancing one-step text-to-image models through reward-based noise optimization. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024. URL https://openreview.net/forum?id=MXY0qsGgeO.

Ying Fan, Olivia Watkins, Yuqing Du, Hao Liu, Moonkyung Ryu, Craig Boutilier, Pieter Abbeel, Mohammad Ghavamzadeh, Kangwook Lee, and Kimin Lee. Reinforcement learning for fine-tuning text-to-image diffusion models. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. URL https://openreview.net/forum?id=8OTPepXzeh.

Zhihan Gao, Xingjian Shi, Boran Han, Hao Wang, Xiaoyong Jin, Danielle Maddix, Yi Zhu, Mu Li, and Yuyang Wang. PreDiff: Precipitation nowcasting with latent diffusion models. In *Advances in Neural Information Processing Systems*, volume 36, pp. 78621–78656, 2023.

Zhengyang Geng, Mingyang Deng, Xingjian Bai, J Zico Kolter, and Kaiming He. Mean flows for one-step generative modeling. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=uWj4s7rMnR.

Morton E Gurtin, Eliot Fried, and Lallit Anand. *The mechanics and thermodynamics of continua*. Cambridge university press, 2010.

Jerry Yuyang Huang, Justin Lin, Sheel Shah, Kartik Nair, and Nicholas Matthew Boffi. How to guide your flow: Few-step alignment via flow map reward guidance. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=gEogY6cPrL.

Jisung Hwang, Yunhong Min, Jaihoon Kim, I Shen, Minhyuk Sung, et al. Noisetilt: Noise-tilted reverse kernels for diffusion reward alignment. *arXiv preprint arXiv:2606.18066*, 2026.

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

Yang Jin, Zhicheng Sun, Ningyuan Li, Kun Xu, Kun Xu, Hao Jiang, Nan Zhuang, Quzhe Huang, Yang Song, Yadong MU, and Zhouchen Lin. Pyramidal flow matching for efficient video generative modeling. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=66NzcRQuOq.

Rafal Karczewski, Markus Heinonen, Alison Pouplin, Søren Hauberg, and Vikas K Garg. The spacetime of diffusion models: An information geometry perspective. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=qCsbYJZRA5.

Jeongsol Kim, Hongeun Kim, Jian Wang, and Jong Chul Ye. Gradient-free noise optimization for reward alignment in generative models. *arXiv preprint arXiv:2605.11347*, 2026a.

Yeongmin Kim, Donghyeok Shin, Byeonghu Na, Minsang Park, Richard Lee Kim, and Il chul Moon. Lookahead sample reward guidance for test-time scaling of diffusion models. In *Forty-third International Conference on Machine Learning*, 2026b. URL https://openreview.net/forum?id=IbRm6gwmew.

Seo Taek Kong, Weina Wang, and R Srikant. Noise schedule design for diffusion models: An optimal control perspective. *arXiv preprint arXiv:2605.21911*, 2026.

Tuomas Kynkäänniemi, Miika Aittala, Tero Karras, Samuli Laine, Timo Aila, and Jaakko Lehtinen. Applying guidance in a limited interval improves sample and distribution quality in diffusion models. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 37, 2024. URL https://openreview.net/forum?id=nAIhvNy15T.

Black Forest Labs, Stephen Batifol, Andreas Blattmann, Frederic Boesel, Saksham Consul, Cyril Diagne, Tim Dockhorn, Jack English, Zion English, Patrick Esser, Sumith Kulal, Kyle Lacey, Yam Levi, Cheng Li, Dominik Lorenz, Jonas Müller, Dustin Podell, Robin Rombach, Harry Saini, Axel Sauer, and Luke Smith. Flux.1 kontext: Flow matching for in-context image generation and editing in latent space, 2025. URL https://arxiv.org/abs/2506.15742.

Marvin Li and Sitan Chen. Critical windows: non-asymptotic theory for feature emergence in diffusion models. In *Forty-first International Conference on Machine Learning*, 2024. URL https://openreview.net/forum?id=a8ZpjLJuKk.

Zeyang Li, Kaveh Alim, and Navid Azizan. Hardflow: Hard-constrained sampling for flow-matching models via trajectory optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2026.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=PqvMRDCJT9t.

Luping Liu, Chao Du, Tianyu Pang, Zehan Wang, Chongxuan Li, and Dong Xu. Improving long-text alignment for text-to-image diffusion models. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=2ZK8zyIt7o.

Alexander Lobashev, Dmitry Guskov, Victor Kawasaki-Borruat, and Maria Larchenko. Caffarelli regularity and hierarchical phase boundaries in diffusion models, 2026. URL https://openreview.net/forum?id=HadqLI0x1V.

Zeqian Long, Mingzhe Zheng, Kunyu Feng, Xinhua Zhang, Hongyu Liu, Harry Yang, Linfeng Zhang, Qifeng Chen, and Yue Ma. Follow-your-shape: Shape-aware image editing via trajectory-guided region control. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=uGaR7L3Z1E.

Yuhang Ma, Xiaoshi Wu, Keqiang Sun, and Hongsheng Li. Hpsv3: Towards wide-spectrum human preference score. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 15086–15095, October 2025.

Yimeng Min and Carla P Gomes. Divergence-suppressing couplings for rectified flow. *arXiv preprint arXiv:2605.17733*, 2026.

Julio M. Ottino. *The Kinematics of Mixing: Stretching, Chaos, and Transport*. Cambridge University Press, 1989.

Frank Permenter and Chenyang Yuan. Interpreting and improving diffusion models from an optimization perspective. In *Forty-first International Conference on Machine Learning*, 2024. URL https://openreview.net/forum?id=o2ND9v0CeK.

Gabriel Raya and Luca Ambrogioni. Spontaneous symmetry breaking in generative diffusion models. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. URL https://openreview.net/forum?id=lxGFGMMSVl.

Antonio Sclocchi, Alessandro Favero, Noam Itzhak Levi, and Matthieu Wyart. Probing the latent hierarchical structure of data via diffusion models. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=0GzqVqCKns.

