NEURAL INFORMATION PROCESSING SYSTEMS

Hall 1-4
Wed 07:00 CST
Sydney Poster Session 3

Rectifying Categorical Flows on Statistical Manifolds for One-Step Generation

Xiaohuan Jia[1] Renzhe Xu[2] Xiao Wang[2] Jiayun Wu[3] Shaohua Fan[1]

[1]Inner Mongolia University [2]Shanghai University of Finance and Economics
[3]Carnegie Mellon University







Carnegie Mellon University

Categorical Generation

Categorical generation has a wide range of applications.

Text Generation[1]

Random tokens · step 0/8

many water during of to water put without …

Intermediate · step 4/8

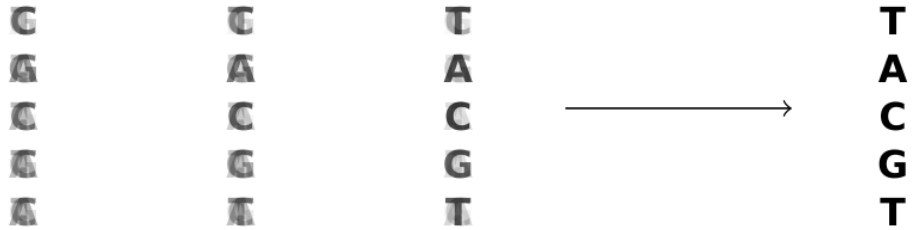
them those later, as be first man people …

Generated text · step 8/8

Many years later, as he faced the firing …

DNA Sequence Generation[2]

Dirichlet Flow Matching



Sequence-generation schematic

[1] Sahoo et al. The Diffusion Duality. ICML 2025.
[2] Stark et al. Dirichlet Flow Matching with Applications to DNA Sequence Design. ICML 2024.

Existing Methods for Categorical Generation

These methods require **multi-step sampling**, making generation costly.

Discrete diffusion

D3PM [3]
Structured discrete transitions.

SEDD [4]
Score entropy for probability ratios.

MDLM [5]
Masked-token denoising.

Flow matching

Discrete Flow Matching [6]
Discrete probability transport.

Fisher-Flow [7]
Fisher–Rao geodesics on the sphere.

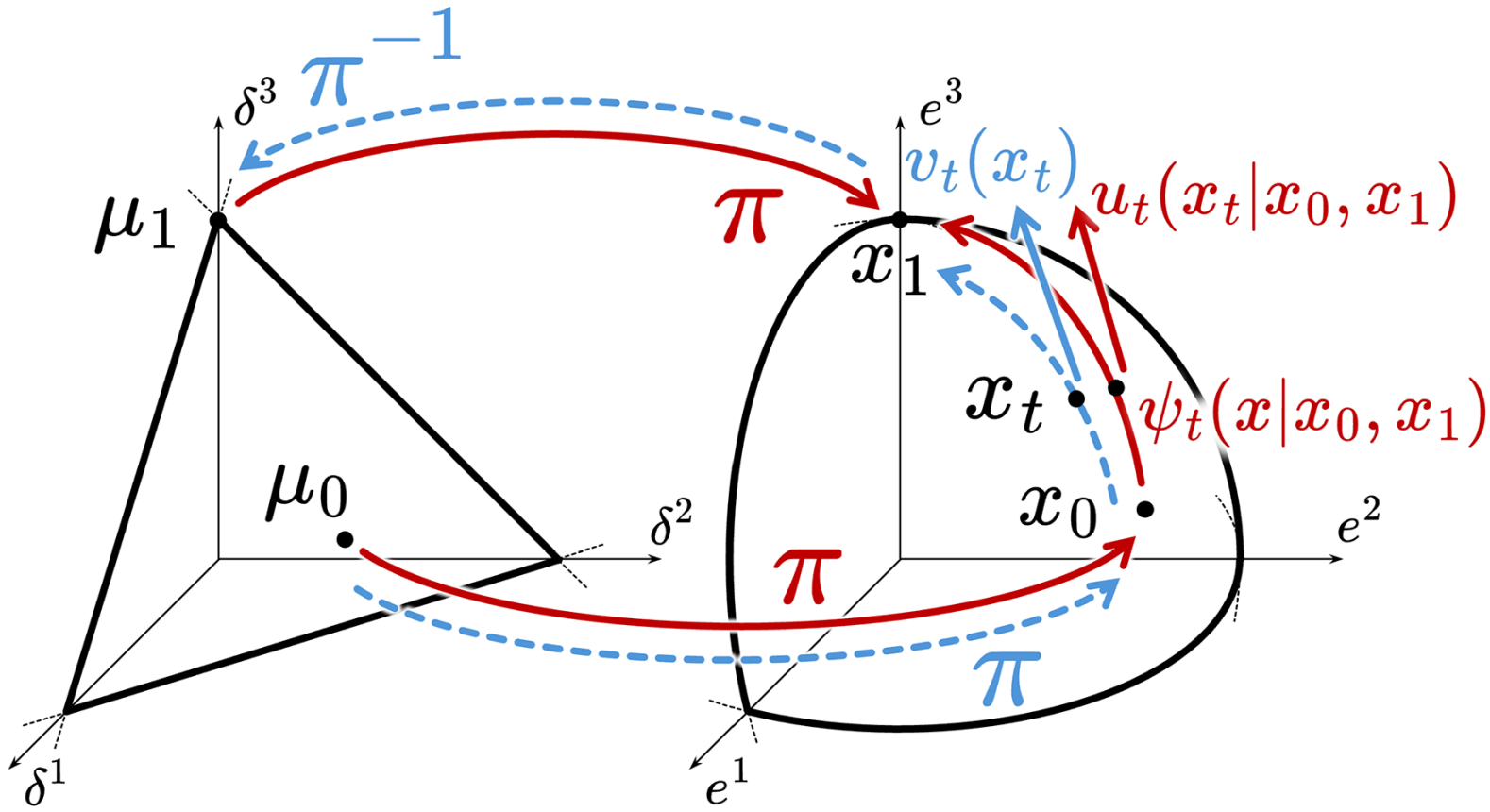
FM with General Discrete Paths [8]
General paths; kinetic-optimal velocities.

