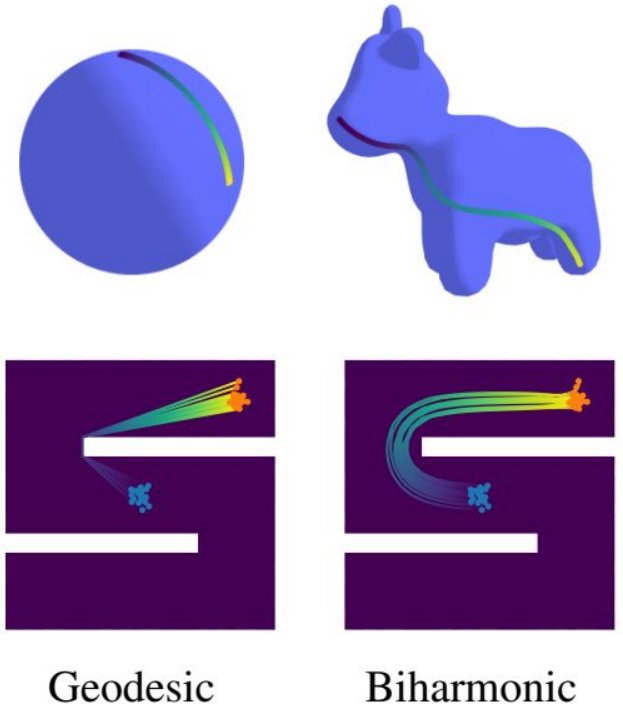
Flow Matching on General Geometries

Ricky T. Q. Chen and Yaron Lipman
ICLR 2024

Learning generative flows that respect geometry

Presented by: Zhixuan Zhou



Geodesic: shortest surface path

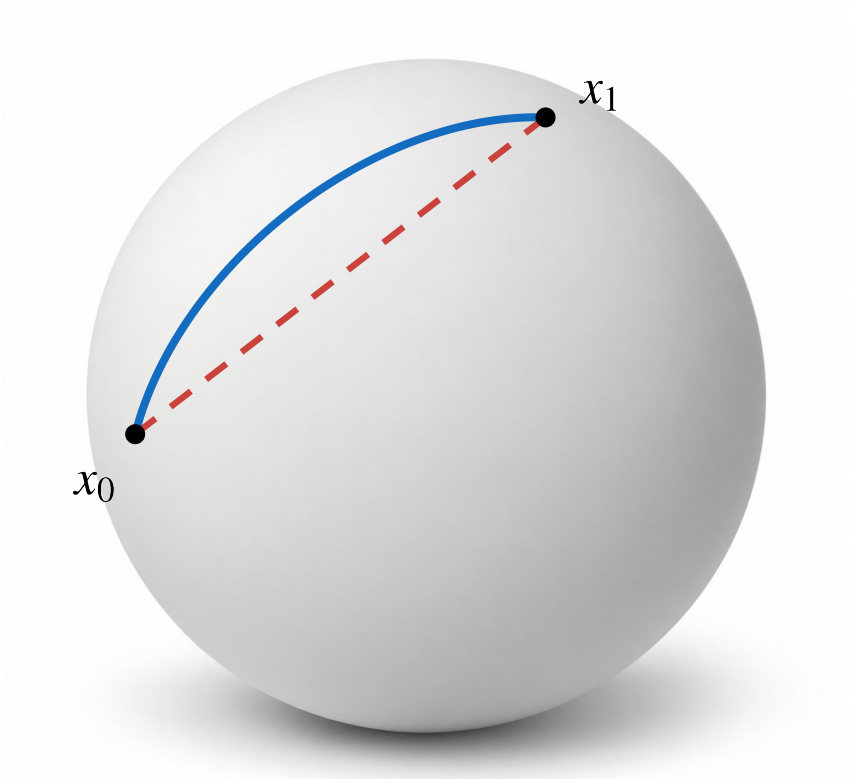
Biharmonic: a distance built from smooth patterns on the surface

Figure: Chen & Lipman (2024), Fig. 1. arxiv.org/abs/2302.03660

Motivation:

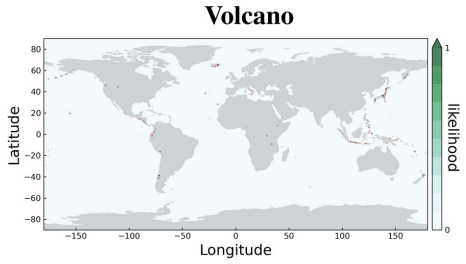
Non-Euclidean data need geometry-aware paths

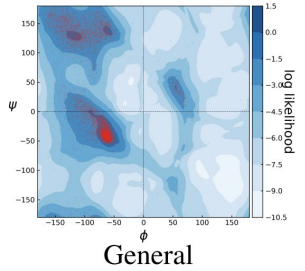
The path must respect the domain



Geodesic
Shortest route on the surface

Euclidean chord
Straight route leaves the surface







Earth locations on S^2

Geographic events lie on the Earth's surface.

Molecular rotation angles

Rotation around a bond: 0° and 360° coincide.

Density on a mesh

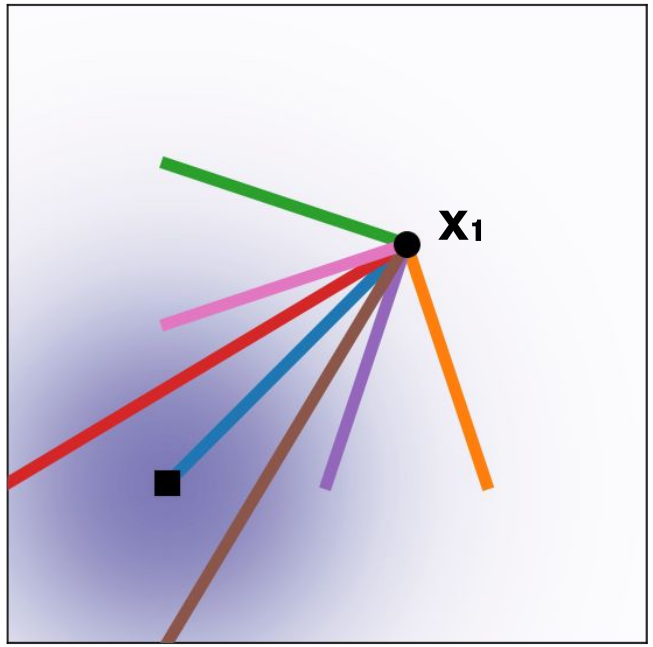
Curvature and topology constrain transport

Original sphere schematic, style inspired by De Bortoli et al. (2022), Fig. 1. Data: Chen & Lipman (2024), Figs. 4, 8, 9.

Euclidean Flow Matching:

Learn a velocity field by regression

Straight paths to a data point



Base density p(x_0)

Training: match a known velocity

Sample x_0 ~ p, x_1 ~ q, and t ~ Uniform(0,1)

p: noise distribution; q: data distribution; t: interpolation time.

$$x_t = (1-t)x_0 + tx_1$$ known position

$$\dot{x}_t = x_1 - x_0$$ known velocity

$$\mathcal{L}_{\mathrm{RCFM}}(\theta) = \mathbb{E}_{t,\,x_0,\,x_1}\|v_t(x_t) - (x_1 - x_0)\|^2$$