Utkarsh Utkarsh, Pengfei Cai, Alan Edelman, Rafael Gomez-Bombarelli, and Christopher Vincent Rackauckas. Physics-constrained flow matching: Sampling generative models with hard constraints. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=cf4etwjY7n.

Mark Veillette, Siddharth Samsi, and Chris Mattioli. SEVIR: A storm event imagery dataset for deep learning applications in radar and satellite meteorology. In *Advances in Neural Information Processing Systems*, volume 33, pp. 22009–22019, 2020.

Zhengchao Wan, Qingsong Wang, Gal Mishne, and Yusu Wang. Elucidating flow matching ODE dynamics via data geometry and denoisers. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=f5czhqYK3H.

Luran Wang, Chaoran Cheng, Yizhen Liao, Yanru Qu, and Ge Liu. Training free guided flow-matching with optimal control. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=61ss5RA1MM.

Jiazheng Xu, Xiao Liu, Yuchen Wu, Yuxuan Tong, Qinkai Li, Ming Ding, Jie Tang, and Yuxiao Dong. Imagereward: Learning and evaluating human preferences for text-to-image generation. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. URL https://openreview.net/forum?id=JVzeOYEx6d.

Jeongyong Yang, Seunghwan Jang, and SooJean Han. Safeflowmatcher: Safe and fast planning using flow matching with control barrier functions. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=refcXHU1Nh.

A PROOF OF PROPOSITION 3.1

Proof. The proof proceeds in four steps. We first derive the velocity field in closed form (Step 1). Differentiating this expression shows that the Jacobian is the sum of an isotropic term and a rank-one term along $\mathbf{e}_{\boldsymbol{\mu}}$ (Step 2). Then, we show that reward gradient aligns with $\mathbf{e}_{\boldsymbol{\mu}}$, which proves claim (i) (Step 3). Claim (ii) then follows immediately with $\mathbf{u_t} = \mathbf{e}_{\boldsymbol{\mu}}$ (Step 4).

STEP 1: CLOSED FORM OF THE VELOCITY FIELD

Posterior mean. Conditioned on the component label $s \in \{+, -\}$, we have $\boldsymbol{x}_1 = s\boldsymbol{\mu} + \sigma\boldsymbol{\xi}$ with $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$. Hence $\boldsymbol{x}_t = t(s\boldsymbol{\mu}+\sigma\boldsymbol{\xi})+(1-t)\boldsymbol{x}_0$ becomes Gaussian with mean $st\boldsymbol{\mu}$ and covariance $\big[(1-t)^2 + t^2\sigma^2\big]\boldsymbol{I}$. The marginal is therefore again a symmetric mixture,

$$\boldsymbol{x}_t \sim \tfrac{1}{2}\mathcal{N}(+t\boldsymbol{\mu}, \rho_t^2\boldsymbol{I}) + \tfrac{1}{2}\mathcal{N}(-t\boldsymbol{\mu}, \rho_t^2\boldsymbol{I}), \qquad \rho_t^2 := (1-t)^2 + t^2\sigma^2. \tag{11}$$

By Bayes' rule,

$$\ln \frac{p(+ \mid \boldsymbol{x}_t)}{p(- \mid \boldsymbol{x}_t)} = \ln \frac{\mathcal{N}(\boldsymbol{x}_t; +t\boldsymbol{\mu}, \rho_t^2\boldsymbol{I})}{\mathcal{N}(\boldsymbol{x}_t; -t\boldsymbol{\mu}, \rho_t^2\boldsymbol{I})} = \frac{\|\boldsymbol{x}_t + t\boldsymbol{\mu}\|^2 - \|\boldsymbol{x}_t - t\boldsymbol{\mu}\|^2}{2\rho_t^2} = \frac{2t\,\boldsymbol{\mu}^\top \boldsymbol{x}_t}{\rho_t^2}, \tag{12}$$

so that the posterior mode balance is

$$m(\boldsymbol{x}, t) := p(+ \mid \boldsymbol{x}_t = \boldsymbol{x}) - p(- \mid \boldsymbol{x}_t = \boldsymbol{x}) = \tanh\big(z(\boldsymbol{x}, t)\big) \quad \text{where} \quad z(\boldsymbol{x}, t) := \frac{t\,\boldsymbol{\mu}^\top \boldsymbol{x}}{\rho_t^2}. \tag{13}$$

Then, the conditional mean for given mode s becomes

$$\mathbb{E}[\boldsymbol{x}_1 \mid \boldsymbol{x}_t = \boldsymbol{x},\ s] = s\boldsymbol{\mu} + \frac{t\sigma^2}{\rho_t^2}\,(\boldsymbol{x} - st\boldsymbol{\mu}) = \frac{t\sigma^2}{\rho_t^2}\,\boldsymbol{x} + s\,\frac{(1-t)^2}{\rho_t^2}\,\boldsymbol{\mu}, \tag{14}$$

Finally, averaging over the posterior equation 13 gives,

$$\hat{\boldsymbol{x}}_1(\boldsymbol{x}, t) = \mathbb{E}[\boldsymbol{x}_1 \mid \boldsymbol{x}_t = \boldsymbol{x}] = \frac{t\sigma^2}{\rho_t^2}\,\boldsymbol{x} + \frac{(1-t)^2}{\rho_t^2}\,\tanh\big(z(\boldsymbol{x}, t)\big)\,\boldsymbol{\mu}. \tag{15}$$

Velocity field. For the linear interpolant, $x_0 = (\boldsymbol{x}_t - t\boldsymbol{x}_1)/(1-t)$ implies $\boldsymbol{x}_1 - \boldsymbol{x}_0 = (\boldsymbol{x}_1 - \boldsymbol{x}_t)/(1-t)$, so the marginal velocity is $\boldsymbol{v}(\boldsymbol{x}, t) = \mathbb{E}[\boldsymbol{x}_1 - \boldsymbol{x}_0 \mid \boldsymbol{x}_t = \boldsymbol{x}] = (\hat{\boldsymbol{x}}_1(\boldsymbol{x}, t) - \boldsymbol{x})/(1-t)$. Substituting equation 15 and simplifying with $t\sigma^2 - \rho_t^2 = (1-t)\big(t\sigma^2 - (1-t)\big)$ gives the decomposition

$$\boldsymbol{v}(\boldsymbol{x}, t) = \frac{t\sigma^2 - (1-t)}{\rho_t^2}\,\boldsymbol{x} + \frac{1-t}{\rho_t^2}\,\tanh\big(z(\boldsymbol{x}, t)\big)\,\boldsymbol{\mu}. \tag{16}$$

The first term is an isotropic linear drift whose scalar coefficient depends only on time, while the second acts only along $\boldsymbol{\mu}$ and drives the selection between the two modes.

