Primal-Dual Flow Matching for Sample-Wise Constrained Generation

NeurIPS 2026

Zhengyan Huan[1,2] Zhengyan.Huan@tufts.edu

Peter Y. Lu[1,2] Peter.Lu@tufts.edu

Shuchin Aeron[1,2] Shuchin.Aeron@tufts.edu

1: Department of Electrical and Computer Engineering, Tufts University
2: NSF AI Institute for Artificial Intelligence and Fundamental Interactions

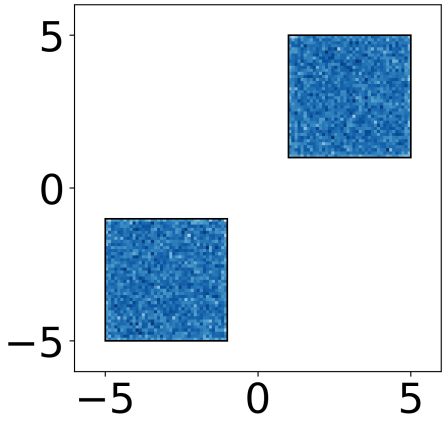
November 12, 2026

Project Page

Constraint-Aware Generation

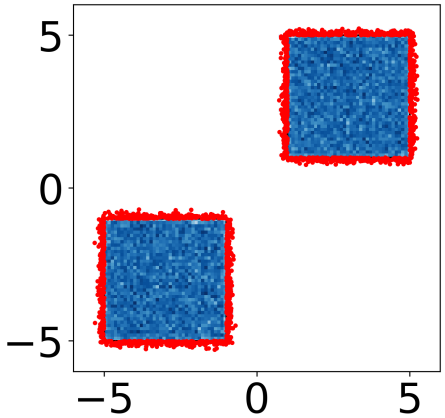
A model trained to match the data can still generate samples outside the feasible set.

Clean data



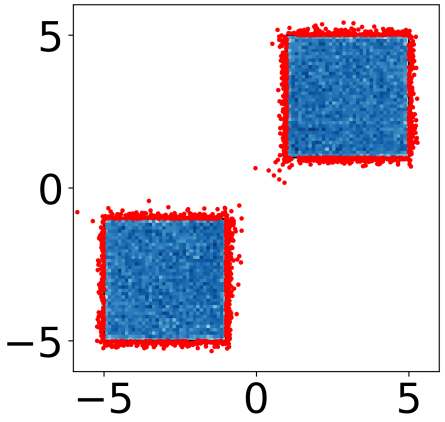
$P_{\text{clean}}(\mathcal{C}) = 1$

Noisy training data



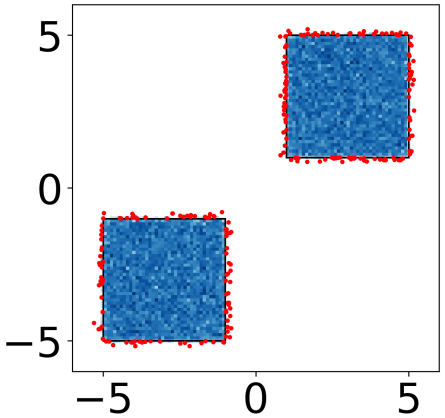
$Q(\mathcal{C}) \approx 96\%$

Standard FM



$P_{\text{FM}}(\mathcal{C}) = 96.37\%$

Optimal Solution



$P^*(\mathcal{C}) = 99.5\%$

Red points are infeasible. Standard flow matching falls short of the target $b = 99.5\%$.

Objective

Develop a constraint-aware generation method that

- preserves **distributional fidelity** and meets **a user-specified constraint satisfaction rate**;
- requires **only binary feedback** to assess feasibility (existing works [1],[2],[3] deal with certain types of constraints);
- provides provable theoretical guarantees (Theorems 3.5 and 4.5 in the paper).

Problem Formulation

Settings

- $x \in \mathbb{R}^d$: the sample;
- Q: the training distribution;
- P: the modeled distribution with density p;
- $\mathcal{C} \subset \mathbb{R}^d$: the constraint set;
- $b \in (0, 1)$: the user-specified target constraint satisfaction rate.