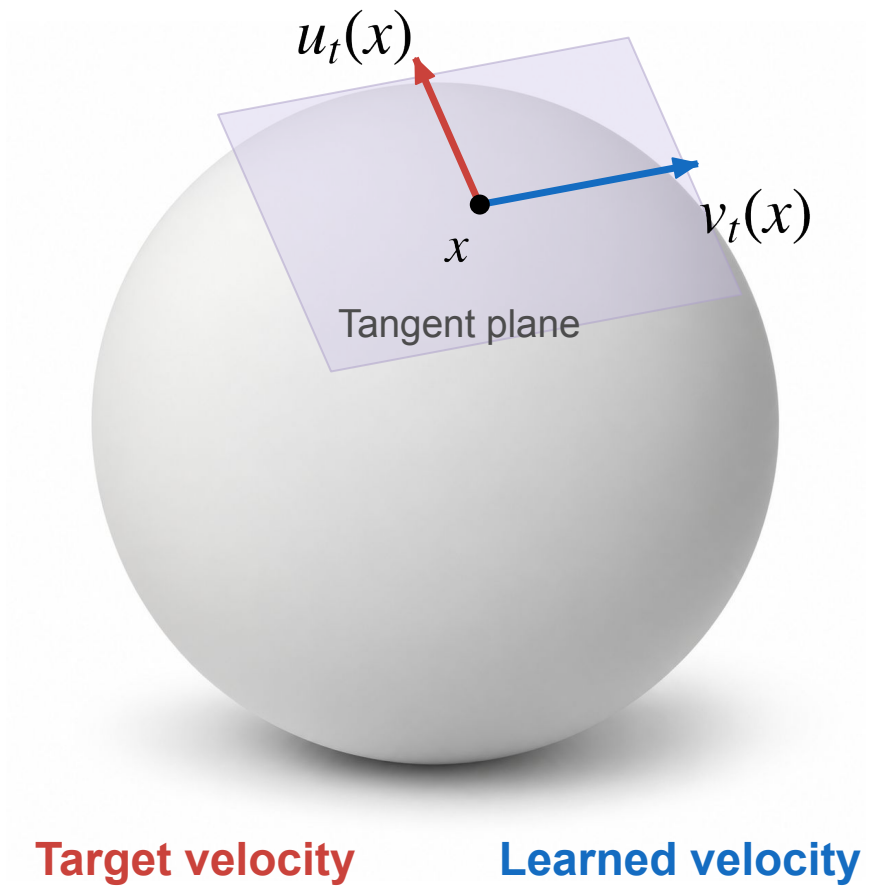
Generation: follow the learned field

$$\frac{dx_t}{dt} = v_t(x_t), \qquad x_0 \sim p$$

Straight-path example: Lipman et al. (2023), Fig. 3. Adapted notation: Chen & Lipman (2024), Sec. 3.1.

Riemannian Flow Matching:

Match velocities that respect the surface



What is Riemannian manifold?

Euclidean space is flat everywhere.
A manifold looks Euclidean locally, but may curve globally.
The metric g measures lengths and angles in each tangent space.

Valid states stay on the surface

Each point on the sphere is a valid state.

Valid velocities lie in the tangent plane

Target and learned velocities both point along the surface at the current position.

The RFM loss compares these velocities

$$\mathcal{L}_{\mathrm{RFM}}(\theta) = \mathbb{E}_{t,\, p_t(x)} \| v_t(x) - u_t(x) \|_g^2$$

Original schematic, style inspired by De Bortoli et al. (2022), Fig. 1. Notation and loss: Chen & Lipman (2024), Sec. 2, Eq. (3).

Why conditional flow matching works:

Conditional supervision gives the same gradient

RFM: marginal sampling and target velocity

$$\mathcal{L}_{\mathrm{RFM}}(\theta) = \mathbb{E}_{t,\, x_t \sim p_t} \left[\| v_\theta(t, x_t) - u_t(x_t) \|_g^2 \right]$$

The marginal target averages over possible destinations and is hard to compute.

RCFM: sample one data endpoint and use its conditional target

$$\mathcal{L}_{\mathrm{RCFM}}(\theta) = \mathbb{E}_{t,\, x_1 \sim q,\, x_t \sim p_t(\cdot \mid x_1)} \left[\| v_\theta(t, x_t) - u_t(x_t \mid x_1) \|_g^2 \right]$$