STEP 2: RANK-ONE ANISOTROPY OF THE JACOBIAN

For brevity we write $a_t := \big(t\sigma^2 - (1-t)\big)/\rho_t^2$ for the coefficient of the isotropic term. Differentiating equation 16 with $\nabla_{\boldsymbol{x}} z(\boldsymbol{x}, t) = t\boldsymbol{\mu}/\rho_t^2$ and $\tanh' = \operatorname{sech}^2$ yields

$$\partial_{\boldsymbol{x}} \boldsymbol{v}(\boldsymbol{x}, t) = a_t\,\boldsymbol{I} + b_t(\boldsymbol{x})\,\mathbf{e}_{\boldsymbol{\mu}}\mathbf{e}_{\boldsymbol{\mu}}^\top \qquad \text{where} \qquad b_t(\boldsymbol{x}) := \frac{\|\boldsymbol{\mu}\|^2\, t(1-t)}{\rho_t^4}\operatorname{sech}^2\big(z(\boldsymbol{x}, t)\big) \ \geq\ 0, \tag{17}$$

with the rank-one term arises from $\boldsymbol{\mu}\,\boldsymbol{\mu}^\top = \|\boldsymbol{\mu}\|^2 \mathbf{e}_{\boldsymbol{\mu}}\mathbf{e}_{\boldsymbol{\mu}}^\top$. Thus the jacobian $\boldsymbol{J}(\boldsymbol{x}, t) = \partial_{\boldsymbol{x}} \boldsymbol{v}(\boldsymbol{x}, t)$ is isotropic apart from a rank-one term along $\mathbf{e}_{\boldsymbol{\mu}}$. Specifically, it has eigenvalue $a_t + b_t(\boldsymbol{x})$ along $\mathbf{e}_{\boldsymbol{\mu}}$ and a_t on the orthogonal complement $\mathbf{e}_{\boldsymbol{\mu}}^\perp$.

STEP 3: PROOF OF (1)

We analyze the normalized reward direction $\boldsymbol{w}_t := \nabla_{\boldsymbol{x}_t} r(\boldsymbol{x}_1)/\|\nabla_{\boldsymbol{x}_t} r(\boldsymbol{x}_1)\|$ by separating $t = 1$ and $t \in [0, 1)$.

Terminal ($t = 1$). At $t = 1$, the claim follows directly from the chain rule:

$$\nabla_{\boldsymbol{x}_1} r(\boldsymbol{x}_1) = \varphi'(\boldsymbol{e}_{\boldsymbol{\mu}}^\top \boldsymbol{x}_1)\, \boldsymbol{e}_{\boldsymbol{\mu}}. \tag{18}$$

Since $\varphi' > 0$, the scalar factor is strictly positive, immediately yielding $\boldsymbol{w}_1 = \mathbf{e}_{\boldsymbol{\mu}}$.

Induction ($t \in [0, 1)$). Fix $t \in [0, 1)$ and let $t = t_0 < t_1 < \cdots < t_K = 1$ be an arbitrary time grid on which the trajectory is generated by Euler updates $\boldsymbol{x}_{t_{k+1}} = \boldsymbol{x}_{t_k} + h_k\, \boldsymbol{v}(\boldsymbol{x}_{t_k}, t_k)$, $h_k := t_{k+1} - t_k$. Writing $\boldsymbol{J}_k := \boldsymbol{J}(\boldsymbol{x}_{t_k}, t_k)$, adjoint method propagates the reward gradient backward using $\partial \boldsymbol{x}_{t_{k+1}} / \partial \boldsymbol{x}_{t_k} = \boldsymbol{I} + h_k \boldsymbol{J}_k$, which becomes:

$$\nabla_{\boldsymbol{x}_{t_k}} r(\boldsymbol{x}_1) = \big(\boldsymbol{I} + h_k \boldsymbol{J}_k\big)^\top \nabla_{\boldsymbol{x}_{t_{k+1}}} r(\boldsymbol{x}_1), \qquad k = K-1, \ldots, 0. \tag{19}$$

It therefore suffices to show that each factor $(\boldsymbol{I} + h_k \boldsymbol{J}_k)^\top$ maps $\mathbf{e}_{\boldsymbol{\mu}}$ to a positive multiple of itself. By equation 17, $\boldsymbol{J}_k$ is symmetric with $\boldsymbol{J}_k \mathbf{e}_{\boldsymbol{\mu}} = \big(a_{t_k} + b_{t_k}(\boldsymbol{x}_{t_k})\big)\mathbf{e}_{\boldsymbol{\mu}}$, hence

$$\big(I + h_k \boldsymbol{J}_k\big)^\top \mathbf{e}_{\boldsymbol{\mu}} = \big(1 + h_k\, (a_{t_k} + b_{t_k}(\boldsymbol{x}_{t_k}))\big)\, \mathbf{e}_{\boldsymbol{\mu}}. \tag{20}$$

Under the step-size condition $1 + h_k\, (a_{t_k} + b_{t_k}(\boldsymbol{x}_{t_k})) > 0$ for all k, backward induction from the terminal boundary $\nabla_{\boldsymbol{x}_1} r(\boldsymbol{x}_1) = \varphi'(\mathbf{e}_{\boldsymbol{\mu}}^\top \boldsymbol{x}_1)\, \mathbf{e}_{\boldsymbol{\mu}}$ gives

$$\nabla_{\boldsymbol{x}_t} r(\boldsymbol{x}_1) = \varphi'(\mathbf{e}_{\boldsymbol{\mu}}^\top \boldsymbol{x}_1) \prod_{k=0}^{K-1} \big(1 + h_k\, (a_{t_k} + b_{t_k}(\boldsymbol{x}_{t_k}))\big)\, \mathbf{e}_{\boldsymbol{\mu}}, \tag{21}$$

with a strictly positive prefactor, so $\boldsymbol{w}_t = \boldsymbol{e}_{\boldsymbol{\mu}}$ for all $t \in [0, 1)$. The argument uses only that $\mathbf{e}_{\boldsymbol{\mu}}$ is a common eigenvector of every $\boldsymbol{J}_k$; it does not require the Jacobians along the trajectory to commute.