Statistical Flow Matching (SFM) [9]
Flows on statistical manifolds.

[3] Austin et al. Structured Denoising Diffusion Models in Discrete State-Spaces. NeurIPS 2021.
[4] Lou et al. Discrete Diffusion Modeling by Estimating the Ratios of the Data Distribution. ICML 2024.
[5] Sahoo et al. Simple and Effective Masked Diffusion Language Models. NeurIPS 2024.
[6] Gat et al. Discrete Flow Matching. NeurIPS 2024.
[7] Davis et al. Fisher Flow Matching for Generative Modeling over Discrete Data. NeurIPS 2024.
[8] Shaul et al. Flow Matching with General Discrete Paths: A Kinetic-Optimal Perspective. ICLR 2025.
[9] Cheng et al. Categorical Flow Matching on Statistical Manifolds. NeurIPS 2024.

Statistical Flow Matching[9]

Model categorical distributions on the sphere using Fisher–Rao geometry.



[9] Cheng et al. Categorical Flow Matching on Statistical Manifolds. NeurIPS 2024.

Why Statistical Flows Need Many Steps

Averaging directions at crossings yields **curved learned trajectories.**

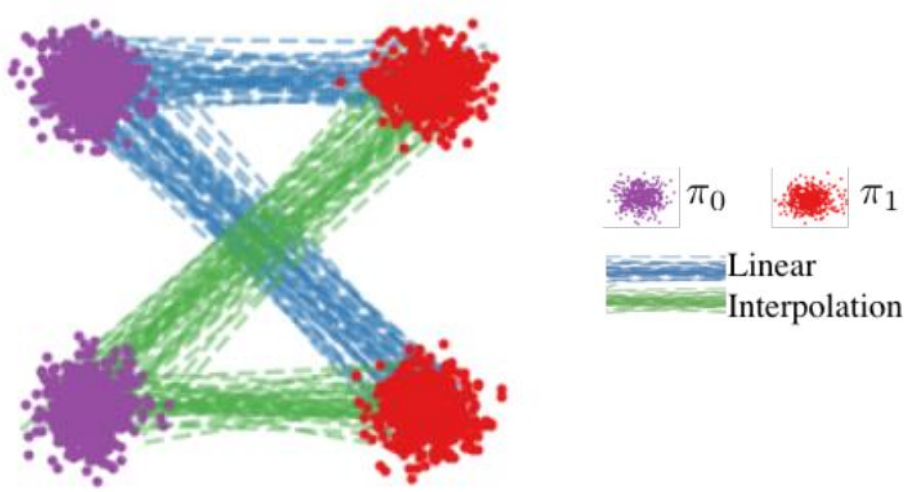
$$(X_0, X_1) \sim \pi(p_0) \times \pi(q)$$