Same Gradient

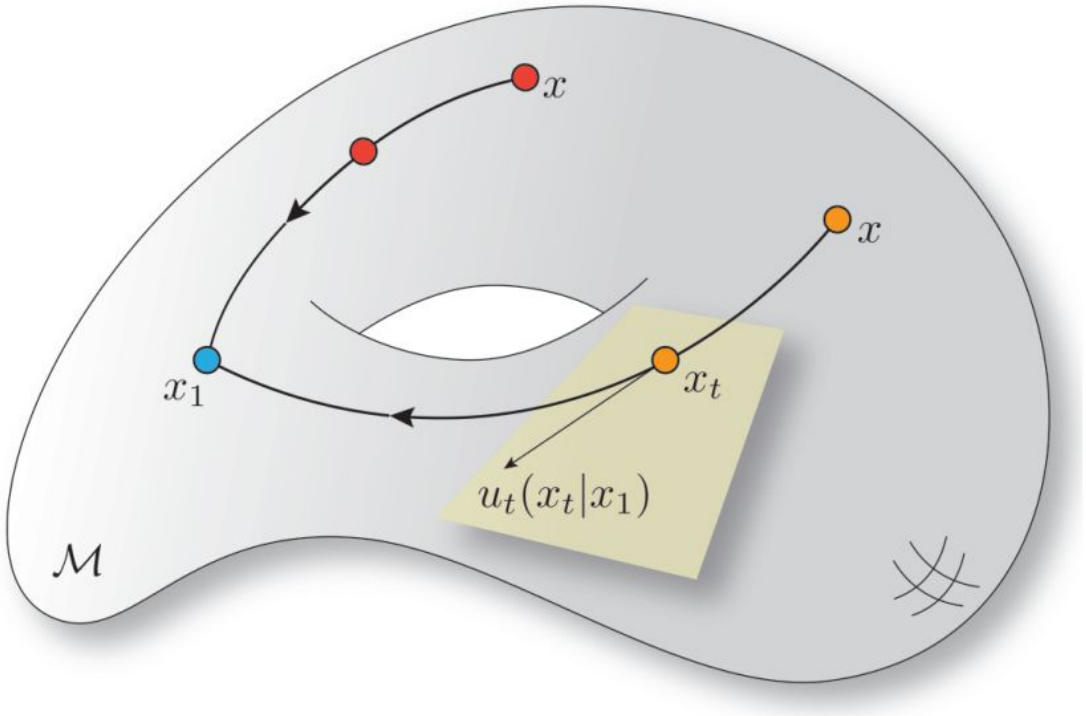
$$\nabla_\theta \mathcal{L}_{\mathrm{RFM}}(\theta) = \nabla_\theta \mathcal{L}_{\mathrm{RCFM}}(\theta)$$

Next: how do we construct a computable conditional target?

Chen & Lipman (2024), Sec. 3.1, Eqs. (3), (8), and Appendix A. Blue: marginal. Orange: conditional.

Constructing the conditional target:

Choose a rule for reaching each data point



Different starting points, one shared target.

Choose a function d(x, y)

x is the current point; y is the target.

1. **Never negative**
   The value can decrease toward zero.
2. **Zero only when x = y**
   Zero identifies the target point.
3. **A nonzero slope before arrival**
   There is a local direction that reduces d.

Make d shrink at a constant rate

$$d(x_t, x_1) = (1 - t)\, d(x_0, x_1)$$

t = 0: start. t = 1: arrival.

Chen & Lipman (2024), Fig. 2, Sec. 3.2 and Eq. (12).

From a path rule to a training target:

Turn the shrinking rule into a target velocity

1 Differentiate the constructed distance

The construction on the previous slide fixes the rate of change of d.

$$\frac{\mathrm{d}}{\mathrm{d}t} d(x_t, x_1) = -d(x_0, x_1)$$

2 Apply the chain rule

The chain rule gives an identity for the same rate, expressed using velocity.

$$\frac{\mathrm{d}}{\mathrm{d}t} d(x_t, x_1) = \langle \nabla d(x_t, x_1), u_t \rangle_g$$

u_t is the target velocity.