Combining both cases gives $\boldsymbol{w}_t = \boldsymbol{e}_{\boldsymbol{\mu}}$ for all $t \in [0, 1]$.

STEP 4: PROOF OF (II)

By (i) and equation 17, the directional expansion rate equation 2 becomes:

$$\lambda_{\boldsymbol{w}}(\boldsymbol{x}_t, t) = \boldsymbol{e}_{\boldsymbol{\mu}}^\top \boldsymbol{J}(\boldsymbol{x}_t, t)\, \boldsymbol{e}_{\boldsymbol{\mu}} = \frac{t\sigma^2 - (1-t)}{\rho_t^2} + \frac{\|\boldsymbol{\mu}\|^2\, t(1-t)}{\rho_t^4}\, \mathrm{sech}^2\Big(\frac{t\, \boldsymbol{\mu}^\top \boldsymbol{x}_t}{\rho_t^2}\Big), \tag{22}$$

which is equation 5. □

B DERIVATION OF THE TRANSITION TIMES

We continue with the notation of Appendix A, writing $\lambda_{\boldsymbol{w}}(\boldsymbol{x}, t) = a_t + b_t(\boldsymbol{x})$ as in equation 17. We derive the three transition times along a sampling trajectory: t_e marks the first time the expansion region opens; t_c, the time a given trajectory leaves it; t_σ the expansion by the isotropic term after mode commitment.

B.1 EXPANSION ONSET t_e

For fixed t, $\lambda_{\boldsymbol{w}}$ depends on the state only through $\mathrm{sech}^2(t\, \boldsymbol{\mu}^\top \boldsymbol{x} / \rho_t^2)$, so it is maximized on the hyperplane $\{\boldsymbol{x} : \boldsymbol{\mu}^\top \boldsymbol{x} = 0\}$. We write $\lambda_{\boldsymbol{w}}^{max}$ for this maximum value:

$$\lambda_{\boldsymbol{w}}^{\max}(t) = \frac{t\sigma^2 - (1-t)}{\rho_t^2} + \frac{\|\boldsymbol{\mu}\|^2\, t(1-t)}{\rho_t^4}. \tag{23}$$

The onset time t_e is the first moment at which $\lambda_{\boldsymbol{w}}^{\max}$ crosses from negative to positive. Specifically, multiplying equation 23 by ρ_t^4 presents the equation for t_e as follows:

$$g(t) := \big(t\sigma^2 - (1-t)\big)\rho_t^2 + \|\boldsymbol{\mu}\|^2\, t(1-t) = 0. \tag{24}$$

In the regime $t\sigma \ll 1 - t$, which holds throughout the early stage for well-separated modes, $\rho_t^2 \to (1-t)^2$ and g factors as $(1-t)\big[\|\boldsymbol{\mu}\|^2 t - (1-t)^2\big]$. Hence, we obtain simplified equation as,

$$t^2 - (2 + \|\boldsymbol{\mu}\|^2)\, t + 1 = 0. \tag{25}$$

whose smaller root becomes

$$t_e = \frac{2 + \|\boldsymbol{\mu}\|^2 - \|\boldsymbol{\mu}\|\sqrt{\|\boldsymbol{\mu}\|^2 + 4}}{2}. \tag{26}$$

Finally, rationalizing the numerator gives the form as follows:

$$t_e = \frac{2}{2 + \|\boldsymbol{\mu}\|^2 + \|\boldsymbol{\mu}\|\sqrt{\|\boldsymbol{\mu}\|^2 + 4}}. \tag{27}$$

B.2 Contraction Boundary $\mathcal{B}$ and Crossing Time t_c

Second transition t_c varies across trajectories. We therefore represent it as a curve in the $(\boldsymbol{x}, t)$ plane and define t_c as the time at which a trajectory crosses this curve. Setting $\lambda_{\boldsymbol{w}}(\boldsymbol{x}, t) = 0$ in equation 5 gives

$$\operatorname{sech}^2\Big(\frac{t\|\boldsymbol{\mu}\|\, \boldsymbol{e}_{\boldsymbol{\mu}}^\top \boldsymbol{x}}{\rho_t^2}\Big) = \frac{\rho_t^2\,(1 - t - t\sigma^2)}{\|\boldsymbol{\mu}\|^2\, t(1-t)}. \tag{28}$$

Inverting equation 28 yields the transition boundary $\mathcal{B} = \{(\boldsymbol{x}, t) : f(\boldsymbol{x}, t) = 0\}$ of equation 7,

$$f(\boldsymbol{x}, t) = |\boldsymbol{e}_{\boldsymbol{\mu}}^\top \boldsymbol{x}| - \frac{\rho_t^2}{t\|\boldsymbol{\mu}\|}\, \cosh^{-1}\sqrt{\frac{\|\boldsymbol{\mu}\|^2\, t(1-t)}{\rho_t^2\,(1 - t - t\sigma^2)}}\,, \tag{29}$$

which is real-valued exactly on $t \in [t_e, t_\sigma)$, with $f < 0$ in the expansion region and $f > 0$ in the contraction region. As a trajectory commits to a mode, it crosses $\mathcal{B}$ at

$$t_c(\boldsymbol{x}_0) \coloneqq \inf\{\, t > t_e : f(\boldsymbol{x}_t, t) \geq 0 \,\}, \tag{30}$$

where $\lambda_{\boldsymbol{w}}$ turns negative.