Independent coupling

$$v_\theta(X_t, t) \approx \mathbb{E}[u_t^S(X_t \mid X_0, X_1) \mid X_t]$$

Averaged directions

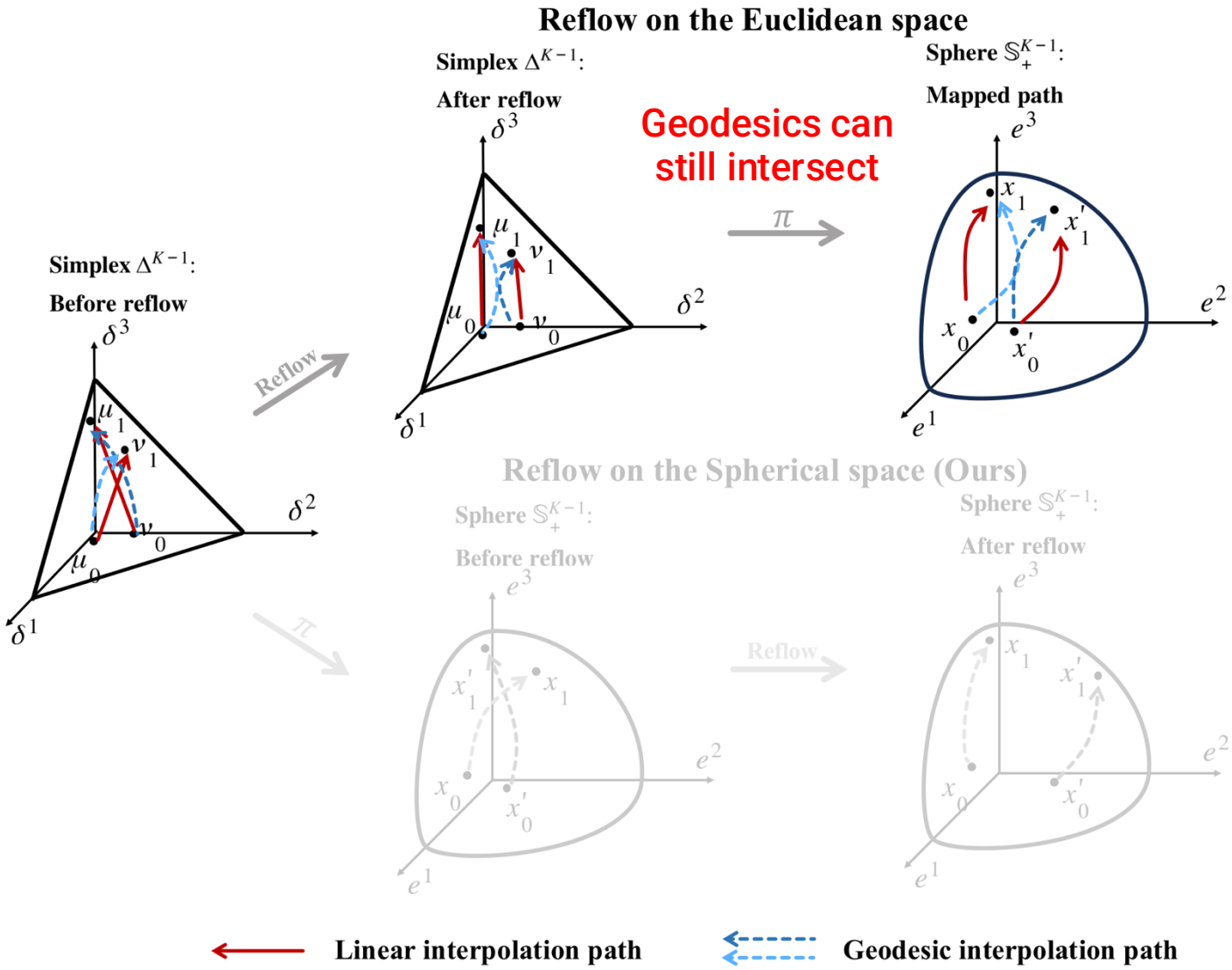
Rectified Flow[10]
Euclidean illustration