3 Combine (1) and (2), then minimize the norm

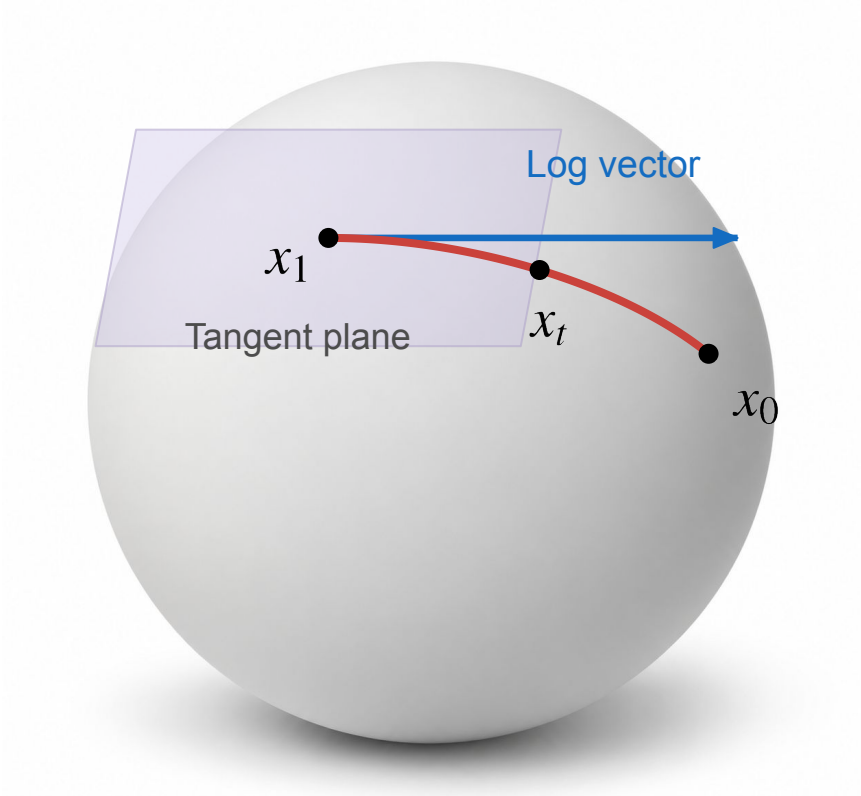
Equate the two rates, then minimize the velocity norm to obtain the expression below.

$$u_t(x_t \mid x_1) = -d(x_0, x_1) \frac{\nabla d(x_t, x_1)}{\|\nabla d(x_t, x_1)\|_g^2}$$

Chen & Lipman (2024), Eqs. (12)–(13), Theorem 3.1. Linear schedule: d decreases at a constant rate.

When geodesics are known:

Compute training points directly



Start at t = 0; reach the target at t = 1.

Geodesic interpolation

$$x_t = \exp_{x_1}((1-t)\log_{x_1}(x_0))$$

Log: manifold → tangent space

Exp: tangent space → manifold

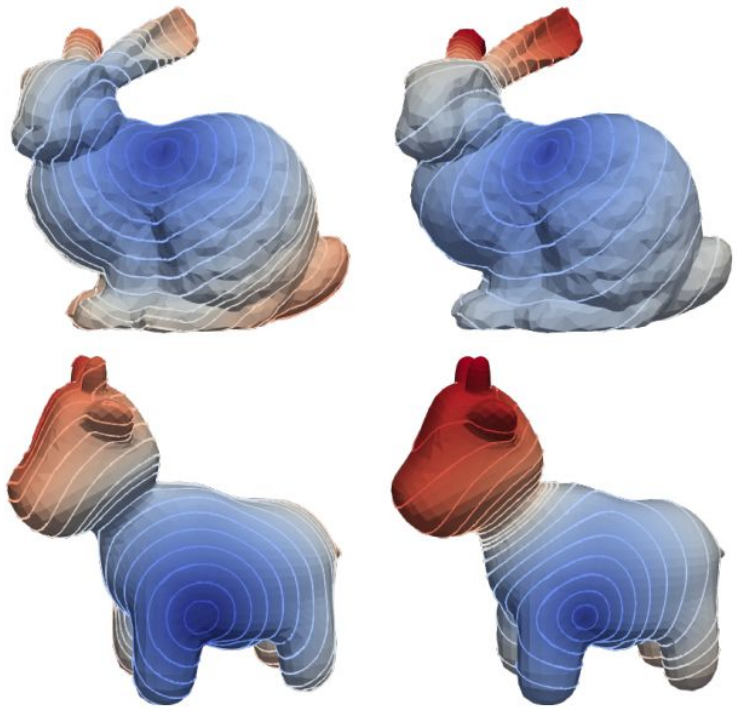
This gives the training position and velocity directly.