B.3 Re-expansion Time t_σ

As the trajectory further settles into a mode, the mode-selection term becomes negligible. Applying the inequality $\operatorname{sech}(z) \leq 2e^{-|z|}$ yields

$$0 \leq b_t(\boldsymbol{x}_t) \leq \frac{4\|\boldsymbol{\mu}\|^2\, t(1-t)}{\rho_t^4} \exp\!\left(-\frac{2t\,|\boldsymbol{\mu}^\top \boldsymbol{x}_t|}{\rho_t^2}\right), \tag{31}$$

where the bound decays exponentially once $t\,|\boldsymbol{\mu}^\top \boldsymbol{x}_t| \gg \rho_t^2$. In this committed regime $\lambda_{\boldsymbol{w}}(\boldsymbol{x}_t, t) \approx a_t$, the numerator alone determines the sign, which gives

$$t_\sigma = \frac{1}{1 + \sigma^2}. \tag{32}$$

Since the numerator increases monotonically, $\lambda_{\boldsymbol{w}} > 0$ for $t > t_\sigma$. This final expansion phase reflects the terminal spreading of each mixture component toward its variance σ^2.

C Experimental Details

C.1 Toy Experiment

To verify the phase transitions, we consider the two-mode Gaussian mixture with dimension $D = 8$ as follows:

$$\boldsymbol{x}_1 \sim \tfrac{1}{2}\mathcal{N}(\mu\, \boldsymbol{e}_1,\ \sigma^2 \boldsymbol{I}_D) + \tfrac{1}{2}\mathcal{N}(-\mu\, \boldsymbol{e}_1,\ \sigma^2 \boldsymbol{I}_D), \qquad \mu = 2.5,\ \ \sigma = 0.25,$$

where $\boldsymbol{I}_D$ is the D-dimensional identity matrix and $\boldsymbol{e}_1$ is the first standard basis vector, so that the two components are separated by 2μ along a single axis with isotropic covariance. We train a velocity field $\boldsymbol{v}_{\boldsymbol{\theta}}$ with standard conditional flow matching: we sample $\boldsymbol{x}_0 \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_D)$ and $t \sim \mathrm{Unif}[0, 1]$, form the linear interpolant $\boldsymbol{x}_t = (1-t)\boldsymbol{x}_0 + t\boldsymbol{x}_1$, and regress $\boldsymbol{v}_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t)$ onto the target velocity $\boldsymbol{x}_1 - \boldsymbol{x}_0$ under an MSE loss. The velocity field is an MLP with three hidden layers of width 64 and GELU activations, taking t concatenated to $\boldsymbol{x}_t$ as input, trained with Adam optimizer with learning rate 10^{-3} with batch size as 256 for 2×10^4 iterations.

C.2 Solving PDEs with Physical Hard Constraint

C.2.1 PDE and Constraint Operators

Notation. Within this section $\boldsymbol{u}$ denotes the PDE solution field, $\boldsymbol{\xi} \in \Omega$ the spatial coordinate and τ the physical time of the PDE; the flow time of Alg. 1 remains $t \in [0,1]$, and the reward direction of Section 3.1 always carries a subscript or a hat ($\boldsymbol{u}_t$, $\hat{\boldsymbol{u}}_t$). A generated sample $\boldsymbol{x}_1$ is the discretised space–time field, and we write $h(\boldsymbol{u})$ for a residual evaluated on the field that $\boldsymbol{x}_1$ represents. The reaction coefficient of the reaction–diffusion system is written κ to avoid collision with ρ_t of Section 3. We reuse the PDE settings of PCFM (Utkarsh et al., 2025); the flow models are described in Section C.2.2.

Constraint set and projection. For every system the hard constraint $\mathcal{C}$ of Alg. 1 is the zero set of a single residual, stacking a conditioning block (an initial or boundary condition) on top of a conservation block,

$$\mathcal{C} = \{\boldsymbol{x}_1 : h(\boldsymbol{x}_1) = 0\}, \qquad h(\boldsymbol{u}) = \begin{bmatrix} h_{\mathrm{IC}}(\boldsymbol{u}) \\ h_{\mathrm{CL}}(\boldsymbol{u}) \end{bmatrix} \in \mathbb{R}^m, \qquad \boldsymbol{J} := \nabla h(\boldsymbol{u})^\top \in \mathbb{R}^{m\times n}, \tag{33}$$

so that blocks acting on overlapping entries of the field are satisfied simultaneously rather than in sequence. The operator $\Pi_{\mathcal{C}}$ is the minimum-norm correction onto $\mathcal{C}$, realised by the Gauss–Newton iteration

$$\boldsymbol{u}^{(0)} = \boldsymbol{x}_1, \qquad \boldsymbol{u}^{(j+1)} = \boldsymbol{x}_1 - \boldsymbol{J}^\top\big(\boldsymbol{J}\boldsymbol{J}^\top\big)^{-1}\Big(h\big(\boldsymbol{u}^{(j)}\big) + J\big(\boldsymbol{x}_1 - \boldsymbol{u}^{(j)}\big)\Big), \qquad \Pi_{\mathcal{C}}(\boldsymbol{x}_1) := \boldsymbol{u}^{(j^\star)}, \tag{34}$$

stopped at the first index $j^\star$ with $\|h(\boldsymbol{u}^{(j)})\| < \epsilon$; a single step is exact whenever h is affine.