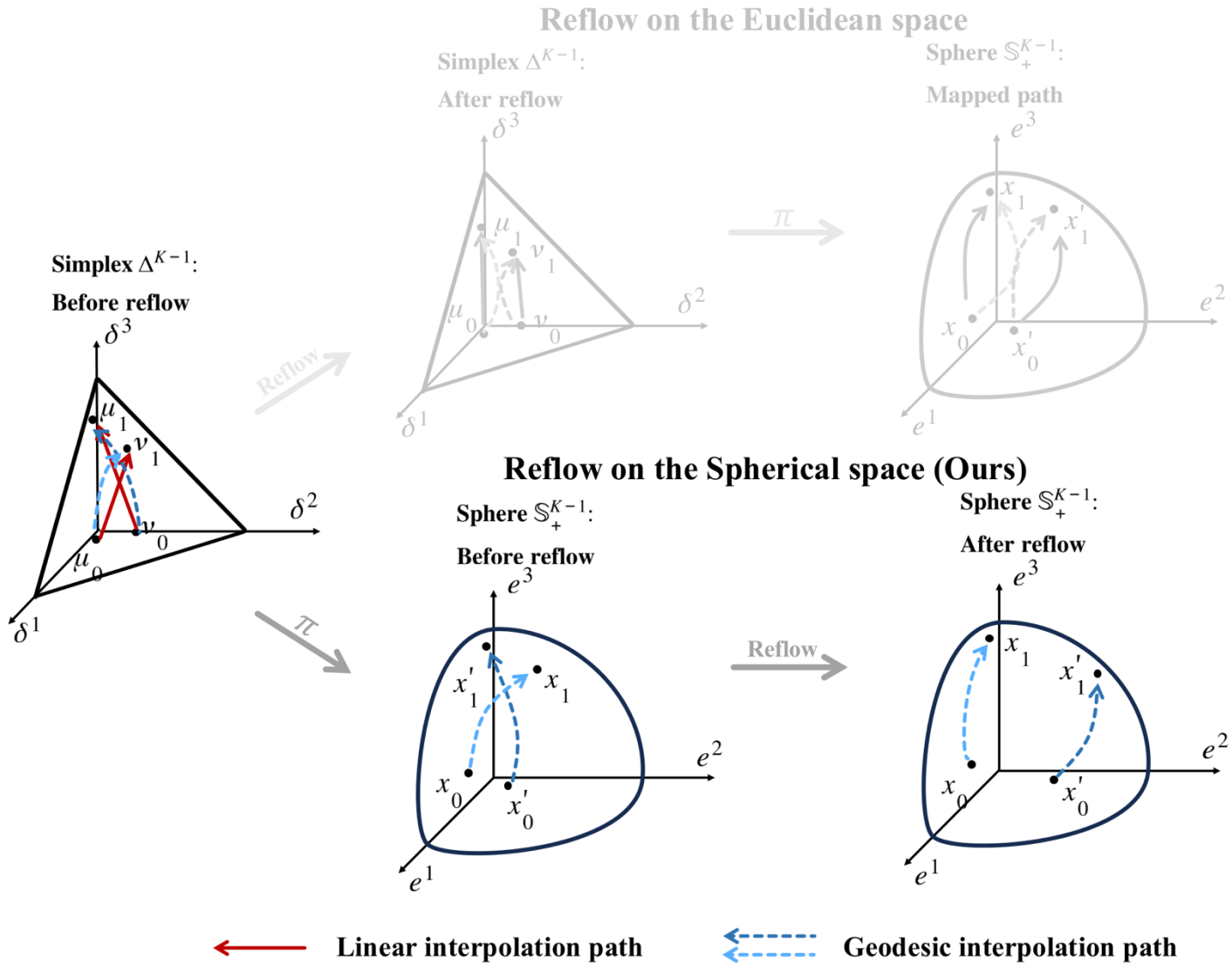
Independent pairs

[10] Liu et al. Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow. ICLR 2023.