The population-level problem is given as

$$\underset{P}{\text{minimize}} \quad \mathbb{E}_{x \sim Q}\big[-\log p(x)\big] \quad \text{s.t.} \quad b - P(\mathcal{C}) \leq 0, \tag{OPT}$$

Here $P(\mathcal{C}) = \int_{\mathcal{C}} p(x)\,\mathrm{d}x$. The set $\mathcal{C}$ is accessible only through a binary membership oracle:

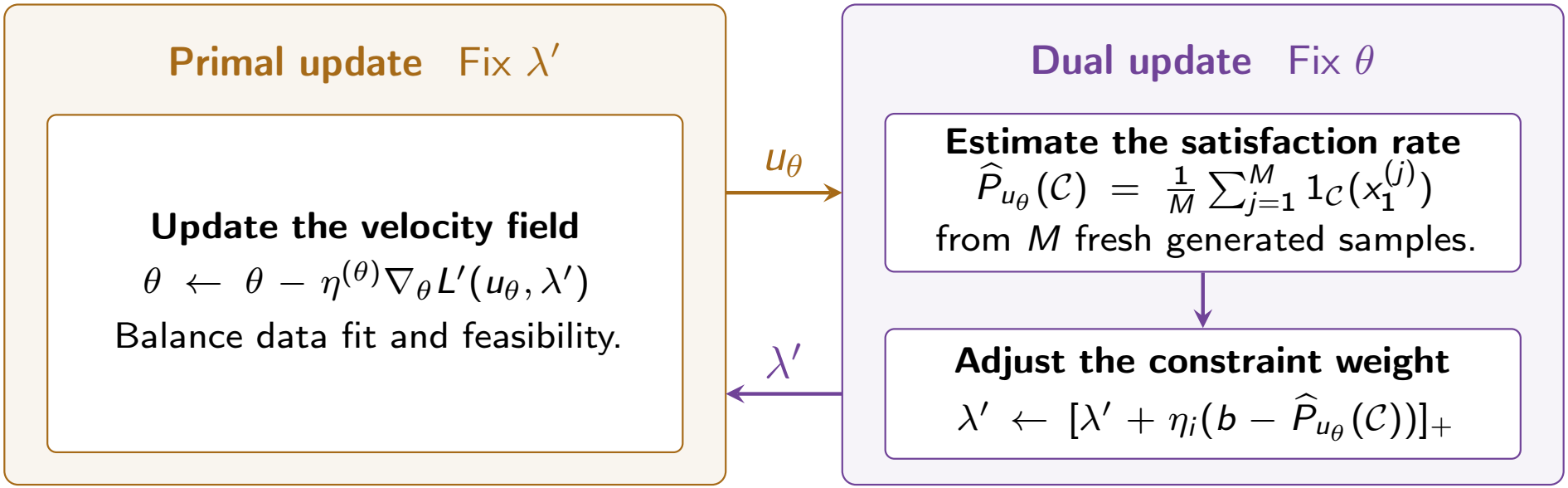
$$1_{\mathcal{C}}(x) = \begin{cases} 1, & x \in \mathcal{C}, \\ 0, & x \notin \mathcal{C}. \end{cases}$$

Primal-Dual Flow Matching (PDFM)

Define the Lagrangian L and its parameterized surrogate L' for practical optimization,

$$L(p, \lambda) = \mathbb{E}_{x\sim Q}[-\log p(x)] + \lambda\,(b - P(\mathcal{C}))\,;$$
$$L'(u_\theta, \lambda') = \mathcal{L}_{\mathrm{FM}}(\theta) + \lambda'\big(b - P_{u_\theta}(\mathcal{C})\big), \quad \mathcal{L}_{\mathrm{FM}}(\theta) = \mathbb{E}_{t,x_0,x_1}\left[\|u_\theta(x_t, t) - (x_1 - x_0)\|^2\right].$$

Primal-Dual Update:



- Initialize with $\lambda' = 0$;
- Stop when both the velocity field u_θ and the dual variable λ' stabilize.

Difficulty in computing $\nabla_\theta P_{u_\theta}(\mathcal{C})$

- $\Phi_{u_\theta,1}(x_0)$: the generated sample at $t=1$ starting from x_0 under u_θ.
- $\nabla_\theta P_{u_\theta}(\mathcal{C}) = \nabla_\theta \int p_0(x_0) 1_{\mathcal{C}}\left[\Phi_{u_\theta,1}(x_0)\right] dx_0$.