Heat equation. The heat equation is taken on $\xi \in [0, 2\pi)$ with periodic boundaries and $\tau \in [0,1]$, as $\partial_\tau u = \alpha\, \partial_{\xi\xi} u$ with $u(\xi, 0) = \sin(\xi + \varphi)$, $\alpha \sim \mathcal{U}(1,5)$ and $\varphi \sim \mathcal{U}(0,\pi)$, on a 100×100 space–time grid. Periodicity makes the spatial integral invariant, so total mass is conserved and

$$h_{\mathrm{Heat}}(\boldsymbol{u}) = \begin{bmatrix} u(\xi,0) - u_{\mathrm{IC}}(\xi) \\ \int_\Omega u(\xi,\tau)\, d\xi - \int_\Omega u(\xi,0)\, d\xi \end{bmatrix}, \qquad \Omega = [0, 2\pi). \tag{35}$$

Both blocks are linear, so equation 34 terminates in one step.

Inviscid Burgers. We use $\partial_\tau u + \partial_\xi\big(\frac{1}{2}u^2\big) = 0$ on $\xi \in [0,1]$, $\tau \in [0,1]$, with a smoothed step initial condition $u(\xi,0) = \big(1 + \exp((\xi - p)/\epsilon)\big)^{-1}$ at location $p \sim \mathcal{U}(0.2, 0.8)$ with $\epsilon = 0.02$, a Dirichlet condition $u(0,\tau) = u_{\mathrm{bc}}$ with $u_{\mathrm{bc}} \sim \mathcal{U}(0,1)$ on the left and a zero-gradient Neumann condition on the right, on a 101×101 grid. Total mass is not invariant here, and the conservation block is the accumulated form of the flux balance, $h^{\mathrm{CL}}_{\mathrm{Burgers}}(u) = M(\tau) - M(0) + \int_0^\tau \frac{1}{2}\big(u(1,s)^2 - u(0,s)^2\big)\, ds$, where $M(\tau) := \int_0^1 u(\xi,\tau)\, d\xi$.[1] We condition on the initial condition and additionally enforce local conservation,

$$h_{\text{Burgers-IC}}(u) = \begin{bmatrix} u(\xi,0) - u_{\mathrm{IC}}(\xi) \\ R^{(k)}_{\mathrm{Flux}}(u) \\ h^{\mathrm{CL}}_{\mathrm{Burgers}}(u) \end{bmatrix}, \qquad R^{(k)}_{\mathrm{Flux}}:\ u_i^{j+1} = u_i^j - \frac{\Delta\tau}{\Delta\xi}\,\Delta F_i^j, \tag{36}$$

where $R^{(k)}_{\mathrm{Flux}}$ unrolls $k = 5$ steps of the Godunov update, $\Delta F_i^j = F(u_i^j, u_{i+1}^j) - F(u_{i-1}^j, u_i^j)$, with numerical flux

$$F(u_-, u_+) = \begin{cases} \min\big(\frac{1}{2}u_-^2,\ \frac{1}{2}u_+^2\big), & u_- \le u_+, \\ \frac{1}{2}u_-^2, & u_- > u_+,\ \frac{u_- + u_+}{2} > 0, \\ \frac{1}{2}u_+^2, & u_- > u_+,\ \frac{u_- + u_+}{2} \le 0. \end{cases} \tag{37}$$

The residual is nonlinear in the state, and its middle block embeds a discretisation of the solver rather than a physical invariant alone.

Reaction–diffusion. We use the nonlinear system on $\xi \in [0,1]$ in flux form, $\partial_\tau u = \kappa\, u(1-u) - \partial_\xi f$ with $f = -\nu\, \partial_\xi u$ and $(\kappa, \nu) = (0.01, 0.005)$, under Neumann conditions fixing the flux at the two

[1] PCFM (Utkarsh et al., 2025) states this block as constant total mass; the nonlinear form above is the one implemented in PCFM's official code, and total mass is measurably not conserved in the released solutions.

Table 3: Sampler settings for the two projection-based baselines. Guided interpolation is reported as the triple (penalty strength λ, inner step size, inner iterations).

	Heat Equation	Burgers (IC)	Reaction–Diffusion
D-Flow			
Constraints enforced	IC+CL	IC+CL+flux	IC+CL
Learning rate	7e-2	1e-3	5e-3
Iteration / Steps	20/50	20/50	20/50
OC-Flow			
Constraints enforced	IC+CL	IC+CL+flux	IC+CL
Learning rate	1e-4	1e-4	5e-4
Iteration / Steps	40/25	40/25	40/25
ECI			
Constraints enforced	IC	IC	IC
Mixing steps n_{mix}	1	7	5
Prior resampling interval	1	7	5
PCFM			
Constraints enforced	IC+CL	IC+CL+flux	IC+CL
Newton updates per step	2	1	1
Guided interpolation	off	$(10^{-3},\ 5,\ 20)$	$(10^{-3},\ 0.01,\ 20)$
Proposed			
Constraints enforced	IC+CL	IC+CL+flux	IC+CL
Learning rate	5e-3	7e-3	1e-2
Iteration/Steps	30/5	30/20	30/20

ends, $f|_{\xi=0} = g_L$ and $f|_{\xi=1} = g_R$, on a 128×100 grid.[2] The total mass $M(\tau)$ is no longer invariant but is driven by the reaction term and the boundary fluxes, giving

$$h_{\mathrm{RD}}(u) = \begin{bmatrix} u(\xi,0) - u_{\mathrm{IC}}(\xi) \\ M(\tau) - M(0) - \kappa \displaystyle\int_0^\tau \int_0^1 u(1-u)\, d\xi\, ds - \int_0^\tau \big(g_L(s) - g_R(s)\big)\, ds \end{bmatrix}. \tag{38}$$

The boundary fluxes are not supplied to the residual but recovered from the sample itself, by one-sided differences of $-\nu\, \partial_\xi u$ at the two ends, so h_{RD} constrains x_1 without reference to the boundary conditions being generated. The conservation block is nonlinear, so J depends on the state and the iteration in equation 34 is required.