Why Euclidean Reflow Fails



Spherical Reflow on Statistical Manifolds



Geodesic Interpolation on the Sphere

Fisher–Rao geodesics become great circles.

$$\mu \in \Delta^{K-1} \longrightarrow X = \pi(\mu) = (\sqrt{\mu_1}, \ldots, \sqrt{\mu_K})$$

$$X_t = \frac{\sin\big((1-t)\theta\big)}{\sin\theta} X_0 + \frac{\sin(t\theta)}{\sin\theta} X_1$$

$$\theta = \arccos(X_0^\top X_1)$$

Geodesic distance

Spherical Reflow: Recouple and Retrain

Recouple endpoints, then fit **geodesics.**

$$Z_1^1 = \mathrm{ODESolve}(v_{\hat{\theta}}^{(0)}, Z_0^1, 0 \to 1)$$

$$\mathcal{L}_{\mathrm{ReSFlow}}^{(1)} = \mathbb{E}\left[\left\| v_\theta^{(1)}(Z_t^1, t) - u_t^S(Z_t^1 \mid Z_0^1, Z_1^1) \right\|_2^2\right]$$

$$v_\theta^{\tan}(X_t, t) = v_\theta(X_t, t) - \langle v_\theta(X_t, t), X_t \rangle X_t$$

Tangent field **Normal component**

What Spherical Reflow Guarantees

Preserve marginals and reduce transport cost.

3.4 Theoretical Guarantees of Spherical Reflow

To ground our geometric intuition, we establish formal theoretical guarantees for the spherical reflow process based on the framework of non-Euclidean probability transport. Detailed proofs for all statements are provided in Appendix D.

First, a valid generative model must ensure that modifying the trajectories does not alter the target data distribution. We confirm that spherical reflow strictly preserves the marginal distributions.

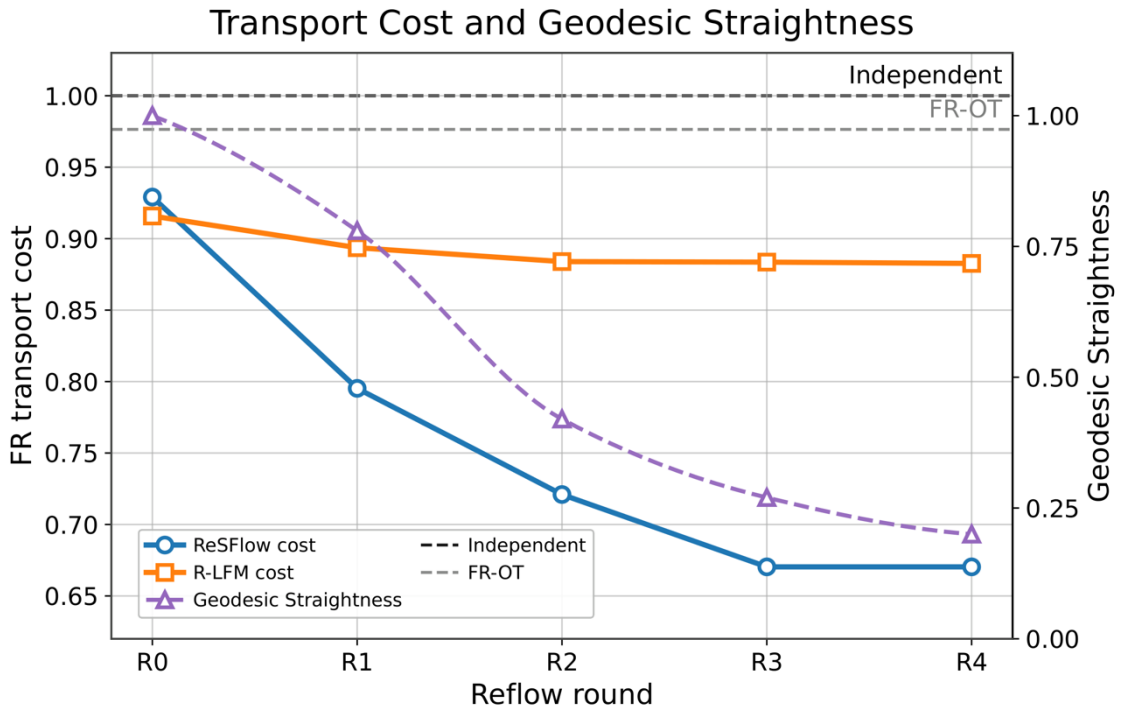
Theorem 3.1 (Spherical Marginal Preservation; Informal; See Theorem D.1)**.** *Let X_t be the spherical geodesic interpolation process. If the flow is rectifiable, its induced spherical rectified flow Z_t preserves the marginals:* $\mathrm{Law}(Z_t) = \mathrm{Law}(X_t)$ *for all* $t \in [0, 1]$.