Learnable $f_\phi : \mathbb{R}^d \times [0,1] \to [0,1]$ as a surrogate for $1_{\mathcal{C}}\left[\Phi_{u_\theta,1}(x_0)\right]$

$$\nabla_\theta P_{u_\theta}(\mathcal{C}) \approx N \mathbb{E}_{i,x_{i\Delta t}}\left[\nabla_\theta u_\theta(x_{i\Delta t}, i\Delta t) \nabla_u f_\phi(x_{i\Delta t} + u\Delta t, (i+1)\Delta t)|_{u=u_\theta(x_{i\Delta t}, i\Delta t)}\right].$$

- $\Delta t = 1/N$ for N Euler steps, $i \sim \mathrm{Unif}\{0, \ldots, N-1\}$.
- $x_{i\Delta t}$: the Euler rollout up to time $i\Delta t$ following u_θ.
- f_ϕ: a **learned differentiable classifier** that predicts from (x_t, t) whether following u_θ will end in $\mathcal{C}$ at $t=1$.

f_ϕ's Training

Use the same terminal label $y = 1_{\mathcal{C}}(x_{N\Delta t}) = 1_{\mathcal{C}}(x_1)$ for each step of a rollout.
Fit f_ϕ by minimizing binary cross-entropy over ϕ.

$$\phi^\star \in \arg\min_\phi \mathbb{E}_{i,x_{i\Delta t}}\left[-y \log f_\phi(x_{i\Delta t}, i\Delta t) - (1-y)\log\left(1 - f_\phi(x_{i\Delta t}, i\Delta t)\right)\right].$$

.

PDE-Constrained Generation

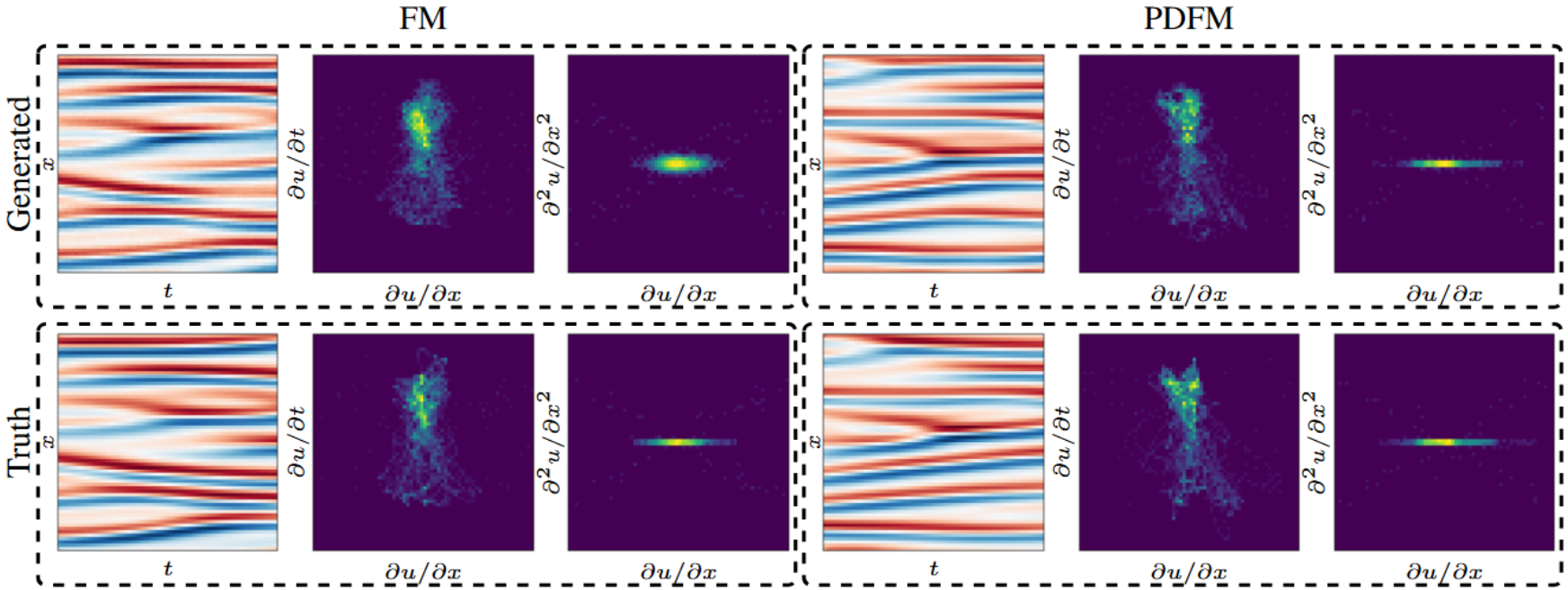
Task. Generate 64×64 space–time samples of the chaotic Kuramoto–Sivashinsky equation.