C.2.2 Experimental Details

Setup. In our method, we employ Gauss–Newton projection (Utkarsh et al., 2025) for the hard constraints, as the projection $\Pi_{\mathcal{C}}$ in contraction phase. For each system, every method samples from a single pretrained flow model, so that the comparison isolates the effect of the sampler rather than that of the generative model. All three systems are one-dimensional in space, and the corresponding models follow the pretraining configuration of PCFM (Utkarsh et al., 2025): an FNO backbone with a Matérn Gaussian process prior ($\nu = 0.5$, lengthscale 10^{-3}) in place of white noise. The Burgers model is used as pretrained, whereas the heat and reaction–diffusion models are additionally fine-tuned under a reweighted flow-time distribution. Both baselines described below operate by correcting an estimate of the terminal sample at each Euler step, and differ in how this correction is computed and in which constraints it is able to enforce. Table 3 reports the settings used for both baselines.

C.3 Data Assimilation

Navier–Stokes. We use the Navier–Stokes dataset and the pretrained drift model of FlowDAS (Chen et al., 2025), following the setting of Chen et al. (2024). The vorticity ω of an incompressible flow

[2] PCFM (Utkarsh et al., 2025) writes this equation with an advective term $-\nu\, \partial_\xi u$; we use the diffusive flux form implemented in PCFM's official code, which is what the conservation law below follows from.

on the torus $[0, 2\pi]^2$ evolves as

$$d\omega + \mathbf{U} \cdot \nabla\omega \, d\tau = \nu\Delta\omega \, d\tau - \alpha\,\omega \, d\tau + \varepsilon \, d\xi, \qquad \mathbf{U} = \nabla^{\perp}\psi, \quad -\Delta\psi = \omega, \tag{39}$$

with $\nu = 10^{-3}$, $\alpha = 0.1$, $\varepsilon = 1$, where τ is the physical time and $d\xi$ is a white-in-time random forcing acting on a few Fourier modes. Each trajectory consists of 50 snapshots at 128×128 resolution, and the vorticity is normalized by its average pixel norm, in which RMSE is computed. We assimilate 8 held-out test trajectories, starting from their first $L = 10$ snapshots. The super-resolution operator bilinearly downsamples the field to 16×16, and the sparse-observation operator applies a fixed random binary mask that observes 5% of the pixels, shared by all trajectories and methods. Both observations are corrupted by Gaussian noise with standard deviation 0.05.

SEVIR. We use the VIL frames of SEVIR (Veillette et al., 2020) at 128×128 resolution and a 10-minute interval, rescaled to $[0, 1]$, together with the pretrained drift model of FlowDAS. Test sequences are drawn from storm events recorded after June 1, 2019, and we assimilate one sequence from each of 8 distinct events, starting from its first $L = 6$ frames. The observation operator applies a fixed random binary mask that observes 10% of the grid cells, with Gaussian noise of standard deviation 10^{-3}.

Metrics. RMSE is computed between each assimilated state and the ground truth. For CSI, both fields are binarized at a threshold $\gamma \in \{0.3, 0.5\}$, and $\mathrm{CSI}_\gamma = \mathrm{hits}/(\mathrm{hits} + \mathrm{misses} + \mathrm{false\ alarms})$. Each metric is averaged over the reported assimilation steps of a trajectory, and we report its mean and standard deviation over the 8 test trajectories.

Implementation details. FlowDAS uses 500 Euler–Maruyama steps and $J = 25$ Monte Carlo samples for posterior estimation, with the guidance step size $\zeta = 1$ on NS SO and $\zeta = 0.1$ on SEVIR as in FlowDAS (Chen et al., 2025). On NS SR, we use $\zeta = 0.5$, the best value among $\{0.25, 0.5, 1, 2, 4\}$, which improves over $\zeta = 2$ used in FlowDAS (next-frame RMSE 0.188 vs. 0.207). OC-Flow, D-Flow and our method optimize their controls with AdamW on a coarse time grid. On NS, all three use 20 grid points, 80 iterations and a learning rate of 5×10^{-3}. On SEVIR, they use 300 iterations and a learning rate of 2×10^{-3}, with 20 grid points for OC-Flow and D-Flow and 35 for ours, and the gradient norm is clipped at 10^{-5} for OC-Flow and ours. For our method, $\hat{\lambda}_{\boldsymbol{w}}$ is estimated with a finite-difference Jacobian–vector product (step size 0.05) along an unguided trajectory to obtain $\hat{t}_{\mathrm{peak}}$ and x_{peak}, and $\hat{t}_c$ is updated every $n_c = 10$ iterations.

C.3.1 Three-Step Roll-out

Table 4: Data assimilation results averaged over the first three assimilation steps ($k = 3$), the setting shown in Fig. 4 and Fig. 6. Protocol, metrics and test trajectories are identical to Table 2, which reports the next frame ($k = 1$). Best results are in **bold**.

	Navier–Stokes (RMSE↓)		SEVIR, SO 10%		
Method	SR 8×	SO 5%	RMSE↓	$\mathrm{CSI}_{0.3}$ ↑	$\mathrm{CSI}_{0.5}$ ↑
FlowDAS	0.238±0.039	0.209±0.030	0.078±0.013	0.783±0.062	0.662±0.039
OC-Flow	0.304±0.039	0.203±0.021	0.111±0.020	0.695±0.088	0.575±0.047
D-Flow	0.310±0.053	0.297±0.054	0.093±0.016	0.739±0.072	0.604±0.044
Ours	**0.196**±0.030	**0.174**±0.013	**0.070**±0.010	**0.800**±0.064	**0.665**±0.051

Fig. 4 and Fig. 6 visualize the third assimilated frame, whereas Table 2 reports the next frame only. Table 4 therefore reports the same metrics averaged over the first three assimilation steps, on all three tasks. Our method attains the best value in every column, and its margin widens relative to the next frame: the RMSE reduction with respect to FlowDAS grows from 10.2%, 12.3% and 4.4% at $k = 1$ to 17.4%, 16.6% and 11.2% at $k = 3$ on NS SR $8\times$, NS SO 5% and SEVIR, respectively. On SEVIR, our CSI margin over FlowDAS also widens, from 0.823 vs 0.819 at $k = 1$ to 0.800 vs 0.783 at $k = 3$ for $\mathrm{CSI}_{0.3}$, consistent with the storm structure preserved in Fig. 6.