To formalize trajectory straightness, we define the Spherical Transport Cost for a coupling (X_0, X_1) on $\mathbb{S}_+^{K-1}$ as $\mathbb{E}\big[d_{\mathbb{S}}(X_0, X_1)^2\big] = \mathbb{E}[\arccos(X_0^\top X_1)^2]$, which measures the expected squared geodesic distance. Reflow theoretically guarantees the optimization of this transport cost.

Theorem 3.2 (Contraction of Spherical Transport Cost; Informal; See Theorem D.2)**.** *Let (Z_0^k, Z_1^k) be the coupling induced by the k-th reflow iteration. The spherical transport cost is non-increasing:*

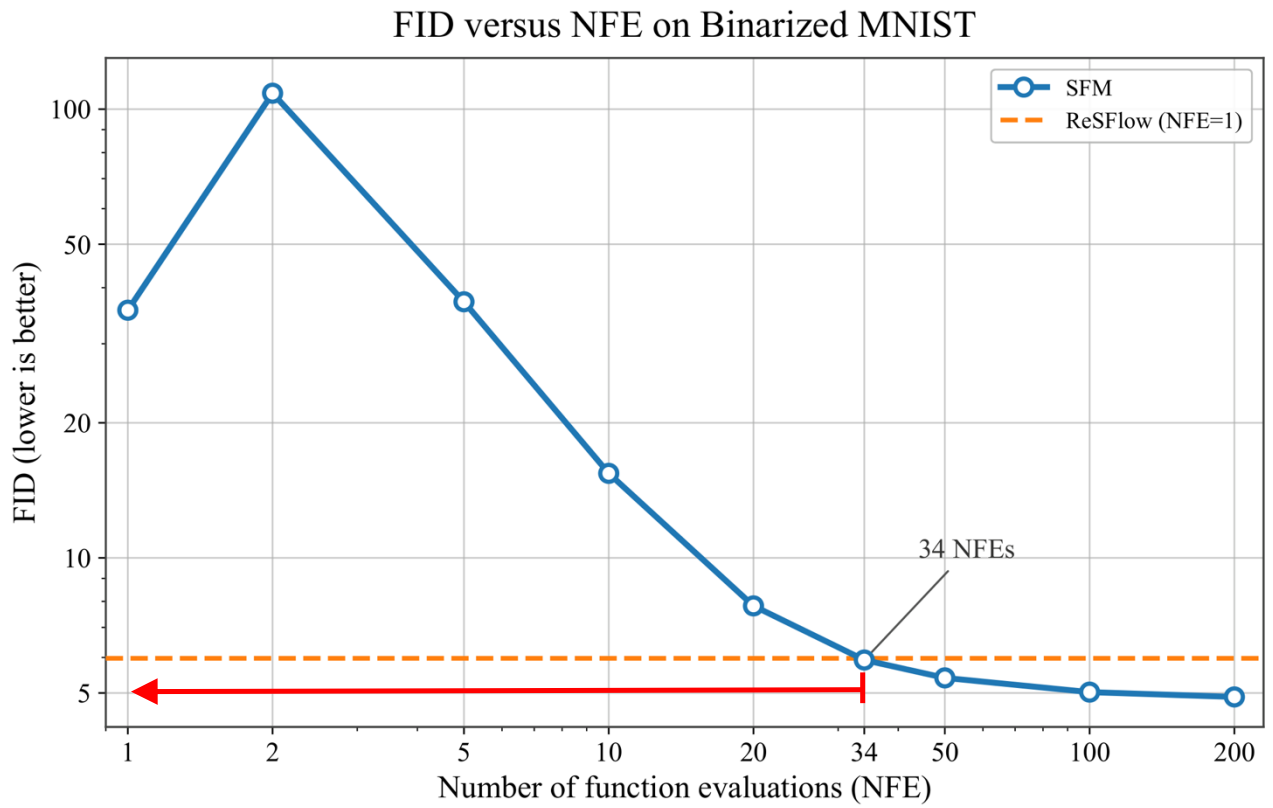
$$\mathbb{E}\big[d_{\mathbb{S}}(Z_0, Z_1)^2\big] \le \mathbb{E}\big[d_{\mathbb{S}}(X_0, X_1)^2\big]\,. \tag{12}$$

Proof. Proof sketch in Appendix D.3: The ODE trajectories have finite length, and the geodesic connecting the recoupled endpoints represents the absolute shortest path on the manifold, enforcing a monotonic decrease in transport cost. □