Chen & Lipman (2024), Eq. (15), Table 5. Diagram: geodesic interpolation using the target as reference.

Complex surfaces:

Build a distance from surface patterns

On meshes, shortest paths are costly to compute.



Colors and contours show distance to a fixed point on the surface.

1 Precompute surface patterns once

Each $\varphi_i(x)$ is a pattern's value at point x.

2 Compare values: biharmonic distance

$$d(x,y)^2 = \sum_{i=1}^{k} \frac{(\varphi_i(x) - \varphi_i(y))^2}{\lambda_i^2}$$

Keep k patterns; weight coarse patterns more strongly.
λ_i is the eigenvalue associated with pattern i.

3 Use this distance in the target field (P7)

Use its gradient; solve the conditional ODE.

Chen & Lipman (2024), Sec. 3.3, Eq. (16), Fig. 3 (first two columns). Conditional target: Eq. (13).

Method wrap-up:

Euclidean and Riemannian Flow Matching

Ingredient	Euclidean FM	Riemannian FM
State space	$\mathbb{R}^n$	$(\mathcal{M}, g)$
Premetric	$d(x, y) = \|x - y\|_2$	$d_g(x, y)$ or $d_w(x, y)$ Geodesic or spectral distance
Conditional path	$x_t = (1 - t)x_0 + tx_1$	$x_t = \psi_t(x_0 \mid x_1)$ $d(x_t, x_1) = (1 - t)d(x_0, x_1)$
Target velocity	$u_t(x_t \mid x_1) = x_1 - x_0$	$u_t(x_t \mid x_1) = -\frac{d(x_t, x_1)}{1 - t} \frac{\nabla_g d(x_t, x_1)}{\|\nabla_g d(x_t, x_1)\|_g^2}$
Conditional loss	$\mathbb{E}_{t, x_0, x_1} \|v_t(x_t) - \dot{x}_t\|_2^2$	$\mathbb{E}_{t, x_0, x_1} \|v_t(x_t) - \dot{x}_t\|_g^2$

$\kappa(t) = 1 - t$; $x_0 \sim p$, $x_1 \sim q$, $t \sim \text{Uniform}(0, 1)$.

Chen & Lipman (2024), Eqs. (10), (12)–(13), (15)–(16). The Euclidean column uses the straight-path special case.

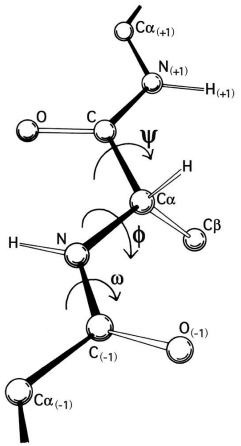
Experimental setup:

Likelihood, scaling, and geometric constraints

Data family	Geometry
Earth / climate locations	Sphere S^2
Protein and RNA angles	Tori T^2 and T^7
Wrapped Gaussian	High-dimensional tori
Bunny, cow, and mazes	Triangular meshes

$T^n = (S^1)^n$

n periodic angles; each S^1 is a circle.
0° and 360° describe the same angle.



Protein: two angles

Torsion angle = rotation around a bond.

80 / 10 / 10 split

Train / validation / test

5 random seeds

Report mean ± standard deviation

Test NLL

Lower is better

Chen & Lipman (2024), Sec. 5, App. H/J. Protein diagram: Wikimedia Commons, “Protein backbone PhiPsiOmega drawing”.

Results on Earth data:

RFM leads on Volcano and Fire

Test negative log-likelihood on the sphere. Lower is better.