C.3.2 Long-range Roll-out

Table 5 extends the NS SR $8\times$ experiment to 10 autoregressive assimilation steps, with all other settings unchanged. Since each assimilated state conditions the next step, errors accumulate for

Table 5: Long-range roll-out on NS SR $8\times$: RMSE averaged over the first k assimilation steps (mean$\pm$std over 8 test trajectories). Best results are in **bold**.

Method	$k = 1$	$k = 3$	$k = 6$	$k = 10$
FlowDAS	0.188$\pm$0.025	0.238$\pm$0.039	0.314$\pm$0.051	0.415$\pm$0.047
OC-Flow	0.247$\pm$0.032	0.304$\pm$0.039	0.362$\pm$0.040	0.416$\pm$0.037
D-Flow	0.229$\pm$0.018	0.310$\pm$0.053	0.440$\pm$0.069	0.601$\pm$0.052
Ours	**0.169**$\pm$0.028	**0.196**$\pm$0.030	**0.246**$\pm$0.030	**0.298**$\pm$0.026

every method, but ours accumulates the slowest. The RMSE reduction of ours relative to FlowDAS grows from 10.2% at $k = 1$ to 17.4%, 21.4% and 28.1% at $k = 3$, 6 and 10, and ours attains lower RMSE than every baseline on all 8 test trajectories for $k \geq 3$. Fig. 5 visualizes the roll-out at $t = 1$, 4, 7 and 10.

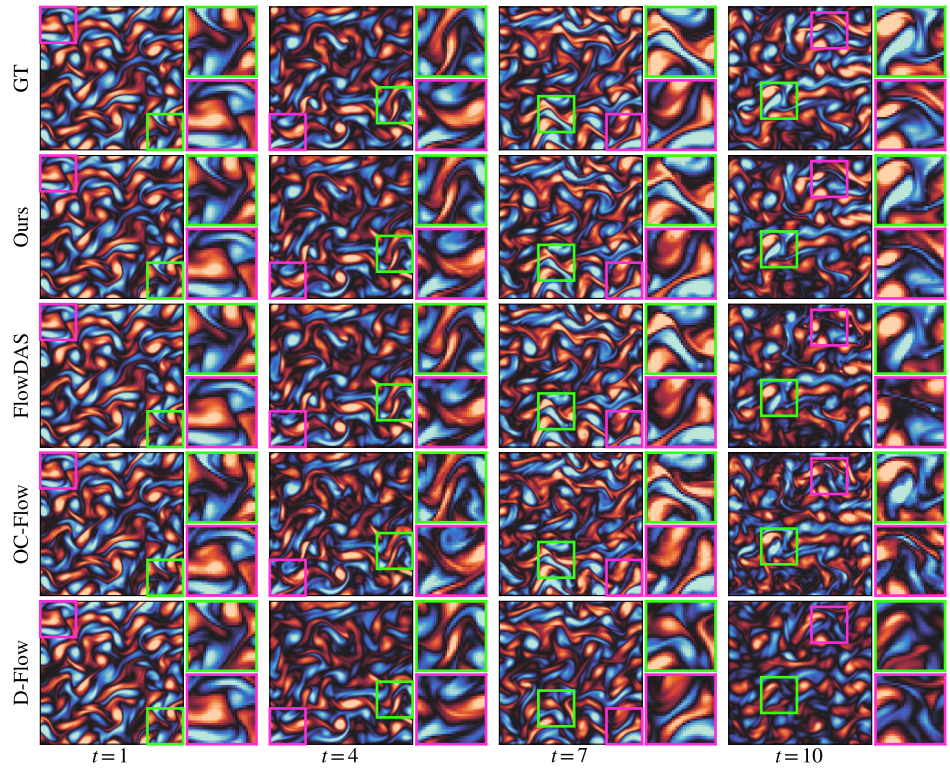

Figure 5: Long-range data assimilation roll-out on NS SR $8\times$ at assimilation steps $k = 1$, 4, 7 and 10. The zoomed patches and the choice of the test trajectory follow Fig. 4, with the median error reduction taken over the four steps shown.

C.3.3 Visualization on SEVIR

Fig. 6 visualizes three-step roll-outs on SEVIR. Within three steps, OC-Flow and D-Flow accumulate spurious low-intensity precipitation over clear-sky regions, whereas ours and FlowDAS preserve the storm structure.

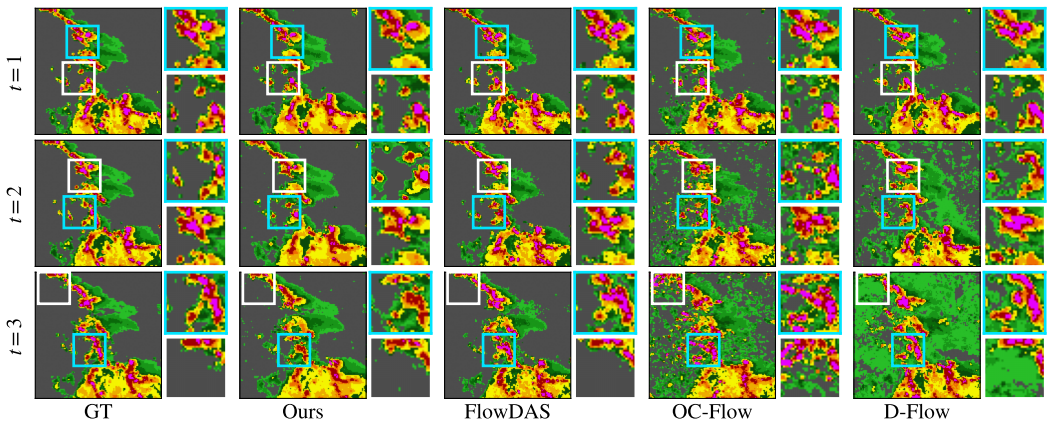

Figure 6: Three-step data assimilation roll-outs on SEVIR (VIL) with 10% sparse observation. Each row is an assimilation step t, and each column a method. The zoomed patches and the choice of the test trajectory follow Fig. 4.