$$\partial_t u + \tfrac{1}{2}\partial_x(u^2) + \partial_x^2 u + \partial_x^4 u = 0, \qquad u = u(x,t).$$

Constraint. Feasible when mean absolute one-step solver residual $\delta_{L_1}(x) < \tau$.

Case 1. Clean data, $Q(\mathcal{C}) = 100\%$, $\tau = 0.033$, target $b = 98\%$.

Case 2. Spatially degraded data, $Q(\mathcal{C}) \approx 39.47\%$, $\tau = 0.05$, target $b = 90\%$.



Generated-Truth example (Case 1).
Top: generated samples and derivative histograms.
Bottom: solver rollouts from the corresponding initial conditions

PDE Results and Conclusion

- $P(\mathcal{C})$ measures the constraint satisfaction rate (higher is better);
- δ_{L_1} measures PDE residual error (lower is better).
- Histogram error (HE) measures distribution discrepancy (lower is better).

	Case 1 ($Q(\mathcal{C}) = 1$)			Case 2 ($Q(\mathcal{C}) \approx 39.47\%$)			
Method	$P(\mathcal{C})$ (%) ↑	δ_{L_1} ↓	HE ↓	$P(\mathcal{C})$ (%) ↑	δ_{L_1} ↓	HE_{clean} ↓	HE_{deg} ↓
FM(Q)	14.19	0.0409	(0.0087, 0.0453)	9.82	0.0622	(0.0094, 0.0465)	(0.0086, 0.0565)
FM($Q \mid \mathcal{C}$)	—	—	—	33.24	0.0539	(0.0135, 0.0477)	(0.0114, 0.0489)
PCFM [4]	6.03	0.0484	(0.0156, 0.1547)	0.06	0.1022	(0.0173, 0.0729)	(0.0186, 0.0823)
FM-RE [5]	43.17	0.0339	(0.0117, 0.0329)	46.33	0.0513	(0.0118, 0.0558)	(0.0097, 0.0570)
PDFM	**97.74**	0.0235	(0.0107, 0.0269)	**90.48**	0.0406	(0.0118, 0.0435)	(0.0103, 0.0457)

Conclusion

PDFM

- targets a user-specified feasibility rate through primal-dual training with binary membership feedback;
- achieves competitive constraint satisfaction and distribution fidelity.

References I

[1] Guan-Horng Liu, Tianrong Chen, Evangelos Theodorou, and Molei Tao.
Mirror diffusion models for constrained and watermarked generation.
In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, Advances in Neural Information Processing Systems, volume 36, pages 42898–42917. Curran Associates, Inc., 2023.

[2] Berthy T. Feng, Ricardo Baptista, and Katherine L. Bouman.
Neural approximate mirror maps for constrained diffusion models.
In The Thirteenth International Conference on Learning Representations, 2025.

[3] Tianyu Xie, Yu Zhu, Longlin Yu, Tong Yang, Ziheng Cheng, Shiyue Zhang, Xiangyu Zhang, and Cheng Zhang.
Reflected flow matching.
In ICML, 2024.

[4] Utkarsh, Pengfei Cai, Alan Edelman, Rafael Gómez-Bombarelli, and Christopher Rackauckas.
Physics-constrained flow matching: Sampling generative models with hard constraints.
In Advances in Neural Information Processing Systems, 2025.

References II

[5] Zhengyan Huan, Jacob Boerma, Liping Liu, and Shuchin Aeron.
Constraint-aware flow matching via randomized exploration.
Transactions on Machine Learning Research, 2026.