Method	Volcano	Earthquake	Flood	Fire
Riemannian CNF	−6.05 ± 0.61	0.14 ± 0.23	1.11 ± 0.19	−0.80 ± 0.54
Moser Flow	−4.21 ± 0.17	−0.16 ± 0.06	0.57 ± 0.10	−1.28 ± 0.05
CNF Matching	−2.38 ± 0.17	−0.38 ± 0.01	**0.25 ± 0.02**	−1.40 ± 0.02
Riemannian Score-Based	−4.92 ± 0.25	−0.19 ± 0.07	0.48 ± 0.17	−1.33 ± 0.06
Riemannian Diffusion*	−6.61 ± 0.96	**−0.40 ± 0.05**	0.43 ± 0.07	−1.38 ± 0.05
RFM (geodesic)	**−7.93 ± 1.67**	−0.28 ± 0.08	0.42 ± 0.05	**−1.86 ± 0.11**

Best reported means on Volcano and Fire

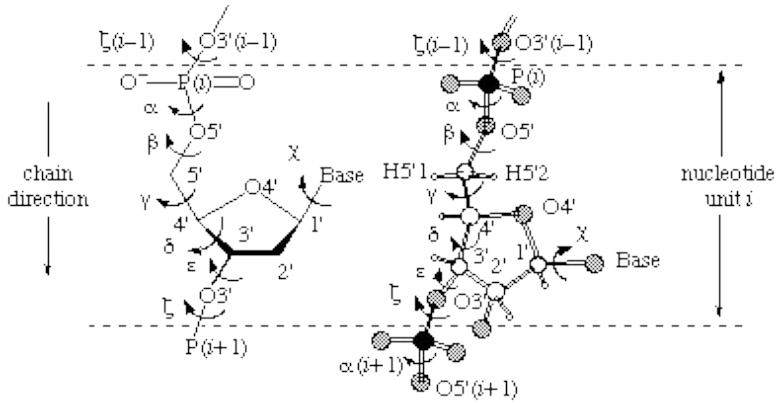
Data: Chen & Lipman (2024), Table 2. Values reproduced as mean ± SD over 5 runs.

Results on tori:

Likelihood and dimensional scaling

RNA: seven torsion angles on T^7

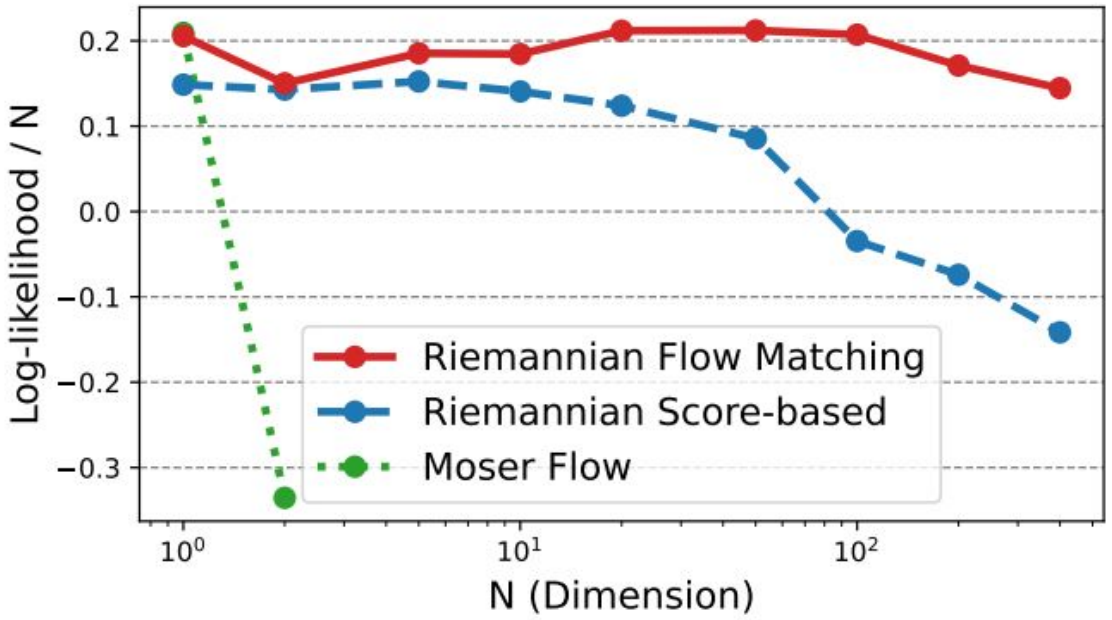
α–ζ: six backbone angles | χ (chi): base–sugar rotation



RNA: 1.50 lower mean NLL

Method	Test NLL
Riemannian Diffusion	−3.70 ± 0.592
RFM	−5.20 ± 0.067

Synthetic high-dimensional tori



Log-likelihood per dimension in bits. Higher is better.