ReSFlow on Binarized MNIST

ReSFlow matches about SFM **34 steps** with **1 step.**



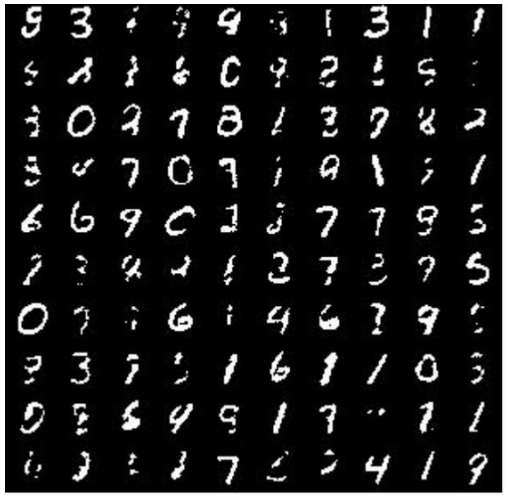
ReSFlow on Binarized MNIST

ReSFlow achieves the **best one-step FID** in this comparison.

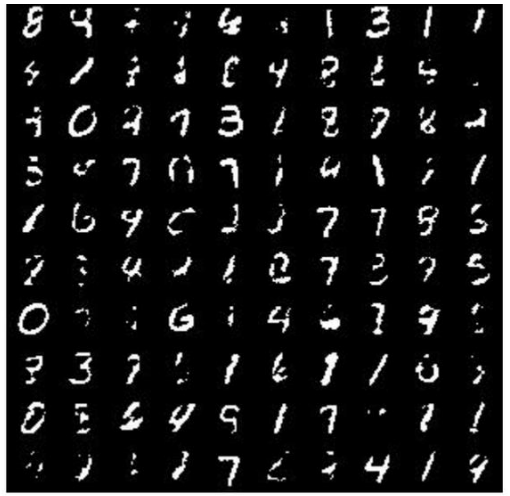
Method	ODE (NFE = adaptive)		Euler one-step (NFE = 1)	
	FID ↓	Time / 1k samples (s) ↓	FID ↓	Time / 1k samples (s) ↓
LinearFM [13]	$5.187 \pm 0.129^{*}$	87.13 ± 0.44	25.317 ± 0.517	$\mathbf{0.406 \pm 0.178}$
SFM [13]	$\mathbf{4.748 \pm 0.092}$	112.81 ± 8.20	35.640 ± 0.487	0.719 ± 0.175
R-LFM	5.681 ± 0.252	$61.19 \pm 0.28^{*}$	$6.261 \pm 0.379^{*}$	0.583 ± 0.244
ReSFlow	5.541 ± 0.252	$\mathbf{36.07 \pm 0.60}$	$\mathbf{5.973 \pm 0.442}$	$0.452 \pm 0.243^{*}$

ReSFlow on Binarized MNIST

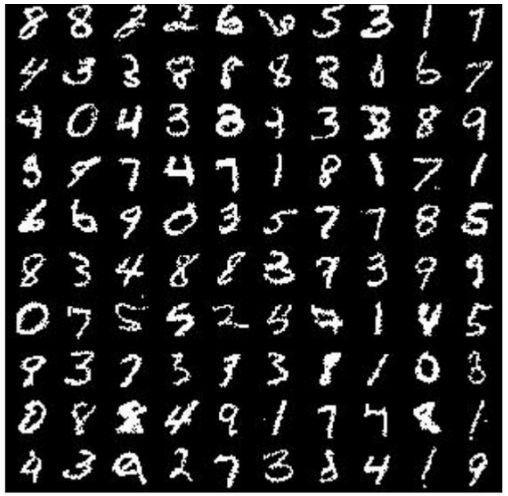
One-step samples from the four compared methods.



LinearFM | Euler



SFM | Euler



R-LFM | Euler



ReSFlow | Euler

ReSFlow on DNA Generation

One-step ReSFlow preserves **ODE-level SP-MSE.**

Method	ODE (NFE=adaptive)		Euler one-step (NFE=1)	
	SP-MSE ↓	Time ↓	SP-MSE ↓	Time ↓
LinearFM [13]	**0.027734 ± 0.000897**	119.04 ± 16.57	0.029449 ± 0.001673*	0.410 ± 0.004*
SFM [13]	0.028015 ± 0.000257*	82.18 ± 4.26	0.035564 ± 0.007768	0.410 ± 0.005
R-LFM	0.052146 ± 0.000781	**67.88 ± 4.17**	0.060504 ± 0.000351	**0.409 ± 0.001**
ReSFlow	0.028743 ± 0.000755	67.94 ± 0.35*	**0.028718 ± 0.000366**	**0.409 ± 0.003**