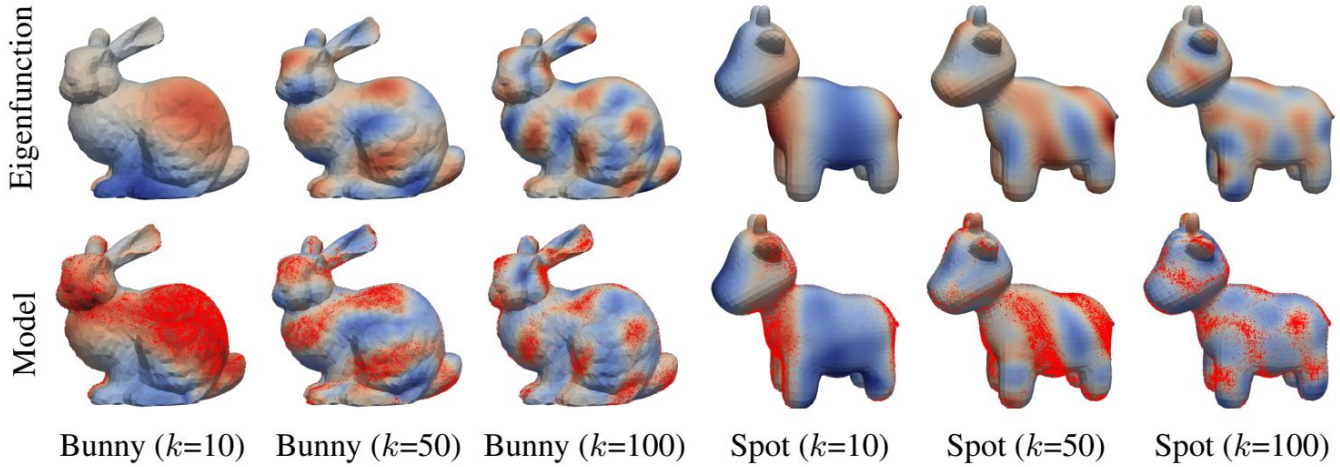
Protein (T^2): RFM has lower mean NLL on 3 of 4 subsets; diffusion leads on Proline.

Chen & Lipman (2024), Table 3 / Fig. 5. RNA diagram: IUPAC–IUB (1982), Fig. 4, polynucleotide conformations.

Results on meshes and mazes:

Complex surfaces and boundaries

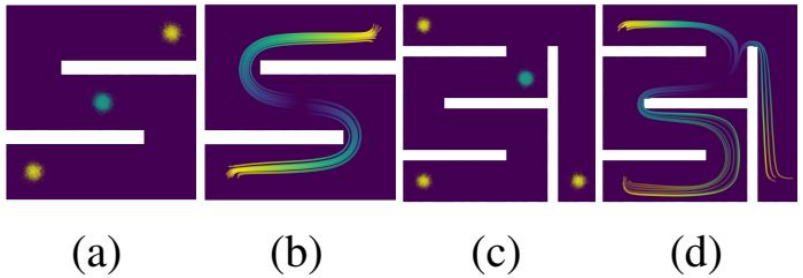
Closed surfaces



Biharmonic premetric with k = 200 modes

Figure labels k = 10, 50, 100 identify target eigenfunctions

Boundaries



k = 30 modes
Neumann boundary conditions

Sample trajectories follow corridors

Figures reproduced from Chen & Lipman (2024), Figs. 4 and 6. Setup: Section 5 and Appendix H.

Critical assessment:

Strengths and limitations

Strengths

More scalable training

No simulation with tractable geodesics; no training-time divergence (Table 1)

Flexible flow design

Premetrics define conditional flows beyond geodesic interpolation

Limitations

Simulation-free has a limited scope

General geometries: spectral preprocessing and conditional ODE solves

Theory depends on the premetric

Finite modes, critical points, and solver error

Assessment: Chen & Lipman (2024), Table 1, Fig. 5, Secs. 3–5 and Appendices E, G, H.

THANK YOU