ReSFlow samples

ODE sampling

```
AAGAGCTAAGCCGCGGGCTCAGGTGCCCTCCCCCCAGCCC
CACAGCCCCGGGGAGCTCCAACGCCACCACCCCCAGGCGG
CAGCCCGGAACACGGGCGCCCCTGCCAGGAAGGATGCGAA
```

One-step Euler (NFE = 1)

```
AAGACCTAAGCCGCGGGCTCAGGTGCCCTCCCCCCAGCCC
CACAGCCCCGGGGAGCTCCAACGCCACCACCCCCAGGCGG
CAGCCCGGAACACGGGCGCCCCTGCCAGGAAGGATGCGAA
```

Paper, Appendix F.4, Table 6 · Exact sequence excerpts; full promoters contain 1,024 bases.

ReSFlow on DNA Generation

One-step sampling takes **0.409 seconds** for 1,000 DNA sequences.

Method	ODE (NFE=adaptive)		Euler one-step (NFE=1)	
	SP-MSE ↓	Time ↓	SP-MSE ↓	Time ↓
LinearFM [13]	**0.027734 ± 0.000897**	119.04 ± 16.57	0.029449 ± 0.001673*	0.410 ± 0.004*
SFM [13]	0.028015 ± 0.000257*	82.18 ± 4.26	0.035564 ± 0.007768	0.410 ± 0.005
R-LFM	0.052146 ± 0.000781	**67.88 ± 4.17**	0.060504 ± 0.000351	**0.409 ± 0.001**
ReSFlow	0.028743 ± 0.000755	67.94 ± 0.35*	**0.028718 ± 0.000366**	**0.409 ± 0.003**

166× faster

ReSFlow samples

ODE sampling

```
AAGAGCTAAGCCGCGGGCTCAGGTGCCCTCCCCCCAGCCC
CACAGCCCCGGGGAGCTCCAACGCCACCACCCCCAGGCGG
CAGCCCGGAACACGGGCGCCCCTGCCAGGAAGGATGCGAA
```

One-step Euler (NFE = 1)

```
AAGACCTAAGCCGCGGGCTCAGGTGCCCTCCCCCCAGCCC
CACAGCCCCGGGGAGCTCCAACGCCACCACCCCCAGGCGG
CAGCCCGGAACACGGGCGCCCCTGCCAGGAAGGATGCGAA
```

Paper, Appendix F.4, Table 6 · Exact sequence excerpts; full promoters contain 1,024 bases.

ReSFlow on Text8

ReSFlow improves ODE-evaluated BPC from **1.830 to 1.261.**

Method	BPC ↓ (adaptive ODE)
LinearFM [13]	2.15448 ± 0.00024
SFM [13]	$1.82998 \pm 0.00028^{*}$
R-LFM	3.26687 ± 0.00012
ReDi [11]	2.30562 ± 0.01010
ReSFlow	$\mathbf{1.26146 \pm 0.00036}$

ReSFlow · ODE sample

```
another_role_ranged_in_from_one_eight_two_
one_le_conseliation_of_louis_nweed_d_law_
in_trlits_quits_teberry_in_children_muslid_
born_herman_lempenn_one_eight_five_four_
whischeuse_leaders_of_several_differences_of_
privileges_sion_new_mans_to_the_war_knew_his_
```

ReSFlow · One-step sample (NFE = 1)

```
into_swy_no_and_they_are_also_itn_ha_is_also_
ttaught_una__fee_be_onf_the_popular_oo_is_
lireaand_for_exampard_altended_fr_a_aerenaen_
so_land_external_links_the_d_rib_ol_blt_
deer_opern_type_tad_seelanes_and_to_cannabse_
_k_vemuns_i__endite_fere_the_level_ard_
```

Paper, Appendix F.4, Table 5 · Original generated text; underscores denote spaces.

Conclusion

- **One-step** categorical generation benefits from **Fisher–Rao geometry.**
- **Spherical reflow** straightens geodesic trajectories.
- ReSFlow enables **one-step sampling** on BMNIST and DNA.

For more details, please join our poster session.

Hall 1-4 (Wed 07:00 CST)
Sydney Poster Session 3

arXiv

code